Journal of Statistical Software
July 2008, Volume 26, Book Review 2. http://www.jstatsoft.org/

Reviewer: Tyler McMillen
California State University at Fullerton

Simulation and Inference for Stochastic Differential Equations: With R Examples

Stefano M. Iacus
Springer-Verlag, New York, 2008.
ISBN 978-0-387-75838. 286 pp. USD 69.50.

Introduction

The aim of this book is to develop a unified framework for simulating stochastic differential equations (SDEs) and estimating parameters of an underlying model from a given time series. This aim is mostly achieved. The book considers one-dimensional SDEs of the form $dX/dt = b\,(t, X) +$ "noise," or

$$dX = b(t, X)\,dt + \sigma\,(t, X)\;dW, \tag{1}$$

where $W(t)$ is the Wiener process. Here $b(t, X)$ is the drift and $\sigma^2(t, X)$ is the variance of the process. Such equations have been successfully used to model an enormous variety of problems in fields as diverse as biology, physics, neuroscience and finance. Most of the examples treated in this volume are drawn from finance. Although only one-dimensional processes are considered, as the author notes, most of the simulation techniques that can be used with one-dimensional processes apply equally well to their multi-dimensional counterparts, and references are provided for multi-dimensional SDE simulations.

The emphasis is on the practical implementation of code written in R. An R package **sde**, available for download from the Comprehensive R Archive Newtwork at `http://CRAN.R-project.org/package=sde`, accompanies the text. This is a powerful package that allows one to run simulations of SDEs and perform various inference routines for estimating parameters. The text of the book integrates theory and explanations of SDE with examples of R code that can be run with the **sde** package.

Although the outlines and main results of the theory of SDEs are presented in the text, it is not suitable as an introduction to SDEs. Most of the results are presented with minimal or no proof. Indeed, many of the explanations will be opaque to the reader with no prior familiarity with SDEs. The author suggests Øksendal (1998) as an introduction. Gardiner's 2004 fine book should also be mentioned, the latest edition of which contains a chapter on the simulation of SDEs. The collection of results in the first chapter of the book under review is quite useful, though, as these are employed throughout the text.

The strength of the book is its second half, on inference, i.e., the estimation of parameters in the drift $b(t,X)$ and variance $\sigma^2(t,X)$. Here, one takes a time series $X(t)$ that obeys the equation (1) for some form of the functions $b(t,X)$ and $\sigma(t,X)$. Then, one estimates the parameters in these functions. For instance, for an Ornstein-Uhlenbeck process, $b(t,X) = \theta_1 + \theta_2 X$ and $\sigma(t,X) = \theta_3$, and the values of $\theta_1, \theta_2, \theta_3$ can be estimated from the time series. With the **sde** package one can easily estimate parameters for a wide variety of underlying model assumptions. Additionally, there is a section on model selection, i.e., deciding which of several possible models gives the best fit to the time series.

Book contents

The book contains four chapters. The first two chapters discuss the theory and simulation of SDEs. The third and fourth are devoted to inference.

Chapter 1 begins with the basic results from the theory of SDEs. Several examples are presented along with simple R code for simulating them. As mentioned above, this section eschews proof, but there are an abundance of references for all of the results.

Chapter 2 covers numerical methods for simulating SDEs. The Euler and Milstein methods are described in some detail. Both of these schemes are built into the **sde** package. Numerous examples are presented, along with R code. The methods described here are not novel, but the R code and explanations of the **sde** package are quite useful.

Chapter 3, on inference, is the largest chapter and the real heart of the book. Throughout this chapter one assumes that a time series is described by an underlying model, e.g., Ornstein-Uhlenbeck, Cox-Ingersoll-Ross, or geometric Brownian motion. Several model examples are described, along with R code. Once one assumes one of these models, one must estimate the parameters in b and σ. Several techniques are explained, some of which are novel. These include pseudo-likelihood methods, approximated likelihood methods, the use of estimating functions, and the generalized method of moments. Bayesian methods are discussed briefly. The most useful aspect of the book is that it explains how to easily implement all of these methods in R, using just a few built-in commands.

Chapter 4 contains short discussions of three topics: (1) model identification, (2) nonparametric estimation, and (3) change-point estimation. I found the first of these the most interesting, since heretofore the author approached each problem by assuming that a given model underlay the process. Here we get to see how to obtain a model, from a number of candidate models, from the time series, and then use the methods of Chapter 3 to estimate the parameters. In nonparametric estimation one estimates b and σ "directly," without assuming any particular form. In the last small section the author discusses how to decide when a parameter in an underlying model has changed. He uses an interesting example from the stock exchange returns from 1971–1974 to illustrate the method.

The appendix contains an introduction to R, and a description of the **sde** package.

Conclusion

This is a useful book for the practitioner with a fairly solid understanding of SDEs and some familiarity with R or a similar language. With these prerequisites the reader will obtain a powerful tool for analyzing SDEs and time series. Moreover, these tools are all integrated into a single framework based on the R package **sde**.

There are a moderate number of typos and a few passages are unclear. For instance, on page 111 an expression for the likelihood function $L_T(\theta)$ is derived, and then it is stated that "θ can be estimated by maximizing $L_T(\theta)$." But it is not explained how to maximize $L_T(\theta)$. However, unclear passages like these are few, and for the most part the book is easy to follow. The material is well organized and it is easy to find the information one seeks. I couldn't find any mathematical errors.

Overall, this is a good book that fills in several gaps. In addition to collecting and summarizing an enormous quantity of theory, it introduces some novel techniques for inference. Statisticians and mathematicians who work with time series should find a place on their shelves for this book.

References

Gardiner CW (2004). *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*. 3rd edition. Springer-Verlag, New York.

Øksendal B (1998). *Stochastic Differential Equations: An Introduction with Applications*. 5th edition. Springer-Verlag, Berlin.

Reviewer:

Tyler McMillen
California State University at Fullerton
Department of Mathematics
Fullerton, CA 92834, United States of America
E-mail: tmcmillen@fullerton.edu

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 26, Book Review 2 *Published:* 2008-07-26
July 2008