

Statistics and Econometrics for Finance

Yong Zeng
Shu Wu *Eds.*

State-Space Models

Applications in Economics
and Finance

 Springer

Statistics and Econometrics for Finance

Series Editors

Ruppert, D.

Fan, J.

Renault, E.

Zivot, E.

For further volumes:

<http://www.springer.com/series/10377>

Yong Zeng • Shu Wu
Editors

State-Space Models

Applications in Economics and Finance

 Springer

Editors

Yong Zeng
Department of Mathematics and Statistics
University of Missouri at Kansas City
Kansas City, Missouri, USA

Shu Wu
Department of Economics
The University of Kansas
Lawrence, Kansas, USA

ISBN 978-1-4614-7788-4 ISBN 978-1-4614-7789-1 (eBook)
DOI 10.1007/978-1-4614-7789-1
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013943592

© Springer Science+Business Media New York 2013

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. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

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.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Since the seminal papers of Kalman (1960, 1961) and the early development in the field of engineering, state-space models have become an increasingly important tool for research in finance and economics in recent years. This book is a collection of contributed papers that reflect this new phase of theoretical developments of state-space models and their applications in economics and finance. We hope that the breadth of research shared in this volume will serve as an inspiration and a valuable reference for future users of state-space models in these two areas.

A generic state-space model consists of two equations. One describes how observable economic variables relate to potentially unobservable *state variables*, and the other describes how state variables evolve over time. Both state and observable variables can be either discrete- or continuous-time stochastic processes.

This book is divided into four parts. In Parts **I** and **II**, we mainly consider discrete-time state-space models. Let Y_t be an $n \times 1$ observable variable and X_t be a $k \times 1$ state variable. Then, a state-space model can be written as follows:

$$Y_t = f(X_t, \theta, \varepsilon_t) \tag{1}$$

$$X_{t+1} = g(X_t, \theta, \eta_{t+1}) \tag{2}$$

where θ is an $m \times 1$ vector of model parameters, and ε_t and η_t are independent and identically distributed random shocks (or noises). $f(\cdot)$ and $g(\cdot)$ are $n \times 1$ and $k \times 1$ (non)linear functions, respectively. Because of its flexibility, many discrete-time models in economics and finance can be represented in this state-space form. These models include autoregressive moving average (ARMA) models, regression models with time-varying coefficients, dynamic factor models (DFM), models with stochastic volatility (SV), regime-switching models, and hidden Markov models (HMM). In these models, the parameter θ is typically unknown. Therefore, one central question in many applications is to estimate θ and to conduct statistical inferences of the model (e.g., hypothesis testing). Researchers are also highly interested in obtaining unbiased and efficient estimates of the underlying state variables X_t .

In Parts **III** and **IV**, we include models where the state variable, $X(t)$, can be a continuous-time finite-state Markov chain as that in hidden Markov or

regime-switching models of Chaps. 8–11. $X(t)$ can also be a general continuous-time Markov process described by a stochastic differential equation as those in Chaps. 12–15. The Markov processes include a geometric Brownian motion (GBM) and a jump-diffusion process with regime switching among others. The observation process, $Y(t)$, can be a continuous-path process (Chaps. 10 and 11), an equally spaced time series (Chap. 14), or an irregularly spaced point process (Chaps. 8, 13, and 15). These models have found many applications in finance such as pricing of credit risk, optimal trading rules and hedging, optimal annuity purchasing or dividend policies, optimal learning in financial markets, and modeling (ultra) high-frequency data.

Below, we provide a more detailed description of each part.

Part I includes three chapters on *Particle Filtering and Parameter Learning in Nonlinear State-Space Models*. The introduction of particle filters has had a major impact on the development of nonlinear and non-Gaussian state-space models. This technique has expanded the range of practical applicability of state-space models to cases with high-dimension state space.

In Chap. 1, Tze Leung Lai and Vibhav Bukkapatanam provide a review of the estimation of the latent state variables using particle filters with known or unknown model parameters. They also present a new adaptive particle filter that uses a computationally efficient Markov Chain Monte Carlo estimate of the posterior distribution of the state-space model parameters in conjunction with sequential state estimation. They describe several applications in finance and economics, including frailty models of portfolio default probabilities, SV models with contemporaneous price and volatility jumps, and hidden Markov models for high-frequency transaction data.

In Chap. 2, Maria Paula Rios and Hedibert Freitas Lopes explore kernel smoothing and conditional sufficient statistics extensions of the auxiliary particle filters (Pitt and Shepard 1999) and bootstrap filters (Gordon, Salmond and Smith 1993). Using simulated data from SV models with Markov switching, they show that the Liu-West particle filter degenerates and has the largest Monte Carlo error, while their auxiliary particle filter extended with sufficient statistics (APF + SS) has a much better performance. Their APF + SS filter takes advantage of recursive sufficient statistics that are sequentially tracked and whose behavior resembles that of a latent state with conditionally deterministic updates. They also assess the performance of the APF + SS filter in sequential estimation in examples with real data.

In Chap. 3, Alexandre J. Chorin, Matthias Morzfeld, and Xuemin Tu review the implicit particle filter. The key idea is to concentrate the particles on the high-probability regions of the target probability density function (pdf) so that the number of particles required for a good approximation of the pdf remains manageable even if the state space has high dimensions. They explain how this idea is implemented, discuss special cases of practical importance, and show the relations of the implicit particle filter with other data assimilation methods. They further illustrate the method with examples such as SV, stochastic Lorenz attractor, stochastic

Kuramoto–Sivashinsky equation, and data assimilations.

Part II includes four chapters on the application of *Linear State-Space Models in Macroeconomics and Finance*.

In Chap. 4, Yulei Luo, Jun Nie, and Eric Young explicitly solve a linear-quadratic macroeconomic model under model uncertainty (due to concerns of model misspecification) and state uncertainty (due to limited information constraint). They show that the model can be mapped to a state-space representation that can be used to quantify the key parameters of model uncertainty. They demonstrate through examples how this framework can be used to study a range of interesting questions in macroeconomics and international finance such as explaining current account dynamics and resolving the international consumption puzzle.

In Chap. 5, Pym Manopimoke estimates a state-space model of the inflation dynamics in Hong Kong. The model allows her to decompose Hong Kong inflation into a stochastic trend and a stationary cycle component that can be driven by both domestic and foreign economic variables such as output gaps. This empirical model is consistent with economic theories of inflation and output, offering new insight into the determination of trend and cyclical inflation in Hong Kong. This is an example of the power of state-space models in empirical macroeconomic research.

In recent years, vector autoregression models (VARs) have become a primary tool for investigating dynamic relationship between multiple economic variables. One challenge, however, is that such relationships are often evolving over time as a result of shifts in government policies or structural changes in the economy. In Chap. 6, Taeyoung Doh and Michael Connolly show that the state-space representation is a useful tool to estimate VARs with time-varying coefficients and/or SV. They show that these models can better capture the changing relationships between important macroeconomic variables.

Chapter 7 is an application to finance. Jun Ma and Mark Wohar use a state-space model to address one important issue regarding sources of stock market volatility. They argue that the existing empirical studies have focused on point estimation and lack robust statistical inference. The authors show that the small signal-to-noise ratio has made the market data contain too little useful information for researchers to reach robust conclusions about the relative importance of different sources of stock market volatility.

Part III includes five chapters on *Hidden Markov Models (HMM)*, *Regime Switching*, and *Mathematical Finance*.

Chapters 8 and 9 are on hidden Markov models and their applications to finance. In Chap. 8, Robert Elliott and Tak Kuen Siu discuss an intensity-based model of portfolio credit risk using a collection of hidden Markov-modulated single jump processes. The model is a dynamic version of a frailty model casted in state-space form, able to describe dependent default risks among firms that are exposed to a common hidden dynamic frailty factor. The authors develop filtering equations and filter-based estimates of the model in recursive forms. They also obtain the joint default probability of reference entities in a credit portfolio as well as the

variance dynamics for both observations and hidden states. In Chap. 9, Xiaojing Xi and Rogemar Mamon develop a weak hidden Markov model (WHMM) for the term structure of interest rates where the means and volatilities of bond yields are governed by a second-order Markov chain in discrete time. The authors use the multivariate filtering technique in conjunction with the EM algorithm to estimate the model parameters. They assess the goodness of fit of the model based on out-of-sample forecasts and apply AIC to determine the optimal number of regimes in their model. They apply the model to a data set of daily Treasury yields in the USA. The empirical results show that their WHMM outperforms the standard HMM in terms of out-of-sample forecasts.

Chapters 10 and 11 are applications of regime-switching models to insurance risk and optimal trading rule. Models with regime switching usually don't have analytical solutions to the associated stochastic control problems. In Chap. 10, Zhuo Jin and George Yin discuss numerical methods for solving stochastic optimization problems involving regime-switching models. They propose a numerical solution to the system of HJB equations based on Markov chain approximation. They show how these regime-switching models can be applied to analyze optimal annuity purchasing and optimal dividend payment strategy problems. In Chap. 11, Eunju Sohn and Qing Zhang study an optimal trading rule problem where the underlying asset price is governed by a mean-reverting process with regime switching. The investor's objective is to buy and sell the asset so as to maximize the overall return. The authors consider the case in which the jump rates of the Markov chain can go to infinite. They study the asymptotic properties of the limit value functions and establish a limiting problem which is easier to solve. They show that the solution to the limiting problem can be used to construct a trading rule that is nearly optimal.

In Chap. 12, Mingming Wang and Allanus Tsoi discuss Constant Proportion Portfolio Insurance (CPPI) problem with jump diffusion. They also consider the associated problem of hedging using both the PDE/PIDE and martingale approaches. In particular, they consider the mean-variance hedging problem when the contingent claim is a function of the CPPI portfolio value.

Part IV includes three chapters on *Nonlinear State-Space Models for High-Frequency Financial Data*.

In Chap. 13, combining classical Kyle and Glosten–Milgrom models, Yoonjung Lee proposes a new state-space modeling framework under asymmetric information. The model is able to describe the interactions among some important variables in financial markets such as the price impact of a trade, the duration between trades, and the degree of information asymmetry. In the model, a private signal is partially revealed through trades, while new public information arrives continuously at the market. In order to set a competitive price that rationally incorporates these two sources of information, the market maker utilizes Bayesian learning. The author derives the corresponding nonlinear filtering equation using anticipative Girsanov transformation. She further proves the existence and uniqueness of the ask and bid prices using an SPDE approach. The pricing rule depends on the actual sequence of order arrivals, not just the total number of buy/sell orders. The price impact of a

trade tends to decrease when the duration between trades gets longer. The speed at which the information gets incorporated into the price depends on the quality of the private signal and the trading rate of informed traders.

Chapter 14 is concerned with volatility estimation and prediction. A popular approach is to use the high-frequency data to estimate volatilities and then fit a low-frequency AR volatility model for forecasting. While the empirical performance of this approach is good, there is a lack of theoretical foundation. In this chapter, Yazhen Wang and Xin Zhang show that, for rather general underlying price and volatility processes, the realized volatility estimators approximately follow a heterogeneous autoregressive model, hence providing theoretical justifications of the popular approach. An important feature of the model is that the two- or multi-scaled realized volatility estimators employed are based on a state-space model, where the prices from high-frequency transactions may include market microstructure noise.

Chapter 15 is concerned with estimating models for ultra-high frequency data. The class of models has a random-arrival-time state-space form that explicitly accommodates market microstructure noises in asset price. Although the model is able to capture stylized facts of tick data, the nonlinear state-space model structure makes parameter estimation a challenge. Cai Zhu and James Huang apply particle Markov Chain Monte Carlo (PMCMC) method to estimate a couple models when the underlying intrinsic value processes follow a GBM or a jump-diffusion process. They show that the PMCMC method is able to yield reasonable estimates of the model parameters and further discuss numeric methods that are able to enhance the efficiency of the algorithm.

We would like to express our gratitude to all the contributors of the book chapters for their efforts in making their research accessible to a wide range readers. We hope the book can lead to more interdisciplinary research among economists, mathematicians, and statisticians. We also would like to thank Yaozhong Hu of University of Kansas, Neng Wang of Columbia University, and Zhenxiao Wu of National University of Singapore for their generous help during the preparation of this book.

We want to thank Brian Foster, a former Springer editor, for his enthusiasm and help in the early stage of this book project. We also want to thank Hannah Bracken, Marc Strauss, Nicolas Philipson, and William Curtis of Springer New York. Their continuing support and commitment throughout the project are highly appreciated. We owe a special thanks to Hannah Bracken for her superb editorial assistance.

Yong Zeng's research is supported by National Science Foundation under the grants of TG-DMS100019 and DMS-1228244 and a grant from the University of Missouri Research Board in 2011. He gratefully acknowledges those grant supports. Finally, we would like to thank our families for their continuous support, which makes the completion of this project possible.

Kansas City, MO
Lawrence, KS

Yong Zeng
Shu Wu

Contents

Part I Particle Filtering and Parameter Learning in Nonlinear State-Space Models

1	Adaptive Filtering, Nonlinear State-Space Models, and Applications in Finance and Econometrics	3
	Tze Leung Lai and Vibhav Bukkapatnam	
1.1	Introduction	3
1.2	Particle Filters in Nonlinear State-Space Models	4
	1.2.1 Bootstrap Filter	4
	1.2.2 Auxiliary Particle Filter	5
	1.2.3 Residual Bernoulli Resampling	6
1.3	Particle Filters with Sequential Parameter Estimation	6
	1.3.1 Liu and West's Filter	7
	1.3.2 Storvik's Filter	8
	1.3.3 Particle Learning	8
	1.3.4 Particle MCMC	9
1.4	A New Approach to Adaptive Particle Filtering	9
	1.4.1 A New MCMC Approach to Sequential Parameter Estimation	10
	1.4.2 Adaptive Particle Filters and Asymptotic Theory	12
1.5	Applications in Finance and Economics	14
	1.5.1 Frailty Models for Corporate Defaults	14
	1.5.2 Stochastic Volatility with Contemporaneous Jumps	17
	1.5.3 State-Space Models for High-Frequency Transaction Data	19
	1.5.4 Other Applications in Finance and Economics	20
1.6	Conclusion	20
	References	20

2	The Extended Liu and West Filter: Parameter Learning in Markov Switching Stochastic Volatility Models	23
	Maria Paula Rios and Hedibert Freitas Lopes	
2.1	Introduction	23
2.1.1	Volatility Models	24
2.1.2	Particle Filters: A Brief Review	26
2.2	Particle Filters with Parameter Learning	28
2.2.1	Kernel Smoothing	28
2.2.2	Sufficient Statistics	30
2.3	Analysis and Results: Simulation Study	32
2.3.1	Simulated Data	33
2.3.2	Exact Estimation Path	33
2.3.3	Estimate Evaluation	36
2.3.4	Economic Insight	48
2.3.5	Robustness	50
2.4	Analysis and Results: Real Data Applications	52
2.4.1	IBOVESPA	53
2.4.2	S&P 500	55
2.5	Conclusions	59
	References	60
3	A Survey of Implicit Particle Filters for Data Assimilation	63
	Alexandre J. Chorin, Matthias Morzfeld, and Xuemin Tu	
3.1	Introduction	63
3.2	Implicit Particle Filters	65
3.2.1	Linear Observation Function and Gaussian Noise	67
3.2.2	Sparse Observations	68
3.2.3	Models with Partial Noise	69
3.2.4	Combined State and Parameter Estimation	70
3.3	Implementations of the Implicit Particle Filter	71
3.3.1	Solution of the Implicit Equation via Quadratic Approximation	71
3.3.2	Solution of the Implicit Equation via Random Maps	72
3.4	Comparison with Other Sequential Monte Carlo Schemes	73
3.4.1	Comparison with the SIR Filter	74
3.4.2	Comparison with Optimal Importance Function Filters	75
3.4.3	Comparison with the Kalman Filter and with Variational Data Assimilation Methods	76
3.5	Applications	77
3.5.1	A Simple Example	77
3.5.2	Stochastic Volatility Model	78
3.5.3	The Stochastic Lorenz Attractor	78
3.5.4	The Stochastic Kuramoto–Sivashinsky Equation	81
3.5.5	Application to Geomagnetic Data Assimilation	83

3.5.6 Assimilation of Ocean Color Data from NASA’s SeaWiFS Satellite 84

3.6 Conclusion 85

References 86

Part II Linear State-Space Models in Macroeconomics and Finance

4 Model Uncertainty, State Uncertainty, and State-Space Models 91
 Yulei Luo, Jun Nie, and Eric R. Young

4.1 Introduction 91

4.2 Linear-Quadratic-Gaussian State-Space Models 92

4.3 Incorporating Model Uncertainty and State Uncertainty 94

4.3.1 Introducing Model Uncertainty 94

4.3.2 Introducing State Uncertainty 95

4.4 Applications 98

4.4.1 Explaining Current Account Dynamics 98

4.4.2 Resolving the International Consumption Puzzle 101

4.4.3 Other Possible Applications 104

4.4.4 Quantifying Model Uncertainty 105

4.4.5 Discussions: Risk-Sensitivity and Robustness Under Rational Inattention 107

4.5 Conclusions 109

Appendix 109

A.1 Solving the Current Account Model Explicitly Under Model Uncertainty 109

References 111

5 Hong Kong Inflation Dynamics: Trend and Cycle Relationships with the USA and China 113
 Pym Manopimoke

5.1 Introduction 113

5.2 Literature Review 115

5.3 Model Specification 117

5.4 Empirical Results 123

5.5 Conclusion 130

References 131

6 The State Space Representation and Estimation of a Time-Varying Parameter VAR with Stochastic Volatility 133
 Taeyoung Doh and Michael Connolly

6.1 Introduction 133

6.2 State Space Representation and Estimation of VARs 134

6.2.1 State Space Representation 134

6.2.2 Estimation of VARs 135

6.3 Application: A Time-Varying Parameter VAR with Stochastic Volatility for Three US Macroeconomic Variables 139

6.3.1 Priors 139

6.3.2	Posterior Simulation	140
6.3.3	Posterior Estimates of Time-Varying Trends and Volatility	140
6.4	Conclusion	144
	References	145
7	A Statistical Investigation of Stock Return Decomposition Based on the State-Space Framework	147
	Jun Ma and Mark E. Wohar	
7.1	Introduction	147
7.2	VAR Variance Decomposition of the Stock Prices	151
7.3	The State-Space Model for Decomposing Stock Prices	154
7.4	The Weak Identification and the Corrected Inference	160
7.5	Conclusion	163
	References	164
 Part III Hidden Markov Models, Regime-Switching, and Mathematical Finance		
8	A HMM Intensity-Based Credit Risk Model and Filtering	169
	Robert J. Elliott and Tak Kuen Siu	
8.1	Introduction	169
8.2	A HMM Frailty-Based Default Model	171
8.3	Filtering Equations for the Hidden Dynamic Frailty Factor	173
8.4	A Robust Filter-Based EM Algorithm	177
8.5	Variance Dynamics	180
8.6	Default Probabilities	181
8.7	Conclusion	182
	References	183
9	Yield Curve Modelling Using a Multivariate Higher-Order HMM ...	185
	Xiaoqing Xi and Rogemar Mamon	
9.1	Introduction	185
9.2	Filtering and Parameter Estimation	188
9.3	Implementation	193
9.4	Forecasting and Error Analysis	197
9.5	Conclusion	202
	References	202
10	Numerical Methods for Optimal Annuity Purchasing and Dividend Optimization Strategies under Regime-Switching Models: Review of Recent Results	205
	Zhuo Jin and George Yin	
10.1	Introduction	205

10.2	Optimal Annuity-Purchasing Strategies	207
10.2.1	Motivation	207
10.2.2	Formulation	208
10.2.3	Constant Hazard Rate	210
10.2.4	General Hazard Rate	211
10.2.5	Examples	212
10.3	Optimal Dividend Payment Policies	214
10.3.1	Motivation	214
10.3.2	Formulation	215
10.3.3	Algorithm	217
10.3.4	Convergence	219
10.3.5	Examples	220
10.4	Concluding Remarks	223
	References	224
11	Trading a Mean-Reverting Asset with Regime Switching: An Asymptotic Approach	227
	Eunju Sohn and Qing Zhang	
11.1	Introduction	227
11.2	Problem Formulation	229
11.3	Properties of the Value Functions	233
11.4	Asymptotic Properties	235
11.5	Further Approximations	238
11.6	A Numerical Example	239
11.7	Concluding Remarks	240
	Appendix	242
	References	244
12	CPPI in the Jump-Diffusion Model	247
	Mingming Wang and Allanus Tsoi	
12.1	Introduction	247
12.2	The Jump-Diffusion Model	248
12.2.1	Density	249
12.2.2	Martingale Measure	251
12.3	The CPPI Strategies	253
12.3.1	The constant multiple case	253
12.3.2	The Time-Varying Multiple Case	259
12.4	The CPPI Portfolio as a Hedging Tool	259
12.4.1	PIDE Approach	260
12.4.2	Martingale Approach	264
12.5	Mean-Variance Hedging	271
12.5.1	The Idea	271
12.5.2	The Problem	273
12.6	Conclusion	275
	References	275

Part IV Nonlinear State-Space Models for High Frequency Financial Data

13	An Asymmetric Information Modeling Framework for Ultra-High Frequency Transaction Data: A Nonlinear Filtering Approach	279
	Yoonjung Lee	
13.1	Introduction	279
13.2	The Model	283
13.2.1	The Information Structure Dynamics	283
13.2.2	Informed Traders' Signal Extraction	285
13.2.3	Order Arrivals	286
13.3	Bayesian Updating of the Market Maker's Beliefs via Filtering	287
13.3.1	Construction of a Reference Measure	289
13.3.2	Filtering Equation	291
13.3.3	Uniqueness of the System	292
13.4	Key Implications of the Model	293
13.4.1	The Quality of the Signal	293
13.4.2	Informed Traders' Trading Rate	294
13.4.3	The Price Impact of a Trade	294
13.5	Parameter Estimation	295
13.5.1	Maximum Likelihood Estimation	296
13.5.2	Parameter Estimation for Simulated Data	297
13.6	Conclusion	298
	Appendix	300
	References	308
14	Heterogenous Autoregressive Realized Volatility Model	311
	Yazhen Wang and Xin Zhang	
14.1	Introduction	311
14.2	High-Frequency Financial Data and Price Model	312
14.3	GARCH and Stochastic Volatility Approximations to the Price Model	313
14.4	The HAR Model for Volatility Processes	314
14.5	The HAR Model for Realized Volatilities	315
14.6	The Temporal Aggregation of AR Processes	317
	References	320
15	Parameter Estimation via Particle MCMC for Ultra-High Frequency Models	321
	Cai Zhu and Jian Hui Huang	
15.1	Introduction	321
15.2	The Model	323
15.2.1	Trading Times	324
15.2.2	Micro-Structure Noise	324
15.2.3	Intrinsic Value Processes	326
15.3	Estimation Method	327
15.3.1	Likelihood Calculation via Simulation	327

- 15.3.2 Importance Sampling 329
- 15.3.3 Sequential Importance Sampling: Particle Filtering 330
- 15.3.4 Particle MCMC 331
- 15.4 Simulation and Empirical Studies 332
 - 15.4.1 Variance Reduction Effect of Particle Filtering Method 332
 - 15.4.2 Simulation Study: GBM Case 333
 - 15.4.3 Comparison Algorithm Under Trading Rules with 1/8 and 1/100 Tick Size 336
 - 15.4.4 Simulation Study: Jump-Diffusion Case 338
 - 15.4.5 Real Data Application 341
- 15.5 Conclusion 342
- References 343

- Index** 345

List of Contributors

Vibhav Bukkapatnam

Department of Management Science and Engineering, Stanford University,
Stanford, CA, USA, e-mail: vibhav@stanford.edu

Alexandre J. Chorin

Department of Mathematics, University of California Lawrence, Berkeley National
Laboratory, Berkeley, CA, USA,
e-mail: chorin@math.berkeley.edu

Michael Connolly

Federal Reserve Bank of Kansas City, Kansas City, MO, USA,
e-mail: Michael.Connolly@kc.frb.org

Taeyoung Doh

Federal Reserve Bank of Kansas City, Kansas City, MO, USA,
e-mail: Taeyoung.Doh@kc.frb.org

Robert J. Elliott

Haskayne School of Business, University of Calgary, Calgary, AB, Canada
School of Mathematical Sciences, University of Adelaide, Adelaide, SA, Australia
Centre for Applied Financial Studies, University of South Australia, Adelaide, SA,
Australia, e-mail: relliott@ucalgary.ca

Jian Hui Huang

Department of Applied Mathematics, The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong, e-mail: majhuang@inet.polyu.edu.hk

Zhou Jin

Center for Actuarial Studies, Department of Economics, The University of
Melbourne, VIC, Australia, e-mail: zjin@unimelb.edu.au

Tze Leung Lai

Department of Statistics, Stanford University, Stanford, CA, USA,
e-mail: lait@stanford.edu

Yoonjung Lee

Boston, MA, USA, e-mail: yoonyung.lee.lin@gmail.com

Hedibert Freitas Lopes

Booth School of Business, University of Chicago, Chicago, IL, USA,
e-mail: hlopes@chicagobooth.edu

Yulei Luo

Department of Economics, The University of Hong Kong, Pokfulam, Hong Kong, China, e-mail: ylo@econ.hku.hk

Jun Ma

Department of Economics, Finance and Legal Studies, University of Alabama, Tuscaloosa, AL, USA, e-mail: jma@cba.ua.edu

Rogemar Mamon

Department of Statistics and Actuarial Sciences, University of Western Ontario, London, ON, Canada, e-mail: rmamon@stats.uwo.ca

Pym Manopimoke

Department of Economics, University of Kansas, Lawrence, KS, USA,
e-mail: pymm@ku.edu

Matthias Morzfeld

Lawrence Berkeley National Laboratory, Berkeley, CA, USA,
e-mail: mmo@math.lbl.gov

Jun Nie

Federal Reserve Bank of Kansas City, Kansas City, MO, USA,
e-mail: jun.nie@kc.frb.org

Maria Paula Rios

Booth School of Business, University of Chicago, Chicago, IL, USA,
e-mail: maria@chicagobooth.edu

Tak Kuen Siu

Cass Business School, City University London, London, UK,
e-mail: Ken.Siu.1@city.ac.uk

Eunju Sohn

Department of Science and Mathematics, Columbia College of Chicago, Chicago, USA, e-mail: esohn@colum.edu

Allanus Tsoi

Department of Mathematics, University of Missouri, Columbia, MO, USA,
e-mail: tsoia@missouri.edu

Xuemin Tu

Department of Mathematics, University of Kansas, Lawrence, KS, USA,
e-mail: xtu@math.ku.edu

Mingming Wang

School of Insurance, The University of International Business and Economics, Beijing, China, e-mail: wiryming@gmail.com

Yazhen Wang

Department of Statistics, University of Wisconsin-Madison, Madison, WI, USA, e-mail: yzwang@stat.wisc.edu

Xiaoqing Xi

Department of Applied Mathematics, University of Western Ontario, London, ON, Canada, e-mail: xxi2@uwo.ca

Mark E. Wohar

Department of Economics, University of Nebraska at Omaha, Omaha, NE, USA, e-mail: mwohar@unomaha.edu

George Yin

Department of Mathematics, Wayne State University, Detroit, MI, USA, e-mail: gyin@math.wayne.edu

Eric R. Young

Department of Economics, University of Virginia, Charlottesville, VA, USA, e-mail: ey2d@virginia.edu

Qing Zhang

Department of Mathematics, University of Georgia, Athens, GA, USA, e-mail: qingz@math.uga.edu

Xin Zhang

Department of Statistics, University of Wisconsin-Madison, Madison, WI, USA, e-mail: zhangxin@stat.wisc.edu

Cai Zhu

Department of Finance, Business School, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, e-mail: czhuaa@ust.hk

Part I
Particle Filtering and Parameter Learning
in Nonlinear State-Space Models

Chapter 1

Adaptive Filtering, Nonlinear State-Space Models, and Applications in Finance and Econometrics

Tze Leung Lai and Vibhav Bukkapatnam

1.1 Introduction

The Kalman filter, which is applicable to linear Gaussian models, and its modifications such as extended Kalman filters, Gaussian sum filters, and unscented Kalman filters for nonlinear state-space models are widely used in engineering. Without relying on local linearization techniques or functional approximations, particle filters are able to handle a large class of nonlinear non-Gaussian state-space models and have become increasingly popular in engineering applications in the past decade. Filtering in state-space models involves sequential computation of the posterior distribution of the latent state x_t given observations y_1, \dots, y_t . Smoothing involves the estimation of the hidden state x_t given observations y_1, \dots, y_n , with $1 \leq t \leq n$. More details are given in Sect. 1.2.

State-space models typically involve unknown parameters that have to be estimated from the data by either maximum likelihood or Bayesian methods. Replacing these unknown parameters in a particle filter by their sequential estimates leads to an adaptive particle filter; see Liu and West (2001) in [31], Storvik (2002) in [38], Carvalho et al. (2010) in [9], Polson et al. (2008) in [35], and Andrieu et al. (2010) in [2]. Section 1.3 reviews existing methods for sequential parameter estimation in state-space models. Section 1.4 describes a new adaptive filter that combines a novel Markov Chain Monte Carlo (MCMC) scheme for sequential parameter estimation with an efficient particle filter to estimate the state x_t . An important advantage of the new approach is that it yields a consistent estimate of the Monte Carlo standard error.

T.L. Lai (✉)

Department of Statistics, Stanford University, Stanford, CA, USA

e-mail: lait@stanford.edu

V. Bukkapatnam

Department of Management Science and Engineering, Stanford University, Stanford, CA, USA

e-mail: vibhav@stanford.edu

Particle filters have powerful and far-reaching applications in state-space models in finance and econometrics, some of which are described in Sect. 1.5. Section 1.6 gives some concluding remarks.

1.2 Particle Filters in Nonlinear State-Space Models

A general state-space model, also called a hidden Markov model (HMM), is defined by the evolution and observation density functions

$$\begin{aligned} x_t | x_{t-1} &\sim f(x_t | x_{t-1}, \theta) \\ y_t | x_t &\sim g(y_t | x_t, \theta) \end{aligned} \quad (1.1)$$

with respect to measures μ and ν , respectively, where θ is a vector of parameters of the model. In the Bayesian formulation, the initial state x_0 has prior density $f(x_0 | \theta)$, and θ has the prior density $\pi(\theta)$ with respect to some measure on the parameter space. In the HMM, x_t is the latent state and y_t is the observed data at time t . The filtering problem is to sequentially estimate the posterior distribution $p(x_t, \theta | \mathcal{Y}_t)$, where $\mathcal{Y}_t = (y_1, y_2, \dots, y_t)$ is the set of observations up to time t . Particle filters approximate this posterior density by using a set of particles $(x_t, \theta)^{(i)}$ with weights $\tilde{w}_t^{(i)}$ ($i = 1, 2, \dots, N$) summing to 1, so that

$$p^N(x_t, \theta | \mathcal{Y}_t)(\cdot) = \sum_{i=1}^N \tilde{w}_t^{(i)} \delta_{(x_t, \theta)^{(i)}}(\cdot) \quad (1.2)$$

in which δ_z is the Dirac delta function with point mass at z . We now describe different approaches for sampling x_s ($s \leq t$) sequentially to form the particle filter in (1.2) when θ is fixed in advance. The weights $\tilde{w}_t^{(i)}$ in (1.2) can be converted to $w_t^{(i)} = 1/N$ by a resampling step that samples with replacement N particles from $\{(x_t, \theta)^{(i)} : 1 \leq i \leq N\}$ with respective weights $\tilde{w}_t^{(i)}$, as in the original proposal by Gordon et al. (1993) in [23] reviewed below. This is called *bootstrap resampling*. An alternative resampling scheme is introduced in Sect. 1.2.3. Resampling can help to mitigate the degeneracy of particles, which will be discussed in Sects. 1.3 and 1.4. Since resampling is optional and also can be used occasionally, we do not include it in the description of the basic algorithm in Sect. 1.2.1. Moreover, we assume the parameter vector to be known and therefore omit it in the rest of this section.

1.2.1 Bootstrap Filter

Originally proposed by Gordon et al. (1993) in [23], the bootstrap filter chooses samples from the prior distribution of the states. The Bayesian update equation for the posterior in a general filter can be written as

$$p(x_t|\mathcal{Y}_t) = \frac{g(y_t|x_t)p(x_t|\mathcal{Y}_{t-1})}{\int g(y_t|x'_t)p(x'_t|\mathcal{Y}_{t-1})d\mu(x'_t)} \quad (1.3)$$

where

$$p(x_t|\mathcal{Y}_{t-1}) = \int f(x_t|x_{t-1})p(x_{t-1}|\mathcal{Y}_{t-1})d\mu(x_{t-1}) \quad (1.4)$$

The bootstrap filter samples from the filtering distribution $p(x_t|\mathcal{Y}_{t-1})$ by first propagating the particles which approximate $p(x_{t-1}|\mathcal{Y}_{t-1})$ using the evolution density $f(x_t|x_{t-1})$, and reweighing the particles thus obtained by the likelihood ratio weights, as summarized in Algorithm 1.

Algorithm 1 : Bootstrap Filter

Initialize particles $\{x_0^{(i)}\}_{i=1}^N$ and the corresponding weights $\{\tilde{w}_0^{(i)} = 1/N\}_{i=1}^N$
for $t=1, 2, \dots, T$ **do**
 for $i=1, 2, \dots, N$ **do**
 a) Propagate particle $x_{t-1}^{(i)}$ to $x_t^{(i)}$ using the evolution density $f(x_t|x_{t-1}^{(i)})$
 b) Update particle weights according to $\tilde{w}_t^{(i)} \propto w_{t-1}^{(i)} g(y_t|x_t^{(i)})$
 end for
end for

1.2.2 Auxiliary Particle Filter

The auxiliary particle filter proposed by Pitt and Shephard (1999) in [33] generates samples from the filtering distribution with density function

$$p(x_t, x_{t-1}|\mathcal{Y}_t) \propto p(x_t|y_t, x_{t-1})p(y_t|x_{t-1})p(x_{t-1}|\mathcal{Y}_{t-1}) \quad (1.5)$$

Particles approximating $p(x_{t-1}|\mathcal{Y}_{t-1})$ are first resampled using weights proportional to the predictive density $p(y_t|x_{t-1})$, and the resampled particles are propagated forward using $p(x_t|y_t, x_{t-1})$. This is convenient only if $p(x_t|y_t, x_{t-1})$ is not difficult to sample from and $p(y_t|x_{t-1})$ is easily available, which is often not the case. Accordingly Pitt and Shephard in [33] suggest to replace $p(y_t|x_{t-1})$ by $p(y_t|\lambda(x_{t-1}))$, in which $\lambda(x_{t-1})$ is the mean, median, or mode of the distribution of x_t given x_{t-1} , and to propagate the resampled particles by sampling from the proposal density $f(x_t|x_{t-1})$, instead of directly from $p(x_t|x_{t-1}, y_t)$. The method is summarized in Algorithm 2.

Algorithm 2 : Auxiliary Particle Filter

Initialize particles $\{x_0^{(i)}\}_{i=1}^N$ and the corresponding weights $\{w_0^{(i)} = 1/N\}_{i=1}^N$

for $t=1, 2, \dots, T$ **do**

for $i=1, 2, \dots, N$ **do**

 a) Resample particle $\tilde{x}_{t-1}^{(i)}$ from $\{x_{t-1}^{(1)}, \dots, x_{t-1}^{(N)}\}$ using the weight $\tilde{w}_t^{(i)} \propto p(y_t | \lambda(x_{t-1}^{(i)}))$

 b) Propagate particle $\tilde{x}_{t-1}^{(i)}$ to $\tilde{x}_t^{(i)}$ using $f(\tilde{x}_t | \tilde{x}_{t-1}^{(i)})$

 c) Resample $x_t^{(i)}$ from $\{\tilde{x}_t^{(1)}, \dots, \tilde{x}_t^{(N)}\}$ using the weight $w_t^{(i)} \propto g(y_t | x_t^{(i)}) / g(y_t | \lambda(x_{t-1}^{(i)}))$

end for

end for

1.2.3 Residual Bernoulli Resampling

Bootstrap resampling has been described in the paragraph preceding Sect. 1.2.1. In fact, the name *bootstrap filter* in Sect. 1.2.1 came from bootstrap resampling that Gordon et al. in [23] used to convert weighted particles to particles with equal weights. *Residual Bernoulli resampling* has been proposed as an alternative to bootstrap resampling and has been shown to often lead to smaller variance for the associated particle filter than the bootstrap resampling scheme. The method is summarized in Algorithm 3.

Algorithm 3 : Residual Resampling Scheme

Input: A set of particles $\{(\tilde{w}_t^{(i)}, \mathcal{P}_t^{(i)}), i = 1, 2, \dots, M\}$

Output: A new set of particles $\{(\frac{1}{M}, \mathcal{P}_t^{(i)}), i = 1, 2, \dots, M\}$

Set $R = \sum_{i=1}^M \lfloor M \tilde{w}_t^{(i)} \rfloor$

for $i=1, 2, \dots, M$ **do**

 – Set $\hat{w}_t^{(i)} = \frac{M \tilde{w}_t^{(i)} - \lfloor M \tilde{w}_t^{(i)} \rfloor}{M - R}$

end for

for $j=1, 2, \dots, M$ **do**

 – Sample $\hat{N}_j \sim \text{mult}(M - R, \hat{w}_t^{(1)}, \hat{w}_t^{(2)}, \dots, \hat{w}_t^{(M)})$

 – Set $N_j = \lfloor M \tilde{w}_t^{(j)} \rfloor + \hat{N}_j$

 – Set $\mathcal{P}_t^{(j)} = \mathcal{P}_t^{(N_j)}$

end for

1.3 Particle Filters with Sequential Parameter Estimation

An important problem which has been studied extensively in recent filtering literature is that of joint parameter estimation and filtering for general state-space models. Traditional methods which incorporate the parameters as part of the latent state vector suffer from severe degeneracy problems due to the absence of state evolution dynamics for the subvector of latent states representing parameters. Methods to address this issue have been considered in [2, 9, 31, 35, 38], and [34]. They are summarized below.

1.3.1 Liu and West's Filter

Liu and West (2001) in [31] suggest to use a kernel smoothing approximation to the posterior density $p(\theta|\mathcal{Z}_{t-1})$ of the unknown parameter θ via a mixture of multivariate normals and to combine it with an auxiliary particle filter described in Algorithm 2. Let $\{x_{t-1}^{(i)}, \theta_{t-1}^{(i)}\}_{i=1}^N$ be a set of particles with weights $w_{t-1}^{(i)}$ ($i = 1, \dots, N$), which approximate $p(x_{t-1}, \theta|\mathcal{Z}_{t-1})$. They approximate the posterior density for θ by

$$p(\theta|\mathcal{Z}_{t-1}) = \sum_{j=1}^N w_{t-1}^{(j)} N(\theta; m^{(j)}, \Sigma_{t-1})$$

where

$$\begin{aligned} m^{(j)} &= a\theta_{t-1}^{(j)} + (1-a)\bar{\theta} \\ \bar{\theta} &= \sum_{j=1}^N \frac{\theta_{t-1}^{(j)}}{N} \\ \Sigma_{t-1} &= (1-a^2) \sum_{j=1}^N \frac{(\theta_{t-1}^{(j)} - \bar{\theta})(\theta_{t-1}^{(j)} - \bar{\theta})'}{N} \end{aligned}$$

The constant a measures the extent of shrinkage of the individual $\theta_{t-1}^{(j)}$ to the overall mean $\bar{\theta}$. It is a tuning parameter whose choice is discussed in [31]. The mixture approximation generates new samples from the current posterior and attempts to avoid particle degeneracy. The method is summarized in Algorithm 4.

Algorithm 4 : Liu and West's Filter

Output: The filtering particles $\{(x_t^{(i)}, \theta_t^{(i)})\}_{i=1}^N$ and the parameter posterior $p(\theta|\mathcal{Z}_t)$, $t = 1, \dots, T$
for $t=1, 2, \dots, T$ **do**

for $i=1, 2, \dots, N$ **do**

 a) Resample $(\tilde{x}_{t-1}, \tilde{\theta}_{t-1})^{(i)}$ from $\{(x_{t-1}^{(j)}, \theta_{t-1}^{(j)})\}_{j=1}^N$ with weights $w_t^{(i)} \propto p(y_t|\lambda(x_{t-1}^{(i)}), m^{(i)})$

 b) Propagate $\tilde{\theta}_{t-1}^{(i)}$ to $\hat{\theta}_t^{(i)}$ using $N(\cdot; m^{(i)}, \Sigma_{t-1})$

 c) Propagate $\tilde{x}_{t-1}^{(i)}$ to $\hat{x}_t^{(i)}$ using $p(x_t|\hat{x}_{t-1}^{(i)}, \tilde{\theta}_t^{(i)})$

 d) Resample $(x_t, \theta_t)^{(i)}$ from $\{(\hat{x}_t, \hat{\theta}_t)^{(j)}\}_{j=1}^N$ with weights $w_t^{(i)} \propto p(y_t|\hat{x}_t^{(i)}, \hat{\theta}_t^{(i)})/p(y_t|\lambda(\hat{x}_{t-1}^{(i)}), m^{(i)})$

end for

end for

1.3.2 Storvik's Filter

Storvik (2002) in [38] considered sequential parameter estimation for a class of state-space models in which the posterior parameter density $p(\theta|\mathcal{X}_t, \mathcal{Y}_t)$ can be written as $p(\theta|s_t)$, where s_t is a low-dimensional set of sufficient statistics that can be recursively updated by $s_t = \mathcal{S}(s_{t-1}, x_t, y_t)$. The method is summarized in Algorithm 5.

Algorithm 5 : Storvik's Filter

Output: The filtering particles $\{(x_t^{(i)}, \theta_t^{(i)})\}_{i=1}^N$ and the parameter posterior $p(\theta|s_t)$, $t = 1, \dots, T$

for $t=1, 2, \dots, T$ **do**

for $i=1, 2, \dots, N$ **do**

 a) Propagate $x_{t-1}^{(i)}$ to $\bar{x}_t^{(i)}$ using $q(x_t|\bar{x}_{t-1}^{(i)}, \theta, \mathcal{Y}_t)$

 b) Resample $(x_t, s_{t-1})^{(i)}$ from $\{(\bar{x}_t, s_{t-1})^{(j)}\}_{j=1}^N$ with weights $w_t^{(i)} \propto \frac{p(y_t|\bar{x}_{t-1}^{(i)}, \theta)p(\bar{x}_t^{(i)}|x_{t-1}^{(i)}, \theta)}{q(\bar{x}_t^{(i)}|x_{t-1}^{(i)}, \theta, \mathcal{Y}_t)}$

 c) Compute sufficient statistics $s_t^{(i)} = \mathcal{S}(s_{t-1}^{(i)}, x_t^{(i)}, y_t)$

 d) Sample $\theta_t^{(i)}$ from $p(\theta|s_t^{(i)})$

end for

end for

1.3.3 Particle Learning

Carvalho et al. (2010) in [9] propose the *particle learning* method which utilizes a resample-propagate scheme similar to auxiliary particle filters and show that the proposed method outperforms the filter of Liu and West of [31] in some comparative studies. Assuming the availability of conditional sufficient statistics s_t to represent the posterior of the parameter vector θ , and conditional sufficient statistics $s_{t,x}$ recursive state and parameter updates. The particles are now represented at each time by $z_t^{(i)}(x_t, s_t, s_{t,x}, \theta)^{(i)}$ and the Bayesian updating equation can be written as

$$p(z_t|\mathcal{Y}_t) = \int p(s_t|x_t, s_{t-1}, y_t)p(x_t|z_{t-1}, y_t)p(z_{t-1}|\mathcal{Y}_t)dx_t dz_{t-1} \quad (1.6)$$

where

$$p(z_{t-1}|\mathcal{Y}_t) \propto p(\mathcal{Y}_t|z_{t-1})p(z_{t-1}|\mathcal{Y}_{t-1}) \quad (1.7)$$

Denoting the updating formulas for s_t and $s_{t,x}$ by $s_t = \mathcal{S}(s_{t-1}, x_t, y_t)$ and $s_{t,x} = \mathcal{H}(s_{t-1,x}, \theta, y_t)$, Algorithm 6 summarizes the resampling and propagation steps in their filter.

Algorithm 6 : Particle Learning

Output: The filtering density approximated by particles $z_t^{(i)}$ and the parameter posterior $p(\theta|s_t)$
for $t=1, 2, \dots, T$ **do**
 1) Resample: $\hat{z}_{t-1}^{(i)}$ from the particles $\{z_{t-1}^{(j)}\}_{j=1}^N$ using weights $w_t^{(i)} \propto p(y_t|z_{t-1}^{(i)})$
 2) Propagate: $\hat{x}_{t-1}^{(i)}$ to $x_t^{(i)}$ using the distribution $p(x_t|\hat{z}_{t-1}^{(i)}, y_t)$
 3) Propagate: Parameter sufficient statistics $s_t^{(i)} = \mathcal{S}(\hat{s}_{t-1}^{(i)}, x_t^{(i)}, y_t)$
 4) Propagate: Sample $\theta^{(i)}$ from $p(\theta|s_t^{(i)})$
 5) Propagate: State sufficient statistics $s_{t,x}^{(i)} = \mathcal{K}(\hat{s}_{t-1,x}^{(i)}, \theta^{(i)}, y_t)$
end for

1.3.4 Particle MCMC

Hybrid methods that combine particle filters with MCMC schemes have been considered in the literature. Important recent developments in this direction are [2, 35], and [34]. Polson et al. (2008) in [35] use a rolling window MCMC algorithm that approximates the target posterior distribution by a mixture of lag- k smoothing distributions. They recast the filtering problem as a sequence of smaller smoothing problems which can be solved using standard MCMC approaches as in [7] and [8]. They exploit, whenever possible, a sufficient statistic structure as in [19] and [38] to perform parameter updates and develop an algorithm with linear computational cost. Andrieu et al. (2010) in [2] introduce the particle MCMC (PMCMC) methods to perform inference on the unknown parameter vector θ . Pitt et al. (2012) in [34] provide further analytic results on PMCMC and show that using auxiliary particle filters in PMCMC schemes may help reduce computation time.

1.4 A New Approach to Adaptive Particle Filtering

In this section, we describe a new adaptive filtering technique, recently introduced in [11, 12], for joint parameter and latent state filtering in particle filters, which provides substantial improvement over previous approaches. The authors in [12] propose an efficient MCMC method to estimate the posterior distribution of the parameters, which can be used in conjunction with traditional particle filter methods that assume the parameters to be known. Bukkapatanam et al. (2012) in [5] provide further development of the methodology and its applications to economics and finance, which will be summarized in Sect. 1.5.

Chan and Lai (2012a) in [11] begin by considering the case where the parameter vector θ is known so that it can be omitted from the notation for particle filters, as in Sects. 1.2.1–1.2.3. They consider more generally the estimation of $\psi_T = \mathbb{E}[\psi(\mathcal{X}_T)|\mathcal{B}_T]$ instead of $\mathbb{E}[\psi(x_T)|\mathcal{B}_T]$, where $\mathcal{X}_T = (x_1, \dots, x_T)$. When bootstrap resampling is performed at every stage, they show that the bootstrap filter estimate $\hat{\psi}_T$ of $\psi_T = \mathbb{E}[\psi(\mathcal{X}_T)|\mathcal{B}_T]$ has a martingale representation