Journal of Statistical Software

October 2015, Volume 67, Book Review 7. doi: 10.18637/jss.v067.b07

Reviewer: James P. Howard, II
University of Maryland University College

Uncertainty Quantification and Stochastic Modeling with MATLAB

Eduardo Souza de Cursi and Rubens Sampaio
ISTE Press, London, 2015.
ISBN 978-1785480048. 442 pp. USD 180 (H).
http://www.sciencedirect.com/science/book/9781785480058

Souza de Cursi and Sampaio's new book, *Uncertainty Quantification and Stochastic Modeling with MATLAB*, presents a collection of methods for measuring, quantitatively, the potential error in simulations and models. Uncertainty quantification is a relatively new field within numerical analysis with applications across engineering, social sciences, and natural sciences as computer-based simulations and complex statistical analyses become core analytical methods. The creators of deployed models, such as high-frequency traders and recommendation system developers, will also find this book applicable.

Uncertainty comes from many potential sources, including failure of the model to capture the full detail of the application, errors in measuring input parameters, and potential numerical errors in the calculations. These uncertainty problems may involve finding the distribution of the optimal point in an optimization problem or the spread of potential solutions to the equation, $A\mathbf{x} = \mathbf{b}$. Many of these questions lie at the heart of numerical analysis and challenge implementers across fields. Understanding the sources of uncertainty, along with their behavior, can allow researchers to estimate the effects and compensate accordingly.

Starting with these uncertainty sources, the authors open with an extensive chapter covering random processes and their implementation in MATLAB. Serving mostly as review material, the chapter introduces randomness, describes several critical random number distributions, leading into a theoretical overview of stochastic simulations. They continue with a review of Monte Carlo simulation processes and a description of entropy and information, followed by a chapter on the representation of random variables. The book then moves through uncertainty quantification in core areas of numerical analysis: linear and nonlinear equations, differential equations, and optimization. These uncertainty problems may involve finding the distribution of the optimal point in an optimization problem or the variance of potential solutions to a linear system.

The book is well organized and mirroring the traditional numerical analysis course. This allows the reader to expand on previous ideas as they read, which may or may not be the case in many books on numerical programming. In addition, the review of probability, statistics,

and Monte Carlo methods hits the right highlights necessary to understand the rest of the text without getting bogged down in the more arcane statistical minutiae. Within this framework, the text properly develops from the source material and clearly presents the quantification of uncertainty in a rigorous manner.

A great drawback to this book is less the writing and more the production. In many instances, algorithm implementations or program listings exceed one page. Frequently, regardless of length, listings cross page boundaries, which makes it more difficult to read and understand. An egregious example of this is listing 5.2, which comprises seven functions across five pages. Several related functions are included together, which is good implementation and design.

Software developers consider it a best practice to ensure that methods fit within a single screen of text, because it reduces total complexity and helps developers understand what is in front of them. We should encourage the same practice in books on numerical analysis for the same reasons. In addition, many practitioners of numerical analysis and computational statistics are neither programmers nor computer scientists, but may be social scientists, engineers, or otherwise have less training in numerical methods. As a result, steps to ease the understanding and readability of example programs should be taken.

That said, the code is well developed and documented. Despite the formatting problems, the code itself is clean, readable, and well-formatted with keywords bolded and indentation intact. Further, each function is documented, with respect to its expected inputs and outputs, in standard MATLAB help format, in the comment section of each function. If the code examples were installed, the standard help functions would return results for the examples. These touches lend credibility to the idea that these code examples are functioning, came from real world applications, and would work if rekeyed from the book into a text editor.

And rekeying them will be necessary. Souza de Cursi and Sampaio correctly state in their introduction that the code included with their book is not ready for actual use and should be used as a guideline for later reimplementation by the reader. This is good because creating complete, error-checking, and comprehensive examples is beyond the feasible for a book. Page limitations and the highly repetitive nature of error checking do not lend themselves to good printing. However, it would help if the code were available to use as a starting point. Internet searches and the publisher's page did not reveal any sort of "download the source" option that would include the code from this book. This is despite the availability of a color plate download made available.

Like many mathematics books, the text is heavy on theory, proof, and equations. This is common in mathematics and statistics and is something we should support as readers, reviewers, and authors. However, in a number of places, there is limited exposition, which can make it difficult to pick up the text at a specific point and find value. Part of this is the nature of mathematics, though more exposition can mitigate the effects, if not eliminate them. The book would be better served with more plain language explanation of the models developed.

Despite these flaws, the material in the text is strong. Cleaning up the formatting and adding more detailed explanations will help this book dramatically. As we can expect a second edition, as with so many other books, these changes will improve the book and create a better reference or even textbook going forward. And as stochastic modeling use will grow, understanding estimation error will grow more and more important. Understanding the error will allow economists, engineers, medical researchers, and others to better understand their numerical results.

Reviewer:

James P. Howard, II
Mathematics and Statistics
University of Maryland University College
E-mail: jh@jameshoward.us or jphoward@faculty.umuc.edu
Twitter: @howardjp
URL: http://jameshoward.us/

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
October 2015, Volume 67, Book Review 7 *Published:* 2015-10-06
doi:10.18637/jss.v067.b07