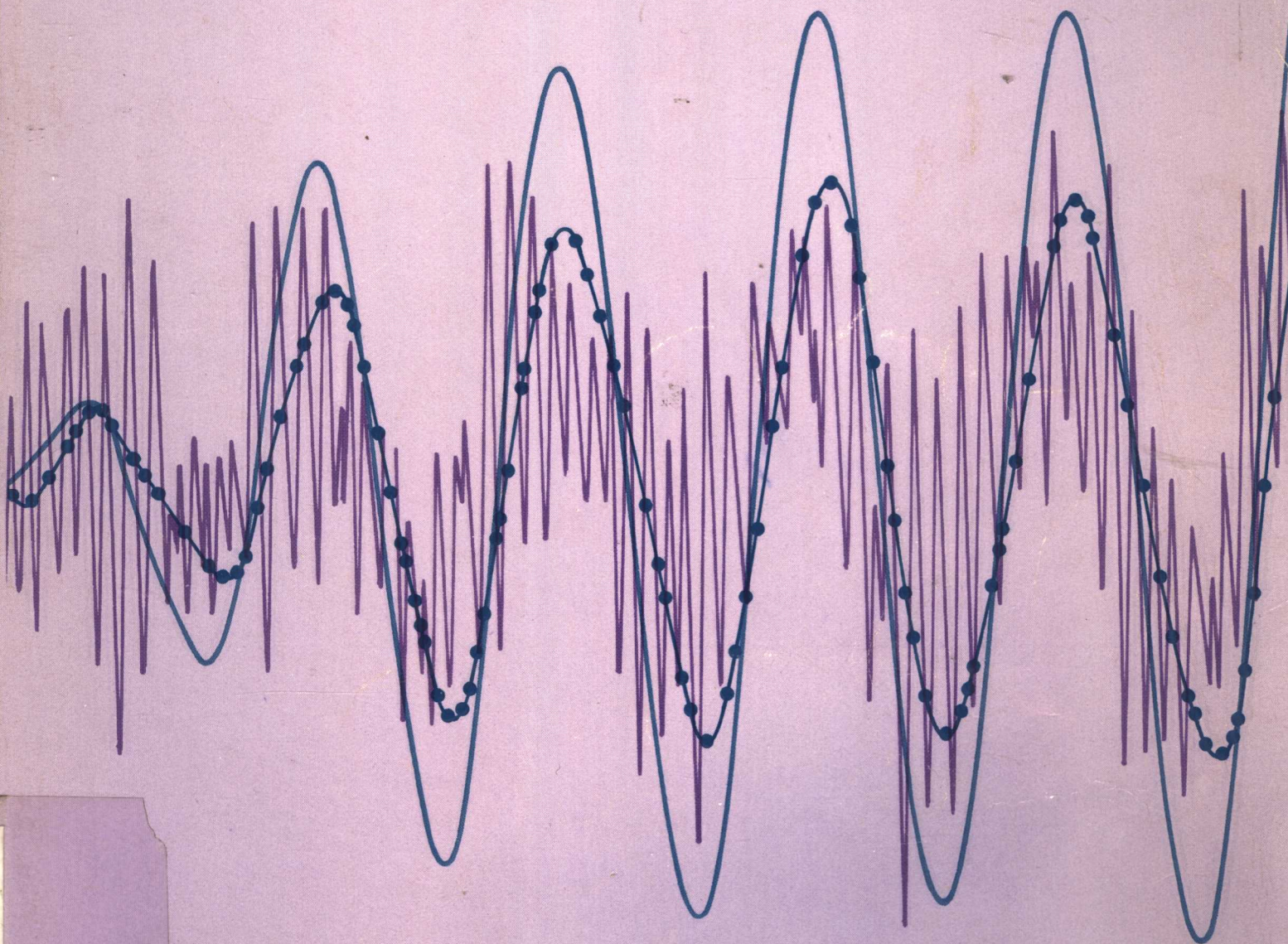


15784

Kalman Filtering: Theory and Application

Edited by
Harold W. Sorenson



**IEEE
PRESS**

A volume in the IEEE PRESS Selected Reprint
Series, prepared under the sponsorship of the
IEEE Control Systems Society.

Kalman Filtering: Theory and Application

Edited by

Harold W. Sorenson

Professor of Engineering Sciences

Department of Applied Mechanics and Engineering Sciences

University of California, San Diego

**A volume in the IEEE PRESS Selected Reprint
Series, prepared under the sponsorship of the
IEEE Control Systems Society.**



IEEE

ESS

**73.7627
K14**

Electrical and Electronics Engineers, Inc., New York

IEEE PRESS

1985 Editorial Board

M. E. Van Valkenburg, *Editor in Chief*

M. G. Morgan, *Editor, Selected Reprint Series*

Glen Wade, *Editor, Special Issue Series*

J. M. Aein	Thelma Estrin	J. O. Limb
J. K. Aggarwal	L. H. Fink	R. W. Lucky
James Aylor	S. K. Ghandhi	E. A. Marcatili
J. E. Brittain	Irwin Gray	J. S. Meditch
R. W. Brodersen	H. A. Haus	W. R. Perkins
B. D. Carroll	E. W. Herold	A. C. Schell
R. F. Cotellessa	R. C. Jaeger	Herbert Sherman
M. S. Dresselhaus		D. L. Vines

W. R. Crone, *Managing Editor*

Hans P. Leander, *Technical Editor*

Teresa Abiuso, *Administrative Assistant*

David G. Boulanger, *Associate Editor*

Copyright © 1985 by
THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.
345 East 47th Street, New York, NY 10017-2394
All rights reserved.

PRINTED IN THE UNITED STATES OF AMERICA

IEEE Order Number: PP03160

Library of Congress Cataloging in Publication Data

Main entry under title:

Kalman filtering.

(IEEE Press selected reprint series)

Pt. 1 consists of papers published in the 1960s;
pt. 2 originally appeared as a special issue for Mar.
1983 of IEEE transactions on automatic control.

Bibliography: p.

Includes indexes.

1. Kalman filtering. I. Sorenson, H. W. (Harold
Wayne), 1936- .

QA402.3.K28 1985 003 85-14253

ISBN 0-87942-191-6

ISBN 0-7803-0421-7 {pbk}

Contents

Introduction	1
Part I: Theoretical Beginnings	5
Section I-A: Historical Survey	6
Least-Squares Estimation: From Gauss to Kalman, <i>H. W. Sorenson (IEEE Spectrum, July 1970)</i>	7
Section I-B: Theoretical Foundations	13
A New Approach to Linear Filtering and Prediction Problems, <i>R. E. Kalman (Transactions of the ASME, Journal of Basic Engineering, March 1960)</i>	16
First Order Error Propagation in a Stagewise Smoothing Procedure for Satellite Observations, <i>P. Swerling (The Journal of the Astronautical Sciences, Autumn 1959)</i>	27
New Results in Linear Filtering and Prediction Theory, <i>R. E. Kalman and R. S. Bucy (Transactions of the ASME, Journal of Basic Engineering, March 1961)</i>	34
A Bayesian Approach to Problems in Stochastic Estimation and Control, <i>Y. C. Ho and R. C. K. Lee (IEEE Transactions on Automatic Control, October 1964)</i>	48
An Innovations Approach to Least-Squares Estimation—Part I: Linear Filtering in Additive White Noise, <i>T. Kailath (IEEE Transactions on Automatic Control, December 1968)</i>	55
Section I-C: General Application Considerations	65
A Statistical Optimizing Navigation Procedure for Space Flight, <i>R. H. Battin (ARS Journal, November 1962)</i>	67
Linear Regression Applied to System Identification for Adaptive Control Systems, <i>R. E. Kopp and R. J. Orford (AIAA Journal, October 1963)</i>	83
Kalman Filtering Techniques, <i>H. W. Sorenson (Advances in Control Systems Theory and Applications, Academic Press, 1966)</i>	90
Section I-D: Model Errors and Divergence	127
On A Priori Statistics in Minimum-Variance Estimation Problems, <i>T. T. Soong (Transactions of the ASME, Journal of Basic Engineering, March 1965)</i>	129
The Effect of Erroneous Models on the Kalman Filter Response, <i>H. Heffes (IEEE Transactions on Automatic Control, July 1966)</i>	133
On the A Priori Information in Sequential Estimation Problems, <i>T. Nishimura (IEEE Transactions on Automatic Control, April 1966)</i>	136
✓ Divergence in the Kalman Filter, <i>F. H. Schlee, C. J. Standish, and N. F. Toda (AIAA Journal, June 1967)</i>	144
✓ Divergence of the Kalman Filter, <i>R. J. Fitzgerald (IEEE Transactions on Automatic Control, December 1971)</i>	151
Section I-E: Divergence Control and Adaptive Filtering	163
Adaptive Filtering, <i>A. H. Jazwinski (Automatica, July 1969)</i>	164
On the Identification of Variances and Adaptive Kalman Filtering, <i>R. K. Mehra (IEEE Transactions on Automatic Control, April 1970)</i>	175
Recursive Fading Memory Filtering, <i>H. W. Sorenson and J. E. Sacks (Information Sciences, January 1971)</i>	185
Section I-F: Computational Considerations	195
Estimation Using Sampled Data Containing Sequentially Correlated Noise, <i>A. E. Bryson, Jr., and L. J. Henrikson (Journal of Spacecraft and Rockets, June 1968)</i>	197
Treatment of Bias in Recursive Filtering, <i>B. Friedland (IEEE Transactions on Automatic Control, August 1969)</i>	201
✓ Discrete Square Root Filtering: A Survey of Current Techniques, <i>P. G. Kaminski, A. E. Bryson, Jr., and S. F. Schmidt (IEEE Transactions on Automatic Control, December 1971)</i>	210
Computational Requirements for a Discrete Kalman Filter, <i>J. M. Mendel (IEEE Transactions on Automatic Control, December 1971)</i>	219
Section I-G: Smoothing	230
A Survey of Data Smoothing for Linear and Nonlinear Dynamic Systems, <i>J. S. Meditch (Automatica, March 1973)</i>	231
Maximum Likelihood Estimates of Linear Dynamic Systems, <i>H. E. Rauch, F. Tung, and C. T. Striebel (AIAA Journal, August 1965)</i>	243

The Optimum Linear Smoother as a Combination of Two Optimum Linear Filters, <i>D. C. Fraser and J. E. Potter (IEEE Transactions on Automatic Control, August 1969)</i>	249
An Innovations Approach to Least-Squares Estimation—Part II: Linear Smoothing in Additive White Noise, <i>T. Kailath and P. Frost (IEEE Transactions on Automatic Control, December 1968)</i>	252
Part II: Application Papers	259
Editorial—On Applications of Kalman Filtering, <i>H. W. Sorenson (IEEE Transactions on Automatic Control, March 1983)</i>	261
Voyager Orbit Determination at Jupiter, <i>J. K. Campbell, S. P. Synnott, and G. J. Bierman (IEEE Transactions on Automatic Control, March 1983)</i>	263
Decoupled Kalman Filters for Phased Array Radar Tracking, <i>F. E. Daum and R. J. Fitzgerald (IEEE Transactions on Automatic Control, March 1983)</i>	276
Utilization of Modified Polar Coordinates for Bearings-Only Tracking, <i>V. J. Aidala and S. E. Hammel (IEEE Transactions on Automatic Control, March 1983)</i>	291
Estimation and Prediction for Maneuvering Target Trajectories, <i>R. F. Berg (IEEE Transactions on Automatic Control, March 1983)</i>	303
Multiconfiguration Kalman Filter Design for High-Performance GPS Navigation, <i>M. H. Kao and D. H. Eller (IEEE Transactions on Automatic Control, March 1983)</i>	314
Nonlinear Kalman Filtering Techniques for Terrain-Aided Navigation, <i>L. D. Hostetler and R. D. Andreas (IEEE Transactions on Automatic Control, March 1983)</i>	325
Application of Multiple Model Estimation to a Recursive Terrain Height Correlation System, <i>G. L. Mealy and W. Tang (IEEE Transactions on Automatic Control, March 1983)</i>	334
Design and Analysis of a Dynamic Positioning System Based on Kalman Filtering and Optimal Control, <i>S. Saelid, N. A. Jenssen, and J. G. Balchen (IEEE Transactions on Automatic Control, March 1983)</i> ...	343
Dynamic Ship Positioning Using a Self-Tuning Kalman Filter, <i>P. T.-K. Fung and M. J. Grumble (IEEE Transactions on Automatic Control, March 1983)</i>	351
On the Feasibility of Real-Time Prediction of Aircraft Carrier Motion at Sea, <i>M. M. Sidar and B. F. Doolin (IEEE Transactions on Automatic Control, March 1983)</i>	362
An Integrated Multisensor Aircraft Track Recovery System for Remote Sensing, <i>W. S. Gesing and D. B. Reid (IEEE Transactions on Automatic Control, March 1983)</i>	368
Bathymetric and Oceanographic Applications of Kalman Filtering Techniques, <i>R. F. Brammer, R. P. Pass, and J. V. White (IEEE Transactions on Automatic Control, March 1983)</i>	376
A Kalman Filtering Approach to Natural Gamma Ray Spectroscopy in Well Logging, <i>G. Ruckebusch (IEEE Transactions on Automatic Control, March 1983)</i>	385
Measurement of Instantaneous Flow Rate Through Estimation of Velocity Profiles, <i>M. Uchiyama and K. Hakomori (IEEE Transactions on Automatic Control, March 1983)</i>	394
Estimation and Prediction of Unmeasurable Variables in the Steel Mill Soaking Pit Control System, <i>V. J. Lumelsky (IEEE Transactions on Automatic Control, March 1983)</i>	403
Application of Kalman Filters to the Regulation of Dead Time Processes, <i>W. L. Bialkowski (IEEE Transactions on Automatic Control, March 1983)</i>	416
On-Line Failure Detection in Nuclear Power Plant Instrumentation, <i>J. L. Tylee (IEEE Transactions on Automatic Control, March 1983)</i>	422
The Application of Kalman Filtering Estimation Techniques in Power Station Control Systems, <i>J. N. Wallace and R. Clarke (IEEE Transactions on Automatic Control, March 1983)</i>	432
Application of Kalman Filtering to Demographic Models, <i>B. G. Leibundgut, A. Rault, and F. Gendreau (IEEE Transactions on Automatic Control, March 1983)</i>	444
Author Index	452
Subject Index	453
Editor's Biography	457

Introduction

THE Kalman Filter! The term evokes many and varied responses among engineers, scientists, and managers who hear it. For some, it is a buzzword that must be used in a proposal when attempting to obtain contractual support for a new effort; for others, it is an answer in search of a problem as they already know of better ways to solve their specific problem; for some, it represents a practical set of procedures that they can use to process numerical data to obtain estimates of parameters and variables whose values are uncertain. This book is addressed to those individuals who know of the Kalman filter, accept that it may provide a computational means to solve their numerical data processing problem, and want to learn more about it.

It is probably not an overstatement to assert that the Kalman filter represents the most widely applied and demonstrably useful result to emerge from the state variable approach of "modern control theory." The application of the basic algorithm and the evolution of the body of algorithms that currently can be said to represent the "Kalman Filter" began remarkably soon after the original publication of the recursive, computational procedure proposed by Kalman and others. The earliest applications were stimulated by the "leap" into space that was beginning in the early sixties. The first proposed application of this revolutionary approach was for many (certainly this author) a NASA report published by Gerald Smith, Stanley Schmidt, and Leonard McGee in 1962 [1]. In the few years following the emergence of this report, there was a flurry of substantial activity, primarily in the aerospace industry, that provided the basis for the Kalman filter now so widely applied.

This book is divided into two, distinct parts. Part I presents the germinal papers appearing during the 1960's that provided serious consideration of the theoretical characteristics of the algorithm and the problems that are encountered in attempting to apply it to practical problems. Part II presents a broad coverage of applications that have been accomplished in the first years of the 1980's. These applications appeared in a Special Issue of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL in March 1983 which is reprinted here in its entirety. The breadth of the application topics speaks clearly to the utility of the method and illustrates the distance that the methodology has come since its initial appearance and proposed applications.

The application papers in Part II are supplemented in Part I by a sampling of the theoretical papers that served to shape the development of the methodology. These papers have not been included as a reminiscence about the *good old days* of the field. Many textbooks have been published that review and synthesize the

results presented in research papers that contributed to the general development. A bibliography listing many of the textbooks appears at the end of this Introduction that should provide the reader with a variety of careful and comprehensive treatments of the subject. While teaching graduate classes for several years and using several of the books as prescribed texts or references, I have been impressed by the positive response of students to the requirement that they read the early papers in the field as a means of understanding more clearly the concerns that stimulated the results that now are presented so logically in the textbooks. This pedantic observation has served as the motivation for the structure of this volume. By making the early papers conveniently available, it is hoped that the reader will be encouraged to peruse some of these papers and to apply the advantage provided to them by the passage of time to recapture the early excitement and the insights that are still valid.

The theoretical papers are organized into several subject areas. The material was organized in this manner with the intention of defining the major ingredients that must be considered in accomplishing a successful application. In a sense, the topics chosen are sufficiently independent that they can serve as a set of coordinate vectors for the consideration of the area. The volume is launched by an historical survey Section I-A that appeared in the *IEEE Spectrum*. This paper, "Least-squares estimation: from Gauss to Kalman" provides a review of the field as it had developed by the time of its appearance. Developments during the one and one-half decades since its creation suggest changes in detail but not of substance. Perhaps, this may be an indication of the rapid development and maturation of the basic methodology.

The first major group of five papers comprises Section I-B of Part I. These papers deal with the fundamental development of the algorithm and its interpretation from important and useful points-of-view. The papers in Section I-C deal with concerns that arise as one attempts to apply the algorithm to a problem. In particular, these papers discuss the manner in which the Kalman filter algorithm, derived by considering linear, dynamic systems, can be applied to nonlinear systems. From these and the other papers referenced in this section arose the *Extended Kalman Filter* which is, really, the subject of this book.

Given the algorithm and its earliest extensions to nonlinear problems, numerical simulation studies pointed to the dismaying conclusion that the estimates and associated covariances produced by the EKF did not necessarily bear a strong or discernible relationship to the actual behavior of the system and the estimation

error. The term *divergence* was invented to define the problem and launched the type of investigation that any developer of an EKF for a specific problem must consider and overcome. Early papers discussing divergence and the effect of model errors that cause it are found in Section I-D.

The methods of divergence control are almost as varied as the imaginations of the people who have addressed the problem. Generally, the procedures that have been proposed are heuristic and *ad hoc*. The procedures that appear to be based on more substantial, theoretical grounds impose assumptions and conditions that permit analysis but which are seldom satisfied in practice. Some of the major approaches that were developed initially are presented in Section I-E. In addition, references containing other approaches and their variations are provided.

Putting aside concerns raised by modeling and approximation errors, the implementation of the EKF algorithm can imply a computational burden, not only in terms of its cycle-time and storage requirements, but also in terms of the effects of numerical errors and their cumulative effects as the recursion proceeds. Assessments of the computational burden and methods for reducing it by exploiting the structure of the model have received consideration. Some early analyses and results are presented in Section I-F. In addition the utilization of sound numerical procedures (e.g., square-root filters) in carrying out the calculations have consistently received attention, beginning with the earliest applications.

The Kalman filter provides an estimate of the state of the system at the current time based on all measurements of the system obtained up to and including the present time. Smoothing is defined as a more general estimation problem in which the state at some or all times in the total sampling interval is estimated from data obtained at times that both precede and succeed the time being considered. While smoothing algorithms have not seen as widespread use as the Kalman filter for a variety of reasons, this presentation cannot be complete without consideration of this more general problem. Different types of smoothing problems can be defined and each problem can be solved with mathematically equivalent but computationally different algorithms. Definitions and different solutions of smoothing problems are described in Section I-G.

The development and analysis of the Kalman filter for linear systems (Section I-B), the consideration of its application to realistic nonlinear systems (Section I-C), the effects of model errors on the filter behavior (Section I-D), methods with which the effect of model errors can be diminished to tolerable levels (Section I-E), and modifications of the basic algorithm to reduce computational burden (Section I-F) provide an outline of the major investigative directions during the evolution of the Kalman filter to its current status. The papers included in Sections I-B through I-G of this volume

represent only a fraction, at best, of the published literature on the subject. A bibliography by Mendel and Giesekeing [2] and the survey paper by Kailath [3] provide additional sources for information about early work. Other references are included in the papers themselves, in the books that are listed in the Bibliography at the end of this Introduction, and in the references given at the end of each section.

Part II of this book contains the application papers that appeared in the Special Issue which was identified earlier. The reader should find at least several papers that are of interest because the specific application being considered. It is informative and, hopefully, illuminating to read each paper from the perspective of the manner in which the author dealt with the general problems considered in the papers of Part I. Many clever solutions are described which, while stemming from the same concerns as stimulated the analyses presented in Part I, illustrate the variety of approaches possible in achieving a successful implementation of a Kalman filter. Alternatively, some applications leave the reader wondering how or why a generic problem was not discussed.

REFERENCES

- [1] G. L. Smith, S. F. Schmidt, and L. A. McGee, "Application of statistical filter theory to the optimal estimation of position and velocity on-board a circumlunar vehicle," NASA-TR-R-135, 1962.
- [2] J. M. Mendel and D. L. Giesekeing, "Bibliography on the linear-quadratic-Gaussian problem," *IEEE Trans. Automat. Contr.*, vol. AC-16, 1971, pp. 847-869.
- [3] T. Kailath, "A view of three decades of linear filtering theory," *IEEE Trans. Inform. Theory*, vol. IT-20, 1974, pp. 146-181.
- [4] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
- [5] R. H. Battin, *Astronautical Guidance*. New York: McGraw-Hill, 1964.
- [6] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*. New York: Academic Press, 1977.
- [7] S. M. Bozic, *Digital and Kalman Filtering*. London: Arnold Publ., 1979.
- [8] R. G. Brown, *Introduction to Random Signal Analysis and Kalman Filtering*. New York: Wiley, 1983.
- [9] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*. Waltham, MA: Ginn, 1969.
- [10] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes and Applications to Guidance*. New York: Interscience, 1968.
- [11] M. H. A. Davis, *Linear Estimation and Stochastic Control*. New York: Halstead Press, 1977.
- [12] R. Deutsch, *Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1965.
- [13] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.
- [14] A. H. Jazwinski, *Stochastic Processes and Filtering*. New York: Academic Press, 1970.
- [15] C. L. Lawson and R. J. Hanson, *Solving Least-Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
- [16] R. C. K. Lee, *Optimal Estimation and Control*. Cambridge, MA: MIT Press, 1964.
- [17] C. T. Leondes, Ed., *Theory and Applications of Kalman Filtering*. NATO AGARDograph, No. 139, 1970.
- [18] C. T. Leondes, Ed., *Control and Dynamic Systems: Advances in Theory and Applications*, vols. 19, 20, 21. New York:

- Academic Press, 1983 and 1984.
- [19] P. B. Liebelt, *An Introduction to Optimal Estimation*. Reading, MA: Addison-Wesley, 1967.
 - [20] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, Vol. 1. New York: Academic Press, 1979.
 - [21] T. P. McGarty, *Stochastic Systems and State Estimation*. New York: Wiley, 1974.
 - [22] J. S. Meditch, *Stochastic Optimal Estimation and Control*. New York: McGraw-Hill, 1969.
 - [23] N. Nahi, *Estimation Theory and Applications*. New York: Wiley, 1969.
 - [24] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*. New York: McGraw-Hill, 1971.
 - [25] F. C. Schweppe, *Uncertain Dynamic Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
 - [26] H. W. Sorenson, *Parameter Estimation: Principles and Problems*. New York: Marcel Dekker Publ., 1980.

Part I

Theoretical Beginnings

Section I-A Historical Survey

Least-squares estimation: from Gauss to Kalman

The Gaussian concept of estimation by least squares, originally stimulated by astronomical studies, has provided the basis for a number of estimation theories and techniques during the ensuing 170 years—probably none as useful in terms of today's requirements as the Kalman filter

H. W. Sorenson University of California, San Diego

This discussion is directed to least-squares estimation theory, from its inception by Gauss¹ to its modern form, as developed by Kalman.² To aid in furnishing the desired perspective, the contributions and insights provided by Gauss are described and related to developments that have appeared more recently (that is, in the 20th century). In the author's opinion, it is enlightening to consider just how far (or how little) we have advanced since the initial developments and to recognize the truth in the saying that we "stand on the shoulders of giants."

The earliest stimulus for the development of estimation theory was apparently provided by astronomical studies in which planet and comet motion was studied using telescopic measurement data. The motion of these bodies can be completely characterized by six parameters, and the estimation problem that was considered was that of inferring the values of these parameters from the measurement data. To solve this problem concerning the revolution of heavenly bodies, the method of least squares was invented by a "young revolutionary" of his day, Karl Friedrich Gauss. Gauss was 18 years old at the time of his first use of the least-squares method in 1795.

As happens even today (e.g., the Kalman filter), there was considerable controversy in the early 19th century regarding the actual inventor of the least-squares method. The conflict arose because Gauss did not publish his discovery in 1795. Instead, Legendre independently invented the method and published his results in 1806 in his book *Nouvelles méthodes pour la détermination des orbites des comètes*. It was not until 1809, in his book *Theoria Motus Corporum Coelestium*, that Gauss published a detailed description of the least-squares method. However, in this treatise Gauss mentions Legendre's discussion of least squares and pointedly refers to his own earlier use (p. 270, *Theoria Motus*)*: "Our principle, which we

have made use of since the year 1795, has lately been published by Legendre in the work *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, 1806, where several other properties of this principle have been explained which, for the sake of brevity, we here omit." This reference angered Legendre who, with great indignation, wrote to Gauss and complained³ that "Gauss, who was already so rich in discoveries, might have had the decency not to appropriate the method of least-squares." It is interesting to note that Gauss, who is now regarded as one of the "giants" of mathematics, felt that he had been eclipsed by Legendre and wrote to a friend saying,³ "It seems to be my fate to concur in nearly all my theoretical works with Legendre. So it is in the higher arithmetic, . . . , and now again in the method of least-squares which is also used in Legendre's work and indeed right gallantly carried through." Historians have since found sufficient evidence to substantiate Gauss' claim of priority to the least-squares method, so it is Legendre rather than Gauss who was eclipsed in this instance and, indeed, in general.

The method of least squares

The astronomical studies that prompted the invention of least squares were described by Gauss in *Theoria Motus*.¹ The following quotation (p. 249) not only describes the basic ingredients for Gauss' studies but captures the essential ingredients for all other data-processing studies. "If the astronomical observations and other quantities on which the computation of orbits is based were absolutely correct, the elements also, whether deduced from three or four observations, would be strictly accurate (so far indeed as the motion is supposed to take place exactly according to the laws of Kepler) and, therefore, if other observations were used, they might be confirmed but not corrected. But since all our measurements and observations are nothing more than approximations to the truth, the same must be true of all calculations resting upon them, and the highest aim of all compu-

* The page numbers here refer to the English translation available from Dover Publications, Inc.¹

tations made concerning concrete phenomena must be to approximate, as nearly as practicable, to the truth. But this can be accomplished in no other way than by a suitable combination of more observations than the number absolutely requisite for the determination of the unknown quantities. This problem can only be properly undertaken when an approximate knowledge of the orbit has been already attained, which is afterwards to be corrected so as to satisfy all the observations in the most accurate manner possible."

Let us briefly reconsider some of the ideas contained in the preceding statement and relate them to "modern" developments.

1. Gauss refers to the number of observations that are absolutely required for the determination of the unknown quantities. The problem of establishing this minimum number of observations is currently discussed in terms of the "observability of the system" and is the subject of many papers; see Refs. 4 and 5, for example.

2. Gauss notes that more observations are required than this minimum because of the errors in the measurements and observations. Thus, he notes the need for "redundant" data to eliminate the influence of measurement errors.

3. Gauss implies that the equations of motion must be exact descriptions, and therefore the problem of dynamic modeling of the system is raised.

4. Gauss requires that approximate knowledge of the orbit be available. This is currently required in virtually all practical applications of Kalman filter theory,⁶ for example, and implies the use of some linearization procedure.

5. Gauss states that the parameter estimates must satisfy the observations in the most accurate manner possible. Thus, he calls for the residuals (that is, the difference between the observed values and the values predicted from the estimates) to be as small as possible.

6. Gauss refers to the inaccuracy of the observations and indicates that the errors are unknown or unknowable and thereby sets the stage for probabilistic considerations. In doing so, he anticipates most of the modern-day approaches to estimation problems.

7. Finally, Gauss refers to the "suitable combination" of the observations that will give the most accurate estimates. This is related to the definition of the structure of an estimation procedure (i.e., linear or nonlinear filtering) and to the definition of the performance criterion. These are extremely important considerations in current discussions of estimation problems.

As stated earlier, Gauss invented and used the method of least squares as his estimation technique. Let us consider Gauss' definition of the method (Ref. 1, page 260). He suggested that the most appropriate values for the unknown but desired parameters are the *most probable values*, which he defined in the following manner: "... the *most probable value* of the unknown quantities will be that in which the sum of the squares of the differences between the actually observed and the computed values multiplied by numbers that measure the degree of precision is a minimum." The difference between the observed and computed measurement values is generally called the *residual*.

To make the discussion more precise, consider the following statement of the estimation problem. Suppose that m measurement quantities are available at discrete

instants of time (t_1, t_2, \dots, t_n) and are denoted at each time t_k as z_k . Suppose that parameters x are to be determined from the data and are related according to

$$z_k = H_k x + v_k \quad (1)$$

where the v_k represent the measurement errors that occur at each observation time. As is seen in Eq. (1), the measurement data and the parameters x are assumed here to be linearly related, thereby making explicit the assumption that Gauss indicated was necessary in the foregoing quotation.

Denote the estimate of x based on the n data samples $\{z_1, z_2, \dots, z_n\}$ as \hat{x}_n . Then, the residual associated with the k th measurement is

$$r_k \triangleq z_k - H_k \hat{x}_n \quad k = 0, 1, \dots, n \quad (2)$$

The least-squares method is concerned with determining the most probable value of x (that is, \hat{x}_n). This most probable value is defined as the value that minimizes the sum of the squares of the residuals. Thus, choose x so that

$$L_n = \frac{1}{2} \sum_{k=0}^n [z_k - H_k x]^T W_k [z_k - H_k x] \quad (3)$$

is minimized. The elements of the matrixes W_k are selected to indicate the degree of confidence that one can place in the individual measurements. As will be explained more fully in the discussion of the Kalman filter, W_k is equivalent to the inverse of the covariance of the measurement noise.

Gauss with his remarkable insight recognized that the simple statement of the least-squares method contains the germ of countless interesting studies. As he says in *Theoria Motus* (Ref. 1, page 269): "The subject we have just treated might give rise to several elegant analytical investigations upon which, however, we will not dwell, that we may not be too much diverted from our object. For the same reason we must reserve for another occasion the explanation of the devices by means of which the numerical calculations can be rendered more expeditious." Judging by the interest in estimation theory over the years, this statement must stand as one of the greatest understatements of all time. In passing, we note that the Kalman filter can be rightfully regarded as an efficient computational solution of the least-squares method.

Gauss did not merely hypothesize the least-squares method; it is interesting to consider his discussion of the problem of obtaining the "most probable" estimate as an introduction to more modern techniques. First, it is significant that he considered the problem from a probabilistic point of view and attempted to define the best estimate as the most probable value of the parameters. He reasoned that errors in the measurements would be independent of each other, so the joint-probability density function of the measurement residuals can be expressed as the product of the individual density functions

$$f(r_0, r_1, \dots, r_n) = f(r_0)f(r_1) \cdots f(r_n) \quad (4)$$

Next, he argued that the density $f(r_k)$ would be a normal density

$$f(r_k) = \frac{\sqrt{\det W}}{(2\pi)^{m/2}} \exp \left[\frac{1}{2} r_k^T W_k r_k \right] \quad (5)$$

although he recognized that one never obtains errors of infinite magnitude, and thus Eq. (5) is not realistic.

However, he rationalized away this difficulty by stating (page 259) that: "The function just found cannot, it is true, express rigorously the probabilities of the errors for since the possible errors are in all cases confined within certain limits, the probability of errors exceeding those limits ought always to be zero while our formula always gives some value. However, this defect, which every analytical function must, from its nature, labor under, is of no importance in practice because the value of our function decreases so rapidly, when $[r_k^T W_k r_k]$ has acquired a considerable magnitude, that it can safely be considered as vanishing."

Gauss proceeded by noting that the maximum of the probability density function is determined by maximizing the logarithm of this function. Thus, he anticipated the *maximum likelihood method*, which was introduced by R. A. Fisher⁷ in 1912 and has been thoroughly investigated up to the present time. It is interesting that Gauss rejected the maximum likelihood method⁸ in favor of minimizing some function of the difference between estimate and observation, and thereby recast the least-squares method independent of probability theory. However, in maximizing the logarithm of the independent and normally distributed residuals, one is led to the least-squares problem defined by Eq. (3).

Kalman filter theory

Let us now leave the early 19th century and enter the 20th century. Consider, briefly, some of the major developments of estimation theory that preceded the introduction of the Kalman filter. As already mentioned, R. A. Fisher introduced the idea of maximum likelihood estimation and this has provided food for thought throughout the subsequent years. Kolmogorov⁹ in 1941 and Wiener¹⁰ in 1942 independently developed a linear minimum mean-square estimation technique that received considerable attention and provided the foundation for the subsequent development of Kalman filter theory.

In Wiener-Kolmogorov filter theory, Gauss' inference that linear equations must be available for the solution of the estimation problem is elevated to the status of an explicit assumption. There are, however, many conceptual differences (as one would hope after 140 years) between Gauss' problem and the problem treated by Wiener and Kolmogorov. Not the least of these is the fact that the latter considered the estimation problem when measurements are obtained continuously, as well as the discrete-time problem. To maintain the continuity of the present discussion, attention shall be restricted to the discrete formulation of Wiener-Kolmogorov (and later Kalman) filter theory.

Consider the problem of estimating a signal s_n , possibly time-varying, from measurement data (z_0, z_1, \dots, z_n) , where the s_n and the $\{z_i\}$ are related through knowledge of the cross-correlation functions. Assume that the estimate of s_n , say $\hat{s}_{n/n}$, is to be computed as a linear combination of the z_i :

$$\hat{s}_{n/n} = \sum_{i=0}^n H_{n,i} z_i \quad (6)$$

The "filter gains" $H_{n,i}$ are to be chosen so that the mean-square error is minimized; that is, choose the $H_{n,i}$ in such a way that

$$M_n = E[(s_n - \hat{s}_{n/n})^T (s_n - \hat{s}_{n/n})] \quad (7)$$

is minimized. It is well known¹¹ that a necessary and sufficient condition for $\hat{s}_{n/n}$ to minimize M_n is that the error in the estimate $(\tilde{s}_{n/n} \triangleq s_n - \hat{s}_{n/n})$ be orthogonal to the measurement data

$$E[\tilde{s}_{n/n} z_i^T] = 0 \quad i = 0, 1, \dots, n \quad (8)$$

This is the Wiener-Hopf equation, which is frequently written as

$$E[s_n z_i^T] = \sum_{j=0}^n H_{n,j} E[z_j z_i^T] \quad i = 0, 1, \dots, n \quad (9)$$

This equation must be solved for the $H_{n,j}$ in order to obtain the gains of the optimal filter. One can rewrite this as a vector-matrix equation whose solution, theoretically speaking, is straightforward. However, the matrix inversion that is required becomes computationally impractical when n is large. Wiener and Kolmogorov assumed an infinite amount of data (that is, the lower limit of the summation is $-\infty$ rather than zero), and assumed the system to be stationary. The resulting equations were solved using spectral factorization.^{9,10,12}

The problem formulated and described here is significantly different from Gauss' least-squares problem. First, no assumption is imposed that the signal is constant. Instead, the signal can be different at each n but can be described statistically by the autocorrelation and cross-correlation functions of the signal and measurement data. Second, instead of arguing that the estimate be the most probable, a probabilistic version of the least-squares method is chosen as the performance index.

It has been found that Eq. (9) is solved in a relatively straightforward manner if one introduces a "shaping filter"^{13,14} to give a more explicit description of the signal. In particular, suppose that the signal and measurement processes are assumed to have the following structure. The measurements are described by

$$\begin{aligned} z_i &= s_i + v_i \\ &= H_i x_i + v_i \end{aligned} \quad (10)$$

where v_i is a white-noise sequence (that is, v_i is both mutually independent and independent of x_i). The system state vector x_i is assumed to be described as a dynamic system having the form

$$x_{i+1} = \Phi_{i+1,i} x_i + w_i \quad (11)$$

where w_i represents a white-noise sequence. Note that if the noise w_i is identically zero and if $\Phi_{i+1,i}$ is the identity matrix, then the state is a constant for all i and one has returned basically to the system assumed by Gauss. With the system described by Eqs. (10) and (11), the known statistics for the initial state x_0 , and the noise sequences $\{w_i\}$ and $\{v_i\}$, one can proceed to the solution of Eq. (9).

Although the weighting function for the filter can be determined, a new solution must be generated for each n . It seems intuitively reasonable that estimates of s_{n+1} (or x_{n+1}) could be derived, given a new measurement z_{n+1} , from $\hat{s}_{n/n}$ and z_{n+1} rather than from $z_0, z_1, \dots, z_n, z_{n+1}$, since $\hat{s}_{n/n}$ is based on the data (z_0, z_1, \dots, z_n) . In 1955 J. W. Follin¹⁵ at Johns Hopkins University suggested a recursive approach based on this idea, which he carried out for a specific system. This approach had immediate appeal and essentially laid the foundation for the developments that are now referred to as the Kalman filter. It is clear (for example, see Ref. 16, p. 129) that Follin's

work provided a direct stimulus for the work of Richard Bucy, which led to his subsequent collaboration with Kalman in the continuous-time version of the filter equations.¹⁷ As frequently happens, the time was ripe for this approach, because several other people independently investigated recursive filter and prediction methods; see, for example, Refs. 18 and 19. Also, the method of stochastic approximation²⁰ was introduced and being studied for related problems²¹ during this period.

Kalman published his first paper on discrete-time, recursive mean-square filtering in 1960.³ It is interesting to note that, analogous to the Gauss-Legendre squabble concerning priority of the least-squares method, there is a difference of opinion concerning the originator of the Kalman filter. Peter Swerling published a RAND Corporation memorandum in 1958¹⁸ describing a recursive procedure for orbit determination. Of further interest is the fact that orbit determination problems provided the stimulus for both Gauss' work and more modern-day developments. Swerling's method is essentially the same as Kalman's except that the equation used to update the error covariance matrix has a slightly more cumbersome form. After Kalman had published his paper and it had attained considerable fame, Swerling²² wrote a letter to the *AIJA Journal* claiming priority for the Kalman filter equations. It appears, however, that his plea has fallen on deaf ears.

The developments beginning with Wiener's work and culminating with Kalman's reflect fundamentally the changes that have occurred in control systems theory during this period. In the "classical control theory," the emphasis was on the analysis and synthesis of systems in terms of their input-output characteristics. The basic tools used for these problems were the Laplace and Fourier transforms. The original formulation and solution of the Wiener-Kolmogorov filtering problem is consistent with this basic approach. More recent developments have stressed the "state-space" approach, in which one deals with the basic system that gives rise to the observed output. It represents in many ways a return to Gauss' approach, since he referred to the dynamic modeling problem as noted earlier. Also, the state-space approach makes use of difference and differential equations rather than the integral equations of the classical approach. Although the two approaches are mathematically equivalent, it seems to be more satisfying to work with differential equations (probably since dynamical systems are generally described in this manner).

At this point, let us summarize the Kalman filtering problem and its solution. The system that is considered is composed of two essential ingredients. First, the state is assumed to be described by

$$\mathbf{x}_{k+1} = \Phi_{k+1,k} \mathbf{x}_k + \mathbf{w}_k \quad (11')$$

and the measurement data are related to the state by

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (10')$$

where $\{\mathbf{w}_k\}$ and $\{\mathbf{v}_k\}$ represent independent white-noise sequences. The initial state \mathbf{x}_0 has a mean value $\hat{\mathbf{x}}_{0/-1}$ and covariance matrix $\mathbf{P}_{0/-1}$ and is independent of the plant and measurement noise sequences. The noise sequences have zero mean and second-order statistics described by

$$\begin{aligned} E[\mathbf{v}_k \mathbf{v}_j^T] &= \mathbf{R}_k \delta_{kj} & E[\mathbf{w}_k \mathbf{w}_j^T] &= \mathbf{Q}_k \delta_{kj} \\ E[\mathbf{v}_k \mathbf{w}_j^T] &= 0 & \text{for all } k, j \end{aligned}$$

where δ_{kj} is the Kronecker delta.

An estimate $\hat{\mathbf{x}}_{k/k}$ of the state \mathbf{x}_k is to be computed from the data $\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_k$ so as to minimize the mean-square error in the estimate. The estimate that accomplishes this is to be computed as an explicit function only of the measurement \mathbf{z}_k and the previous best estimate $\hat{\mathbf{x}}_{k-1/k-1}$. This approach leads to a recursive solution that provides an estimate that is equivalent to the estimate obtained by processing all of the data simultaneously but reduces the data-handling requirements. The estimate of the signal

$$\mathbf{s}_k = \mathbf{H}_k \mathbf{x}_k \quad (12)$$

is given by

$$\hat{\mathbf{s}}_{k/k} = \mathbf{H}_k \hat{\mathbf{x}}_{k/k} \quad (13)$$

The solution of this recursive, linear, mean-square estimation problem can be determined from the orthogonality principle given in Eq. (8), as well as in a large variety of other ways, and is presented below. This system of equations has come to be known as the Kalman filter. The estimate is given as the linear combination of the estimate predicted in the absence of new data, or

$$\hat{\mathbf{x}}_{k/k-1} = \Phi_{k,k-1} \hat{\mathbf{x}}_{k-1/k-1}$$

and the residual \mathbf{r}_k . Thus, the mean-square estimate is

$$\hat{\mathbf{x}}_{k/k} = \Phi_{k,k-1} \hat{\mathbf{x}}_{k-1/k-1} + \mathbf{K}_k [\mathbf{z}_k - \mathbf{H}_k \Phi_{k,k-1} \hat{\mathbf{x}}_{k-1/k-1}] \quad (14)$$

The gain matrix \mathbf{K}_k can be considered as being chosen to minimize $E[(\hat{\mathbf{x}}_k - \mathbf{x}_{k/k})^T(\mathbf{x}_k - \mathbf{x}_{k/k})]$ and is given by

$$\mathbf{K}_k = \mathbf{P}_{k/k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k/k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (15)$$

The matrix $\mathbf{P}_{k/k-1}$ is the covariance of the error in the predicted estimate and is given by

$$\mathbf{P}_{k/k-1} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k/k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k/k-1})^T] \quad (16)$$

$$= \Phi_{k/k-1} \mathbf{P}_{k-1/k-1} \Phi_{k,k-1}^T + \mathbf{Q}_{k-1} \quad (17)$$

The $\mathbf{P}_{k/k}$ is the covariance of the error in the estimate $\hat{\mathbf{x}}_{k/k}$.

$$\mathbf{P}_{k/k} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k/k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k/k})^T] \quad (18)$$

$$= \mathbf{P}_{k/k-1} - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k/k-1} \quad (19)$$

Equations (14), (15), (17), and (19) form the system of equations comprising the Kalman filter.^{2,4}

Kalman filter theory—a perspective

Let us relate elements of this problem to Gauss' earlier arguments. First, Kalman assumes that the noise is independent from one sampling time to the next. But it is clear from Eq. (5) that this is equivalent to assuming that the residual $(\mathbf{z}_k - \mathbf{H}_k \mathbf{x}_k)$ is independent between sampling times and therefore agrees with Gauss' assumption. Next, the noise and initial state are essentially assumed by Kalman to be Gaussian. The linearity of the system causes the state and measurements to be Gaussian at each sampling time. Thus the residual is Gaussian, as Gauss assumed. Therefore, one sees that the basic assumptions of Gauss and Kalman are identical except that the latter allows the state to change from one time to the next. This difference introduces a nontrivial modification to Gauss' problem but one that can be treated within a least-squares framework if the noise $\{\mathbf{w}_k\}$ is considered as the error in the plant model at each stage. In particular, one can formulate the least-squares problem as that of

choosing the estimates $\hat{\mathbf{x}}_{k/k}$ and the plant errors \mathbf{w}_k to minimize the modified least-squares performance index.

$$L_n = \frac{1}{2} (\mathbf{x}_0 - \mathbf{a})^T M_0^{-1} (\mathbf{x}_0 - \mathbf{a}) + \frac{1}{2} \sum_{i=0}^n (\mathbf{z}_i - H_i \mathbf{x}_i)^T R_i^{-1} (\mathbf{z}_i - H_i \mathbf{x}_i) + \frac{1}{2} \sum_{i=0}^{n-1} \mathbf{w}_i^T Q_i^{-1} \mathbf{w}_i \quad (20)$$

subject to the constraint

$$\mathbf{x}_k = \Phi_{k,k-1} \mathbf{x}_{k-1} + \mathbf{w}_{k-1} \quad (21)$$

Note that the first term essentially describes the uncertainty in the initial state. If one has no a priori information, then M_0^{-1} is identically zero and the term vanishes. Similarly, if there is no error in the plant equation, Q_i^{-1} is identically zero so this term vanishes. Then Eq. (20) is seen to reduce to Gauss' least-squares problem, as given in Eq. (3). The weighting matrices M_0^{-1} , R_i^{-1} , and Q_i^{-1} represent the matrix inverses of the a priori covariance matrices if a probabilistic interpretation is desired.

One can obtain a recursive solution to the problem of minimizing (20) by noting that

$$L_n = L_{n-1} + \frac{1}{2} (\mathbf{z}_n - H_n \mathbf{x}_n)^T R_n^{-1} (\mathbf{z}_n - H_n \mathbf{x}_n) + \frac{1}{2} \mathbf{w}_{n-1}^T Q_{n-1}^{-1} \mathbf{w}_{n-1} \quad (22)$$

and by then proceeding inductively starting with $n = 0$ to obtain recursion relations for the least-square estimate.²³ If this is done, the Kalman filter equations are obtained. It is then indicated that, for this linear problem, deterministic least-squares estimation theory and the probabilistically based mean-square estimation theory are equivalent. Further, the problem of minimizing Eq. (22) is seen to give the most probable estimate for this system.

Since the Kalman filter represents essentially a recursive solution of Gauss' original least-squares problem, it is reasonable to consider the substance of Kalman's contribution and attempt to put it into perspective. It cannot be denied that there has been a substantial contribution if for no other reason than the large number of theoretical and practical studies that it has initiated. I suggest that the contribution is significant for two basic reasons:

1. The Kalman filter equations provide an extremely convenient procedure for digital computer implementation. One can develop a computer program using the Kalman filter in a direct manner that (initially, at least) requires little understanding of the theory that led to their development. There are well-established numerical procedures for solving differential equations, so the engineer does not have to be worried about this problem. By contrast, the solution of the Wiener-Hopf equation and the implementation of the Wiener-Kolmogorov filter must be regarded as more difficult or there would have been no need for the Kalman filter. Since Gauss was very concerned with the computational aspects of least-squares applications, one can imagine that he would appreciate the computational benefits of the Kalman filter.

2. Kalman posed the problem in a general framework that has had a unifying influence on known results.

Further, one can analyze the behavior of the estimates within the general framework and thereby obtain significant insights into the results obtained from computational studies. There has been a veritable "explosion" of theoretical papers that have recognizable roots in Kalman's work and thereby testify to the richness of his formulation.

Finally, a third reason for the popularity might be considered, although it is less tangible in character than the other two. It is worth noting that Kalman recognized the potential of his results, whereas others working in the area either did not or were not as successful in communicating the intrinsic worth of recursive filtering to others. One cannot overemphasize the value of recognizing and successfully communicating significant new results.

The Kalman filter, which assumes linear systems, has found its greatest application to nonlinear systems. It is generally used in these problems by assuming knowledge of an approximate solution (as Gauss proposed) and by describing the deviations from the reference by linear equations. The approximate linear model that is obtained forms the basis for the Kalman filter utilization. Commonly, such applications are accomplished with great success but, on occasion, unsatisfactory results are obtained because a phenomenon known as divergence occurs.^{24, 25}

Divergence is said to occur when the error covariance matrix P_k computed by the filter becomes unjustifiably small compared with the actual error in the estimate. When P_k becomes small, it causes the gain matrix to become too small and new measurement data are given too little weight. As a result, the plant model becomes more important in determining the estimate than are the data and any errors in the model can build up over a period of time and cause a significant degradation in the accuracy of the estimate. This happens most commonly when the plant is assumed to be error-free (i.e., $\mathbf{w}_k \equiv 0$ for all k). If the model were perfect and contained no random or model errors, then the vanishing of the error covariance matrix would be desirable and would represent the fact that the state could be determined precisely if sufficient redundant data were processed. However, it is naive at best to assume that any physical system can be modeled precisely, so it is necessary to account for model errors. Thus, it has become good practice⁶ always to include the plant error or noise term \mathbf{w}_k . It should be emphasized that divergence does not occur because of any fault of the filter. If the system were actually linear, Kalman⁴ showed that the filter equations are stable under very reasonable conditions. Thus, the divergence is a direct consequence of the errors introduced by the linear approximation.

To reduce approximation errors, the so-called "extended Kalman filter" is generally used in practice. In this case the nonlinear system is linearized by employing the best estimates of the state vector as the reference values used at each stage for the linearization. For example, at time t_{k-1} , the estimate $\hat{\mathbf{x}}_{k-1/k-1}$ is used as the reference in obtaining the transition matrix $\Phi_{k,k-1}$. This approximation is utilized in Eq. (17) to obtain the Kalman error covariance $P_{k/k-1}$. The estimate is given by

$$\hat{\mathbf{x}}_{k/k-1} = \mathbf{f}_k(\hat{\mathbf{x}}_{k-1/k-1}) \quad (23)$$

where \mathbf{f}_k is used to denote the nonlinear plant equation.

In most cases it is obtained as the solution of an ordinary differential equation that describes the plant behavior. The processing of the data obtained at t_k is accomplished in a similar manner. The estimate $\hat{x}_{k/k-1}$ serves as the reference for the determination of a linear approximation H_k to the nonlinear measurement equation. The matrix H_k is used in Eqs. (15) and (19) to determine the gain and error covariance matrices K_k and $P_{k/k}$. The filtered estimate is then given by

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k[z_k - h_k(\hat{x}_{k/k-1})] \quad (24)$$

where h_k is used to denote the measurement nonlinearity; that is, the measurement is assumed to be described by

$$z_k = h_k(x_k) + v_k \quad (25)$$

Through the use of the extended Kalman filter, one can hope to eliminate or reduce divergence. Note, however, that the $P_{k/k-1}$ and $P_{k/k}$ matrices are still linear approximations of the true error covariance matrices. Further, the nonlinear models f_k and h_k are themselves approximations of the actual physical system, so modeling errors can still exist. Thus, the extended Kalman filter does not insure the elimination of the divergence problem.

In a practical application one does not know the error in the state estimate, so there are grounds for uneasiness in using this method. Of course, the same type of problem must be considered in any least-squares application and we can return to Gauss for the means of judging the behavior of the filter. Recall that he said that the estimates should satisfy all the observations in the most accurate manner possible. Thus, one is led to further consideration of the residuals as a measure of filter performance. Kailath²⁶ pointed out recently that the residual sequence $(z_k - H_k \hat{x}_{k/k-1})$ is a white-noise sequence. Since the residual can be computed explicitly, it can be examined at each stage to verify that the residual (or innovations) sequence has the appropriate statistical characteristics. A method of controlling the divergence, based on the residuals, has been proposed^{27,28} in which the plant and measurement noise covariance matrices Q_k and R_k are chosen in a manner that is appropriate to cause the residuals to have the desired properties. But this method is essentially the same as choosing the least-squares weighting matrices as a reflection of the accuracy of the measurements (or plant). Thus, the least-squares aspect continues to dominate the practical application of the method.

REFERENCES

1. Gauss, K. G., *Theory of Motion of the Heavenly Bodies*. New York: Dover, 1963.
2. Kalman, R. E., "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82D, pp. 35-45, Mar. 1960.
3. Bell, E. T., *Men of Mathematics*. New York: Simon and Schuster, 1961.
4. Kalman, R. E., "New methods in Wiener filtering theory," in *Proc. First Symp. of Engineering Application of Random Function Theory and Probability*. New York: Wiley, 1963, pp. 270-388.
5. Sorenson, H. W., "Controllability and observability of linear, stochastic, time-discrete control systems," in *Advances in Control Systems*, Vol. 6, C. T. Leondes, ed. New York: Academic Press, 1968.
6. Sorenson, H. W., "Kalman filtering techniques," in *Advances in Control Systems*, Vol. 3, C. T. Leondes, ed. New York: Academic Press, 1966.
7. Fisher, R. A., "On an absolute criterion for fitting frequency curves," *Messenger of Math.*, vol. 41, p. 155, 1912.
8. Berkson, J., "Estimation by least-squares and by maximum likelihood," *Proc. Third Berkeley Symp. on Mathematical Statistics, and Probability*, Vol. 1. Berkeley: University of California Press, 1956, pp. 1-11.
9. Kolmogorov, A. N., "Interpolation and extrapolation of stationary random sequences," translated by W. Doyle and J. Selin, Rept. RM-3090-PR, RAND Corp., Santa Monica, Calif., 1962.
10. Wiener, N., *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*. New York: Wiley, 1949.
11. Sorenson, H. W., and Stubberud, A. R., "Linear estimation theory," Chapter 1 in *Theory and Application of Kalman Filtering*; *AGARDograph*, no. 139, 1970.
12. Yaglom, A. M., *An Introduction to the Theory of Stationary Random Functions*, translated by R. A. Silverman. Englewood Cliffs, N.J.: Prentice-Hall, 1962.
13. Bode, H. W., and Shannon, C. E., "A simplified derivation of linear least-squares smoothing and prediction theory," *Proc. IRE*, vol. 38, pp. 417-425, Apr. 1950.
14. Zadeh, L. A., and Ragazzini, J. R., "An extension of Wiener's theory of prediction," *J. Appl. Phys.*, vol. 21, pp. 645-655, July 1950.
15. Carlton, A. G., and Follin, J. W., "Recent developments in fixed and adaptive filtering," *AGARDograph*, no. 21, 1956.
16. Bucy, R. S., and Joseph, P. D., *Filtering for Stochastic Processes with Applications to Guidance*. New York: Interscience, 1968.
17. Kalman, R. E., and Bucy, R. S., "New results in linear filtering and prediction theory," *J. Basic Eng.*, vol. 83D, pp. 95-108, 1961.
18. Swerling, P., "A proposed stagewise differential correction procedure for satellite tracking and prediction," Rept. P-1292, *RAND Corp.* Santa Monica, Calif., Jan. 1958; also published in *J. Astronaut. Sci.*, vol. 6, 1959.
19. Blum, M., "Recursion formulas for growing memory digital filters," *IRE Trans. Information Theory*, vol. IT-4, pp. 24-30, Mar. 1958.
20. Robbins, H., and Munro, S., "A stochastic approximation method," *Ann. Math. Stat.*, vol. 22, pp. 400-407, 1951.
21. Kiefer, J., and Wolfowitz, J., "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.*, vol. 23, pp. 462-466, 1952.
22. Swerling, P., "Comment on 'A statistical optimizing navigation procedure for space flight,'" *AIAA J.*, vol. 1, p. 1968, Aug. 1963.
23. Sorenson, H. W., "Comparison of Kalman, Bayesian and maximum likelihood estimation techniques," Chapter 6 in *Theory and Application of Kalman Filtering*; *AGARDograph*, no. 139, 1970.
24. Schlee, F. H., Standish, C. J., and Toda, N. F., "Divergence in the Kalman filter," *AIAA J.*, vol. 5, pp. 1114-1120, June 1967.
25. Fitzgerald, R. J., "Error divergence in optimal filtering problems," presented at the 2nd IFAC Symposium on Automatic Control in Space, Vienna, Austria, Sept. 1967.
26. Kailath, T., "An innovations approach to least-squares estimation—Part I: linear filtering in additive white noise," *IEEE Trans. Automatic Control*, vol. AC-13, pp. 646-655, Dec. 1968.
27. Jazwinski, A. H., "Adaptive filtering," *Automatica*, vol. 5, pp. 475-485, 1969.
28. Mehra, R. K., "On the identification of variances and adaptive filtering," *Preprints 1969 Joint Automatic Control Conf.*, pp. 494-505.