Preface

Stochastic differential equations model stochastic evolution as time evolves. These models have a variety of applications in many disciplines and emerge naturally in the study of many phenomena. Examples of these applications are physics (see, e.g., [176] for a review), astronomy [202], mechanics [147], economics [26], mathematical finance [115], geology [69], genetic analysis (see, e.g., [110], [132], and [155]), ecology [111], cognitive psychology (see, e.g., [102], and [221]), neurology [109], biology [194], biomedical sciences [20], epidemiology [17], political analysis and social processes [55], and many other fields of science and engineering. Although stochastic differential equations are quite popular models in the above-mentioned disciplines, there is a lot of mathematics behind them that is usually not trivial and for which details are not known to practitioners or experts of other fields. In order to make this book useful to a wider audience, we decided to keep the mathematical level of the book sufficiently low and often rely on heuristic arguments to stress the underlying ideas of the concepts introduced rather than insist on technical details. Mathematically oriented readers may find this approach inconvenient, but detailed references are always given in the text.

As the title of the book mentions, the aim of the book is twofold. The first is to recall the theory and implement methods for the simulation of paths of stochastic processes $\{X_t, t \geq 0\}$ solutions to stochastic differential equations (SDEs). In this respect, the title of the book is too ambitious in the sense that only SDEs with Gaussian noise are considered (i.e., processes for which the writing $\mathrm{d}X_t = S(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$ has a meaning in the Itô sense). This part of the book contains a review of well-established results and their implementations in the R language, but also some fairly recent results on simulation.

The second part of the book is dedicated to the review of some methods of estimation for these classes of stochastic processes. While there is a well-established theory on estimation for continuous-time observations from these processes [149], the literature about discrete-time observations is dispersed (though vaste) in several journals. Of course, real data (e.g., from finance [47],

[88]) always lead to dealing with discrete-time observations $\{X_{t_i}, i = 1, \dots, n\}$, and many of the results from the continuous-time case do not hold or cannot be applied (for example, the likelihood of the observations is almost always unavailable in explicit form). It should be noted that only the observations are discrete whilst the underlying model is continuous; hence most of the standard theory on discrete-time Markov processes does not hold as well.

Different schemes of observations can be considered depending on the nature of the data, and the estimation part of the problem is not necessarily the same for the different schemes. One case, which is considered "natural," is the fixed-Δ scheme, in which the time step between two subsequent observations X_{t_i} and $X_{t_i+\Delta_n}$ is fixed; i.e., $\Delta_n = \Delta$ (or is bounded away from zero) and independent from n. In this case, the process is observed on the time interval $[0, T = n\Delta]$ and the asymptotics considered as $n \to \infty$ (*large-sample asymptotics*). The underlying model might be ergodic or stationary and possibly homogeneous. For such a scheme, the time step Δ might have some influence on estimators because, for example, the transition density of the process is usually not known in explicit form and has to be approximated via simulations. This is the most difficult case to handle.

Another scheme is the "high frequency" scheme, in which the observational step size Δ_n decreases with n and two cases are possible: the time interval is fixed, say $[0, T = n\Delta_n]$, or $n\Delta_n$ increases as well. In the first case, neither homogeneity nor erogidicy are needed, but consistent estimators are not always available. On the contrary, in the "*rapidly increasing experimental design*," when $\Delta_n \to 0$ and $n\Delta_n \to \infty$ but $n\Delta_n^2 \to 0$, consistent estimators can be obtained along with some distributional results.

Other interesting schemes of partially observed processes, missing at random [75], thresholded processes (see, e.g., [116], [118]), observations with errors (quantized or interval data, see, e.g., [66], [67], [97]), or large sample and "small diffusion" asymptotics have also recently appeared in the literature (see, e.g., [222], [217]). This book covers essentially the parametric estimation under the large-sample asymptotics scheme ($n\Delta_n \to \infty$) with either fixed $\Delta_n = \Delta$ or $\Delta_n \to 0$ with $n\Delta_n^k \to 0$ for some $k \geq 2$. The final chapter contains a miscellaneous selection of results, including nonparametric estimation, model selection, and change-point problems.

This book is intended for practitioners and is not a theoretical book, so this second part just recalls briefly the main results and the ideas behind the methods and implements several of them in the R language. A selection of the results has necessarily been made. This part of the book also shows the difference between the theory of estimation for discrete-time observations and the actual performance of such estimators once implemented. Further, the effect of approximation schemes on estimators is investigated throughout the text. Theoretical results are recalled as "Facts" and regularity conditions as "Assumptions" and numbered by chapter in the text.

So what is this book about?

This book is about ready to be used, R-efficient code for simulation schemes of stochastic differential equations and some related estimation methods based on discrete sampled observations from such models. We hope that the code presented here and the updated survey on the subject might be of help for practitioners, postgraduate and PhD students, and researchers in the field who might want to implement new methods and ideas using R as a statistical environment.

What this book is not about

This book is not intended to be a theoretical book or an exhaustive collection of all the statistical methods available for discretely observed diffusion processes. This book might be thought of as a companion book to some advanced theoretical publication (already available or forthcoming) on the subject. Although this book is not even a textbook, some previous drafts of it have been used with success in mathematical finance classes for the numerical simulation and empirical analysis of financial time series.

What comes with the book

All the algorithms presented in the book are written in pure R code but, because of the speed needed in real-world applications, we have rewritten some of the R code in the C language and assembled everything in a package called `sde` freely available on CRAN, the Comprehensive R Archive Network. R and C functions have the same end-user interface; hence all the code of the examples in the book will run smoothly regardless of the underlying coding language. A minimal knowledge of the R environment at the introductory level is assumed, although brief recalls to the main R concepts, limited to what is relevant to this text, are given at the end of the book. Some crucial aspects of implementation are discussed in the main body of the book to make them more effective.

What is missing?

This book essentially covers one-dimensional diffusion processes driven by the Wiener process. Today's literature is vast and wider than this choice. In particular, it focuses also on multidimensional diffusion processes and stochastic differential equations driven by Lévy processes. To keep the book self-contained and at an introductory level and to preserve some homogeneity within the text, we decided to restrict the field. This also allows simple and easy-to-understand R code to be written for each of the techniques presented.

Acknowledgments

I am grateful to Alessandro De Gregorio and Ilia Negri for careful reading of the manuscript, as well as to the anonymous referees for their constructive comments and remarks on several drafts of the book. All the remaining errors contained are of course my responsibility. I also wish to thank John Kimmel for his great support and patience with me during the last three years.

Milan, November 2007 *Stefano M. Iacus*

http://www.springer.com/978-0-387-75838-1

Simulation and Inference for Stochastic Differential Equations
With R Examples
Iacus, S.M.
2008, XVIII, 285 p., Hardcover
ISBN: 978-0-387-75838-1