

Junichiro Hagiwara

Time Series Analysis for the State-Space Model with R/Stan

 Springer

Time Series Analysis for the State-Space Model with R/Stan

Junichiro Hagiwara

Time Series Analysis for the State-Space Model with R/Stan

 Springer

Junichiro Hagiwara
Graduate School / Faculty of Information Science
and Technology
Hokkaido University
Hokkaido
Japan

ISBN 978-981-16-0710-3 ISBN 978-981-16-0711-0 (eBook)
<https://doi.org/10.1007/978-981-16-0711-0>

© Springer Nature Singapore Pte Ltd. 2021

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, 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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To my parent—Shizuo and Sachiko.

Preface

This book is written as a reference for those who intend to perform time series analysis using the R language.

A time series is the sequentially obtained data, such as temperature or stock price, and its analysis requires consideration of the relation among data. Various approaches can be used to conduct this analysis; this book describes both deterministic and stochastic methods. We explain the deterministic method based on the moving average from the explorative viewpoints. In contrast, we explain the probabilistic method based on the state-space model, which has gained increasing attention in recent years, for conducting detailed analysis. R, which is a free software used to perform statistical analysis, is a language suitable to perform time series analysis. This book uses R libraries and related software as well as R.

Because this book presents not only basic topics but also advanced ones, a wide range of target readers, from beginners to experts in the field of time series analysis, will be interested in reading it. This book contains introductory chapters for readers who are not necessarily familiar with statistics and the R language.

One feature of this book is its emphasis on practice. Therefore, majority of the algorithm derivations, such as the derivation of Kalman filtering, is summarized in the appendices, whereas the main body focuses on explaining the meaning of the formulas and how to develop codes based on the provided examples. This type of presentation is based on my experience in this field. Even after understanding the literature on time series analysis, I cannot easily identify the appropriate manner in which the contents can be translated to actual codes at times. Although overcoming such difficulties on the reader's own is a reasonable education policy, I hope that this book can contribute to mitigate such difficulties.

The structure of this book is as follows. First, Chaps. 1–4 present the introduction, where we explain the fundamentals of statistics and handling of the time series data using R after introducing the basic concept of time series analysis; then, we present an example of time series analysis using the exploratory method. Next, Chaps. 5–9 present the fundamental analysis conducted using the state-space model. Chapters 5 and 6 describe the basics of the state-space models. Chapter 7 introduces the Wiener filter, which provides the batch solution for a linear Gaussian state-space model, from a historical viewpoint. Chapter 8 describes the Kalman filter, which

provides the sequential solution for a linear Gaussian state-space model, using the R library **d1m**. Chapter 9 describes a set of typically used linear Gaussian state-space models for conducting detailed analysis. Finally, Chaps. 10–12 present the applied analysis based on the state-space model. Chapter 10 describes a batch solution for the general state-space model using the Markov chain Monte Carlo (MCMC) method. This chapter uses the **Stan** software and also explains the forward filtering backward sampling (FFBS) technique to improve the estimation accuracy using the Kalman filter as a component. Chapter 11 describes a particle filter, which provides a sequential solution for the general state-space model. This chapter also explains an auxiliary particle filter and the Rao–Blackwellization technique to improve the estimation accuracy using the Kalman filter as a component. Chapter 12 presents a specialized analysis with respect to the data exhibiting a structural change. This chapter examines the approach in which a horseshoe distribution is used.

As for the notation, this book provides supplementary notes in a shaded box with respect to the relevant topics. As for fonts, the typewriter typefaces, such as `plot()`, boldface, such as **d1m**, and italic shapes, such as *DLM*, are used for providing code description, library names, and important indices, respectively. All the codes and data used in this book are available from the support site https://github.com/hagijyun/Time_Series_Analysis_4SSM_R_Stan.

Finally, this book has been completed with the cooperation of many people. Although I am very sorry that the names of all these people cannot be listed, I would like to express some acknowledgements.

This book has been written based on the usage of various types of software. I would like to specially thank the development team of R and **Stan** and professor Giovanni Petris of the University of Arkansas, who is one of the authors of the **d1m** library. Because I have gained valuable insights by participating in the Japanese translation of the **Stan** manual <https://github.com/stan-ja> and Japan research meeting on particle filters <https://pf.sozolab.jp/>, I am also grateful to all those involved in these projects. The content of this book is also related to the syllabus of my doctoral course; therefore, I would like to thank professor Yasutaka Ogawa, professor Takeo Ohgane, and associate professor Toshihiko Nishimura of the Hokkaido University. I would like to thank professor Motohiro Ishida of the Tokushima University for the opportunity to write this book. Professor Norikazu Ikoma of the Nippon Institute of Technology, Mr. Kentaro Matsuura, and Dr. Hiroki Itô helped to review the Japanese draft of this book, which was useful for improving its contents; I would like to thank them for their contribution. I would also like to thank Springer’s coordinators, Mr. Yutaka Hirachi and Mr. Rammohan Krishnamurthy, for their cooperation in reviewing and publishing this book.

I have written this book in the hope that sharing my experience would be helpful to everyone who intends to perform time series analysis. If there are any errors in this book, it is my responsibility alone; in such a case, I would like to make corrections through the support site.

Contents

1	Introduction	1
1.1	What Is Time Series Analysis?	1
1.2	Two Approaches in Time Series Analysis	3
1.3	Use of R	4
1.3.1	Library in R and External Software	4
1.3.2	Code and Data in This Book	5
1.4	Notation in This Book	5
	References	5
2	Fundamentals of Probability and Statistics	7
2.1	Probability	7
2.2	Mean and Variance	8
2.3	Normal Distribution	9
2.4	Relation Among Multiple Random Variables	11
2.5	Stochastic Process	14
2.6	Covariance and Correlation	14
2.7	Stationary and Nonstationary Processes	18
2.8	Maximum Likelihood Estimation and Bayesian Estimation	19
	References	21
3	Fundamentals of Handling Time Series Data with R	23
3.1	Object for Handling Time Series	23
3.2	Handling of Time Information	26
	Reference	27
4	Quick Tour of Time Series Analysis	29
4.1	Confirmation of the Purpose and Data Collection	29
4.2	Preliminary Examination of Data	31
4.2.1	Plot with Horizontal Axis as Time	31
4.2.2	Histogram and Five-Number Summary	33
4.2.3	Autocorrelation Coefficient	37
4.2.4	Frequency Spectrum	38
4.3	Model Definition	43

- 4.4 Specification of Parameter Values 45
- 4.5 Execution of Filtering, Prediction, and Smoothing 46
- 4.6 Diagnostic Checking for the Results 51
- 4.7 Guideline When Applying the State-Space Model 56
- References 57
- 5 State-Space Model 59**
 - 5.1 Stochastic Model 59
 - 5.2 Definition of State-Space Model 59
 - 5.2.1 Representation by Graphical Model 61
 - 5.2.2 Representation by Probability Distribution 62
 - 5.2.3 Representation by Equation 63
 - 5.2.4 Joint Distribution of State-Space Model 63
 - 5.3 Features of State-Space Model 65
 - 5.4 Classification of State-Space Models 66
 - References 68
- 6 State Estimation in the State-Space Model 69**
 - 6.1 State Estimation Through the Posterior Distribution 69
 - 6.2 How to Obtain the State Sequentially 70
 - 6.2.1 A Simple Example 70
 - 6.2.2 Conceptual Diagram of Recursion 74
 - 6.2.3 Formulation of Filtering Distribution 75
 - 6.2.4 Formulation of Predictive Distribution 79
 - 6.2.5 Formulation of the Smoothing Distribution 81
 - 6.3 Likelihood and Model Selection in the State-Space Model 83
 - 6.4 Treatment of Parameters in the State-Space Model 84
 - 6.4.1 When Parameters are Not Regarded as Random Variables 85
 - 6.4.2 When Parameters are Regarded as Random Variables 85
 - References 87
- 7 Batch Solution for Linear Gaussian State-Space Model 89**
 - 7.1 Wiener Filter 89
 - 7.1.1 Wiener Smoothing 90
 - 7.2 Example: AR(1) Model Case 92
 - References 95
- 8 Sequential Solution for Linear Gaussian State-Space Model 97**
 - 8.1 Kalman Filter 97
 - 8.1.1 Kalman Filtering 98
 - 8.1.2 Kalman Prediction 102
 - 8.1.3 Kalman Smoothing 105
 - 8.2 Example: Local-level Model Case 107
 - 8.2.1 Confirmation of the Purpose and Data Collection 108
 - 8.2.2 Preliminary Examination of Data 108

8.2.3	Model Definition	108
8.2.4	Specification of Parameter Values	110
8.2.5	Execution of Filtering, Prediction, and Smoothing	115
8.2.6	Diagnostic Checking for the Results	121
	References	126
9	Introduction and Analysis Examples of a Well-Known Component Model in the Linear Gaussian State-Space Model	129
9.1	Combination of Individual Models	129
9.2	Local-Level Model	130
9.2.1	Example: Artificial Local-Level Model	131
9.3	Local-Trend Model	133
9.4	Seasonal Model	135
9.4.1	Approach from the Time Domain	136
9.4.2	Approach from the Frequency Domain	137
9.4.3	Example: CO ₂ Concentration in the Atmosphere	140
9.5	ARMA Model	150
9.5.1	Example: Japanese Beer Production	152
9.6	Regression Model	160
9.6.1	Example: Nintendo's Stock Price	161
9.6.2	Example: Flow Data of the Nile (Considering the Rapid Decrease in 1899)	165
9.6.3	Example: Family Food Expenditure (Considering Effects Depending on the Days of the Week)	167
9.7	Supplement to Modeling	175
	References	176
10	Batch Solution for General State-Space Model	179
10.1	MCMC	179
10.1.1	MCMC Fundamentals	180
10.1.2	Notes on Using the MCMC Method	184
10.2	State Estimation with MCMC	185
10.3	Use of Library	187
10.3.1	Various Libraries	187
10.3.2	Example: Artificial Local-Level Model	188
10.4	Estimation Example in General State-Space Model	194
10.5	Technique for Improving Estimation Accuracy	199
10.5.1	Case in Which the Linear Gaussian State-Space Model is Partially Applicable	200
10.5.2	Example: Artificial Local-Level Model	201
10.5.3	Example: Monthly Totals of Car Drivers in the UK Killed or Injured	207
	References	217

- 11 Sequential Solution for General State-Space Model** 219
 - 11.1 Particle Filter 219
 - 11.1.1 Particle Filtering 221
 - 11.1.2 Particle Prediction 225
 - 11.1.3 Particle Smoothing 226
 - 11.2 State Estimation with Particle Filter 229
 - 11.2.1 Example: Artificial Local-Level Model 229
 - 11.2.2 Attention to Numerical Computation 241
 - 11.3 Use of Library 246
 - 11.4 Estimation Example in General State-Space Model 247
 - 11.4.1 Example: A Well-Known Nonlinear Benchmark Model 247
 - 11.4.2 Application of a Particle Filter 250
 - 11.5 Technique for Improving Estimation Accuracy 253
 - 11.5.1 Auxiliary Particle Filter 254
 - 11.5.2 Case in Which the Linear Gaussian State-Space Model is Partially Applicable 266
 - References 273
- 12 Example of Applied Analysis in General State-Space Model** 277
 - 12.1 Consideration of Structural Change 277
 - 12.2 Approach Using a Kalman Filter (Known Change Point) 279
 - 12.2.1 Time-Invariant Model Studied thus Far 279
 - 12.2.2 Utilizing Prior Information in the Linear Gaussian State-Space Model 279
 - 12.2.3 Numerical Result 280
 - 12.3 Approach Using MCMC (Unknown Change Point) 283
 - 12.3.1 Time-Invariant Model Studied thus Far 283
 - 12.3.2 Use of a Horseshoe Distribution in the General State-Space Model 284
 - 12.3.3 Numerical Result 286
 - 12.4 Approach Using a Particle Filter (Unknown Change Point) 292
 - 12.4.1 Time-Invariant Model Studied thus Far 292
 - 12.4.2 Use of a Horseshoe Distribution in the General State-Space Model 294
 - 12.4.3 Numerical Results 294
 - 12.5 Real-Time Detection for an Unknown Change Point 299
 - References 300
- Appendix A: Library in R and External Software** 303
- Appendix B: Library dlm** 305
- Appendix C: Supplement on Conditional Independence in the State-Space Model** 313

Appendix D: Symbol Assignment in the Linear Gaussian State-Space Model	315
Appendix E: Algorithm Derivation	317
Appendix F: Execution of Particle Filtering with Library	335
Index	343