

Peter Young

CCES

Recursive Estimation and Time-Series Analysis

An Introduction



Springer-Verlag
Berlin Heidelberg New York Tokyo

0212

8562561

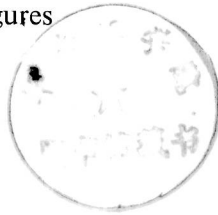
Y1

Peter Young

Recursive Estimation and Time-Series Analysis

An Introduction

With 54 Figures



E8562561

Springer-Verlag
Berlin Heidelberg New York Tokyo 1984

PETER YOUNG

Professor and Head of Department of Environmental Sciences
University of Lancaster
Lancaster, LA1 4YQ, Lancs. England

ISBN 3-540-13677-0 Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-13677-0 Springer-Verlag New York Heidelberg Berlin Tokyo

Library of Congress Cataloging in Publication Data

Young, Peter C.

Recursive estimation and time-series analysis.

(Communications and control engineering series)

Based on lecture notes originally prepared for a NATO summer school on

"The theory and practice of systems modelling and identification" held July 17-28, 1972,
at the Ecole nationale supérieure de l'aéronautique et de l'espace.

Bibliography: p.

Includes indexes.

1. Estimation theory. 2. Time-series analysis.

I. Title. II. Series.

QA276.8.Y68 1984 519.5'44 84-14078

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically those of translation, reprinting, re-use of illustrations, broadcasting, reproduction by photocopying machine or similar means, and storage in data banks. Under § 54 of the German Copyright Law where copies are made for other than private use, a fee is payable to "Verwertungsgesellschaft Wort", Munich.

© Springer-Verlag, Berlin, Heidelberg 1984

Printed in Germany

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Offsetprinting: Mercedes-Druck, Berlin

Bookbinding: Lüderitz & Bauer, Berlin

2061/3020 5 4 3 2 1 0

1725088

Communications and Control Engineering Series

Editors: A. Fettweis · J. L. Massey · M. Thoma

To
Wendy, Timothy, Melanie and Jeremy

8962561



Preface

This book has grown out of a set of lecture notes prepared originally for a NATO Summer School on "*The Theory and Practice of Systems Modelling and Identification*" held between the 17th and 28th July, 1972 at the Ecole Nationale Supérieure de L'Aeronautique et de L'Espace. Since this time I have given similar lecture courses in the Control Division of the Engineering Department, University of Cambridge; Department of Mechanical Engineering, University of Western Australia; the University of Ghent, Belgium (during the time I held the IBM Visiting Chair in Simulation for the month of January, 1980), the Australian National University, and the Agricultural University, Wageningen, the Netherlands. As a result, I am grateful to all the recipients of these lecture courses for their help in refining the book to its present form; it is still far from perfect but I hope that it will help the student to become acquainted with the interesting and practically useful concept of recursive estimation. Furthermore, I hope it will stimulate the reader to further study the theoretical aspects of the subject, which are not dealt with in detail in the present text. The book is primarily intended to provide an introductory set of lecture notes on the subject of recursive estimation to undergraduate/Masters students. However, the book can also be considered as a "theoretical background" handbook for use with the CAPTAIN Computer Package. This 'Computer Aided Program for Time Series Analysis and the Identification of Noisy Systems' was originally conceived by the author in the mid nineteen sixties and was developed initially in the Control Division of the Engineering Department at the University of Cambridge in 1971 on the basis of recursive algorithms developed during the previous six years (Young *et al.*, 1971; Shellswell, 1971). A *command mode* version of the package was developed as an alternative to the original interactive, *conversational mode* version, by the Institute of Hydrology in England (Venn and Day, 1977) and in this form, it has been acquired by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) of Australia for use in its CSIRONET nationwide computing system (Freeman, 1981). More recently, advanced versions of CAPTAIN, based on some of the more sophisticated procedures discussed in the present book, have been developed by the author and his colleagues at the Centre for Resource and Environmental Sciences, (CRES), ANU, for use in the UNIVAC 1100 series computer. A more comprehensive, "user friendly" version of the package is currently under development by the author and Mr. John Hampton in the Department of Environmental Sciences, University of Lancaster, Lancaster, LA1 4YQ, England. This version will be available initially for use on the VAX 11/780 computer. Also a microcomputer version MICROCAPTAIN has recently been developed by the author (see Epilogue) for use on the APPLE II and EPSON HX-20 microcomputers. Any enquiries about any of these CAPTAIN programs should be addressed to me at the above address.

Acknowledgements

The final appearance of this book has been delayed for various reasons. Foremost amongst these are the various changes in location of my family and myself during the preparation of the text. It was begun in 1979, while I was spending a period of study leave at my old College, Clare Hall, in Cambridge; it was continued, very much as a spare time activity between numerous environmental projects, at the Centre for Resource and Environmental Studies, Australian National University, Canberra, between 1980 and 1981; and it has been completed, whenever time has allowed, at the Department of Environmental Science, University of Lancaster, between 1981 and 1983.

The overlong gestation period has meant that, at various times, a number of people have read draft chapters of the book and suggested improvements. In particular, I would like to thank Dr. Paul Steele, Dr. Tony Jakeman, and Miss Christina Sirakoff for their help and encouragement. Mrs. June Harries (Canberra) and Mrs. Glenys Booth (Lancaster) have typed most of the manuscript, often from poorly handwritten drafts, and I am extremely grateful to both of them for their patience and the quality of their work. A number of the diagrams were drawn by Elizabeth Barta.

Of course all the errors, omissions and other deficiencies of the book are my sole responsibility. Like most authors, if I was writing the book again, I would change many aspects of the presentation. Having gone through the not inconsiderable labours of preparing a "camera-ready" manuscript, however, I find such a possibility daunting in the extreme. I hope, therefore, the reader will not be too disappointed in the final outcome.

Finally, in dedicating this book to my wife and family, I am acknowledging their unfailing support, without which the book would never have appeared.

Peter Young
Lancaster, England.
January 1984.

Contents

1. <u>INTRODUCTION</u>	1
2. <u>RECURSIVE ESTIMATION : A TUTORIAL INTRODUCTION</u>	9
2.1 Recursive Estimation of the Mean Value of a Random Variable	10
2.2 Recursive Least Squares Estimation for a Single Unknown Parameter in a Regression Relationship	17
2.3 Summary	23
3. <u>RECURSIVE ESTIMATION AND STOCHASTIC APPROXIMATION</u>	24
3.1 The Recursive Least Squares Algorithm	24
3.2 Connection with Stochastic Approximation	30
3.3 Some Extensions to Stochastic Approximation	37
. Matrix gain SA and optimum algorithms	37
. Continuous-time algorithms	39
. Search algorithms	40
. Acceleration of convergence	40
3.4 Summary	41
4. <u>RECURSIVE LEAST SQUARES REGRESSION ANALYSIS</u>	42
4.1 The RLS Algorithm for the General Linear Regression Model	42
4.2 Some Cautionary Comments : Multiple Collinearity and Errors-in-Variables	49
. Multiple collinearity	50
. Errors-in-variables and the structural model	51
4.3 Summary	54
5. <u>RECURSIVE ESTIMATION OF TIME-VARIABLE PARAMETERS</u>	
<u>IN REGRESSION MODELS</u>	55
5.1 Shaping the Memory of the Estimator	57
. The moving rectangular window	58
. The moving exponential window (exponential forgetting factor)	60
5.2 Modelling the Parameter Variations	66

5.3	Vector Measurements and the Kalman Filter- Estimation Algorithm	75
	. Estimation of the varying mean value for Walgett rainfall data	82
5.4	The Estimation of Rapidly Varying Parameters	84
	. Simple parameter decomposition	84
	. A practical example : Missile parameter estimation	86
	. More complex methods	92
5.5	Simplified 'Constant Gain' Algorithms for Time Variable Parameter Estimation	94
5.6	Recursive Smoothing Algorithms	96
5.7	Statistical Methods for Detecting the Presence of Parameter Variation	98
5.8	Summary	102
6.	<u>THE TIME-SERIES ESTIMATION PROBLEM</u>	104
6.1	The Time-Series Model in the Observation Space	105
6.2	Various Observation Space Model Forms	109
6.3	Least Squares Estimation : Its Advantages and Limitations	118
6.4	Least Squares Estimation : A Special Case	123
6.5	The General Case : The Structural Model	125
6.6	Summary	128
7.	<u>THE INSTRUMENTAL VARIABLE (IV) METHOD OF TIME-SERIES ANALYSIS</u>	129
7.1	The Recursive IV Algorithm for Time-Series Models	130
7.2	The Recursive-Iterative IV Algorithm	134
7.3	Estimation of the Noise Model Parameters : the AML Method	137
7.4	Statistical Properties of the IV and AML Estimates	141
	. IV estimates	141
	. AML estimates	142
7.5	Convergence of the IV and AML Estimators	143

7.6	Identifiability and the Choice of Input Signals	143
	. A special case	144
	. Choice of input signals	146
	. Restrictions on the system to be identified	147
	. The general case	148
	. Noise process identifiability	149
	. Some concluding comments on identifiability	149
7.7	Parametric Variations	150
7.8	A Time-Series Analysis Procedure Based on IV-AML Estimation	151
7.9	Representative Practical Results	152
	. The gas furnace data of Box-Jenkins	152
	. Rainfall-runoff modeling for the Bedford- Ouse River Basin	157
7.10	Summary	166
8.	<u>OPTIMUM INSTRUMENTAL VARIABLE METHODS OF TIME-SERIES MODEL ESTIMATION</u>	168
8.1	The Maximum-Likelihood Method	168
8.2	IV Within the Context of Maximum Likelihood	170
8.3	The AML Method Within the Context of Maximum Likelihood	173
8.4	A Refined Recursive IV-AML Approach to Time- Series Analysis	175
8.5	The Statistical Properties of the Refined IVAML Estimates	184
8.6	Performance of the Refined IVAML Algorithms	185
8.7	A Practical Example : Analysis of Time-Series Tracer Data in Translocation Studies	195
8.8	The Optimum Generalized Equation Error (OGEE) Approach to Time-Series Analysis	198
	. For various common time-series model forms	198
	. For a multiple input, single output (MISO) system model	198
8.9	Summary	203
9.	<u>ALTERNATIVE RECURSIVE APPROACHES TO TIME-SERIES ANALYSIS</u>	205
9.1	Prediction Error (PE) Methods	205
	. Statistical Properties of PE Estimates for the Transfer Function (TF) Model	213

9.2	The Extended Kalman Filter	215
	. A practical example : stochastic-dynamic model for water quality in the Bedford-Ouse River	222
9.3	Maximum Likelihood Estimation in the State Space	226
9.4	Summary	228
10.	<u>RECURSIVE ESTIMATION : A GENERAL TOOL IN DATA ANALYSIS AND STOCHASTIC MODEL BUILDING</u>	231
10.1	Pre-processing of Time-Series Data	231
10.2	Model Structure Identification	233
10.3	Model Parameter Estimation	234
10.4	State Estimation	235
10.5	Self-Adaptive (or Self-Tuning) Estimation, Forecasting and Control	238
10.6	Summary	239
11.	<u>EPILOGUE</u>	241
APPENDIX 1	<u>RELEVANT MATHEMATICAL AND STATISTICAL BACKGROUND MATERIAL</u>	245
A.1.1	<u>Matrix Algebra</u>	245
1.	Matrices	245
2.	Vectors	246
3.	Matrix Addition (or Subtraction)	246
4.	Matrix or Vector Transpose	246
5.	Matrix Multiplication	247
6.	Determinant of a Matrix	248
7.	Partitioned Matrices	249
8.	Inverse of a Matrix	250
9.	Quadratic Forms	251
10.	Positive Definite or Semi-Definite Matrices	252
11.	The Rank of a Matrix	252
12.	Differentiation of Vectors and Matrices	253
A.1.2	<u>Statistics and Probability</u>	254
1.	Discrete Random Variables	254
	. Mean value	
	. Variance	

2. Discrete Random Vectors	255
. Joint probability mass function	
. Marginal probability mass function	
. Mean	
. Covariance Matrix	
3. Conditional Probabilities	256
. Conditional probability mass function	
4. Continuous Random Variables and Vectors	256
. Probability Density Function	
5. The Normal or Gaussian Density Function	257
. Normally distributed random variable (scalar)	
. Normally distributed random vector	
6. Properties of Estimators	258
. Unbiased	
. Asymptotically unbiased	
. Minimum variance, unbiased estimators	
. Probability-in-the-limit (p.lim)	
7. The Likelihood Function and Maximum Likelihood Estimation	259
. Hessian matrix	
8. The Cramer-Rao Lower Bound	261
. Information matrix	
9. Time-Series	261
. Mean and variance	
. Covariance and correlation	
. Covariance matrix	
. White noise	
10. Gauss-Markov Random Sequences	263
. Conditional probability density function	
. Joint probability density function	
A.1.3 <u>Simple Deterministic Dynamic Systems</u>	264
1. First Order, Continuous-Time Linear Dynamic System	264
. Time-constant	
. Steady-state gain	
2. First Order Discrete-Time Linear Dynamic System	264

3. The Discrete-Time State Space Representation of a Deterministic Dynamic System	265
4. Transfer Function Representation of a Single Input, Single Output (SISO) Discrete Dynamic System	266
5. The Infinite Dimensional, Impulse Response Representation of a Linear SISO Discrete Dynamic System	267
6. Differentiation of a TF with respect to a Given Parameter	268
<u>APPENDIX 2 GAUSS'S DERIVATION OF RECURSIVE LEAST SQUARES</u>	270
<u>APPENDIX 3 THE INSTANTANEOUS COST FUNCTION ASSOCIATED WITH THE RECURSIVE LEAST SQUARES ALGORITHM</u>	278
<u>REFERENCES</u>	281
AUTHOR INDEX	291
SUBJECT INDEX	294

Algorithms

I	Recursive Least Squares (scalar RLS)	21
II	Recursive Least Squares (vector RLS)	26
III	Recursive Statistical Regression Algorithm for Regression Models with Multiple Unknown Parameters	46
IV	Recursive Instrumental Variable (IV)	
	. static	53
	. time-series	132
V	RLS with Rectangularly Weighted Past (RWP) Modification	60
VI	RLS with Exponentially Weighted Past (EWP) Modification	63
VII	Time Variable Parameter (TVP) Version of Regression Algorithm with Parameters modeled as Stochastic Gauss-Markov Random Sequences	70
VIII	RLS for Vector Measurements	77
IX	TVP for Vector Measurements	78
X	Kalman Filter	79
XI	Smoothing Version of TVP Algorithm	97
XII	RLS for Special Time-Series Model	124
XIII	Recursive Approximate Maximum Likelihood (AML)	140
XIV	Recursive IV for Time-Series Models with Time- Variable Parameters (Heuristic)	151
XV	Refined Recursive IV (RIV)	177
XVI	Refined Recursive AML (RAML)	178
XVII	Symmetric Gain RIV	183
XVIII	Symmetric Gain RAML	183
XIX	Refined Recursive IVAML for Multiple Input, Single Output (MISO) Time-Series Models	203
XX	Recursive Prediction Error (PE) Algorithm for Transfer Function Time-Series Model	211
XXI	Extended Kalman Filter	220
XXII	Recursive/Iterative Maximum Likelihood Algorithm in the State Space	226
XXIII	Refined IVAML for Time Variable Parameter (TVP) Time-Series Models	235
XXIV	Special Stochastic Approximation (SA) Version of Recursive IV	239

1 Introduction

*Nous traiterons particulièrement le
problème suivant, tant à cause de son utilité
pratique que de la simplicité de la solution*

K.F. Gauss
Méthode des Moindres Carrés
Translation by J. Bertrand, 1855.

The concept of least squares estimation is inextricably linked with the name of Karl Friedrich Gauss. Although Legendre was responsible for the first published account of the theory in 1805 and, indeed, first coined the term "least squares", it was Gauss who developed the method into a statistical tool, embedding it within a statistical framework involving a probabilistic treatment of observational errors.

Gauss's first published exposition on least squares appeared in his famous *Theoria Motus Corporum Coelestium* which appeared in 1809 when he was 31 years of age. But, as D.A. Sprott (1978) has pointed out recently in his excellent review of Gauss's contributions to statistics, the basic ideas were most probably formulated while he was still in his twenties.

In the *Theoria Motus*, the discussion on least squares appears in relation to an important practical problem; namely the estimation of the six constant coefficients or "parameters" that determine the elliptical orbit of a planetary body, on the basis of $n > 6$ observations. His second exposition on the subject was presented in a series of papers (1821, 1823, 1826) which were collected together under the title *Theoria Combinationis Erroribus Minimum Obnoxiae*. Here he abandoned the previous "inferential" treatment delineated in the *Theoria Motus* and concentrated on a "decision theoretic" approach, in which he restricted attention to estimates that are a linear function of the observations. And it was here that he presented perhaps his most famous statistical theorem that, *Among all linear error-consistent estimates, the least squares estimate has minimum mean square error.*

But our current interest lies elsewhere in the *Theoria Combinationis*: in a quite short and apparently little known section of some five pages (Bertrand, 1855, p. 53-58; Trotter, 1957, p. 67-72) Gauss shows how it is possible to find the changes which the most likely values of the unknowns undergo when a new equation (observation) is adjoined and to determine the weights of these new determinations. In other words, and to use more contemporary terminology, he developed an algorithm for sequentially or *recursively* updating the least squares parameter estimates on receipt of additional data.

Bertrand's French translation of Gauss's work appeared in 1855 under the appropriate title "Méthode des Moindres Carrés" and was authorised by Gauss himself.

This translation, together with a commentary provided by the present author, appears in Appendix 2. In addition to its importance in historical terms, Gauss's analysis is interesting because it demonstrates the elegance of his approach and the power of his mind: without the advantages of matrix algebra which, as we shall see, considerably simplify the derivation, Gauss was able to obtain the recursive least squares algorithm with consummate ease.

Gauss's analysis represents the birth of recursive least squares theory; a theory so much ahead of its time that it would lie dormant for almost a century and a half before it was rediscovered on two separate occasions: first by the statistician R.L. Plackett in 1950; and then later and in a more sophisticated form, as the core of the linear filtering and prediction theory evolved by the control and systems theorist R.E. Kalman (1960).

Not surprisingly, perhaps, Plackett's paper went almost unnoticed in the pre-computer age of the early nineteen fifties. Harking back to Gauss, he re-worked the original results in more elegant vector-matrix terms and developed an algorithm for the general case in which additional observations occur in sets $S > 1$. In the present book, like Gauss, we restrict the analysis largely to $S = 1$, although the extension to $S > 1$ is straightforward and is discussed in certain special cases.

Kalman's results, almost certainly obtained without knowledge of either Gauss's or Plackett's prior contributions, were developed within the context of state variable estimation and filter theory, using an argument based on orthogonal projection. Not only were Kalman's results mathematically elegant in providing a computationally straightforward solution to the optimal filtering problem, which had a number of advantages over the earlier Wiener solution (Wiener, 1949), but they also had good potential for practical application. Not surprisingly, therefore, they caused quite a revolution in the automatic control and systems field providing, during the next ten years, a rich source of research material for control and systems analysts. Subsequently, the term *Kalman Filter* has become widely used, not only in the academic and industrial world of automatic control but also in other disciplines such as statistics and economics.

It is now well known that the Kalman filter estimation algorithm can be derived in various ways; via orthogonal projection, as in Kalman's exposition; as well as from the standpoint of maximum likelihood or Bayesian estimation. It can also be developed in various different forms, for application to both discrete (Kalman, 1960) and continuous (Kalman and Bucy, 1961) time-series. But, in all its forms, it has had a profound effect on data processing during the last two decades, being used in applications ranging from trajectory and orbit estimation to the forecasting of economic time-series.

Sprott (1978) has questioned whether the Kalman filter is really a significant 'development' of the Gauss-Plackett recursive least squares algorithm. While it is true that the Gauss-Plackett recursion formulae are an essential aspect of the

Kalman filter equations, it is also clear, as we shall see in this book, that Kalman considerably extended the theory both to allow for the estimation of *time-variable* parameters or *states*, and to handle the analysis of statistically *non-stationary* time-series. Nevertheless the Gauss-Plackett recursion is undoubtedly the central component of the Kalman filter and the basis of most other recursive least squares algorithms. Thus a good understanding of its function in a data processing sense is an essential pre-requisite for the practical application of the algorithm. It is the provision of such understanding, therefore, which is one of the primary aims of this book.

As Gauss pointed out so succinctly in the quotation at the beginning of this chapter, recursive least squares theory is both simple and useful. Here we will exploit this simplicity and take the reader gently through the mysteries of the subject, avoiding wherever possible undue rigour and complexity. In the spirit of Gauss, we will concentrate on mathematical analysis which, while it is often algebraic in form, also has sufficient statistical content to ensure that the reader is fully aware of the important statistical aspects of the results. We will, however, allow ourselves one luxury not available to Gauss and simplify the analysis still further by resort to matrix algebra, assuming that the reader is already acquainted with such analysis; has access to a good text on the subject; or finds that the background notes in Appendix 1 of the book provide sufficient revision.

Finally, to emphasize the practical utility of the various recursive least squares algorithms which will emerge during the analysis, we will provide a number of simulation and practical examples, with applications which range from the man-made world of engineering to the more natural world of ecology and the environment. Many other applications in diverse areas, from economics to hydrodynamics, are discussed in a variety of technical papers produced by the author and his colleagues over the past few years and these are either referred to in the text or listed in the bibliography. Some more recent references are discussed in a short Epilogue at the end of the book.

The text is divided into two major parts: the first is primarily concerned with the estimation of constant or time-variable parameters in general models which are *linear-in-the-parameters*; the second shows how the procedures developed in the first part can be modified to handle the analysis of stochastic time-series and so provide algorithms for the recursive estimation of parameters and states in stochastic dynamic systems. In sympathy with the introductory nature of this book, there has been a conscious attempt to simplify the mathematical analysis as much as possible, particularly in the early chapters, so as to enhance the readability of the book and avoid an overly esoteric presentation. For the reader unfamiliar with some of the mathematics used in the book, Appendix 1 provides background notes, not only on matrix algebra but also on probability and statistics, as well as some very simple concepts in dynamic systems. In all cases, the results in Appendix 1 are chosen