Journal of Statistical Software

April 2016, Volume 70, Book Review 4. doi: 10.18637/jss.v070.b04

Reviewer: Norman Matloff
University of California, Davis

Numerical Python: A Practical Techniques Approach for Industry

Robert Johansson
Apress, New York, 2015.
ISBN 1-4842-0554-9. 487 pp. USD 59.99 (P).
http://www.apress.com/9781484205549

This is a very useful book, especially for *Journal of Statistical Software (JSS)* readers who wish to learn more about numerical methods and/or scientific applications of Python. Though arguably the book focuses more on numerical analysis than on statistics, it has three solid chapters on the latter, plus abundant infrastructure material useful for statisticians. The author, a physicist by background, has more recently been working in the data science area.

Given the rather lengthy size of the book, the author does not attempt to include a basic Python tutorial. The reader is assumed to have some prior experience with the language. However, such background may be quickly acquired from the numerous tutorials on the web (such as my *Fast Lane to Python*, http://heather.cs.ucdavis.edu/pyfast.pdf). The author does recommend the use of development environments, notably **IPython** (now transitioning to **Jupyter**) and **Spyder**, which are introduced briefly in Chapter 1.

Those with just a passing familiarity with Python, though, will find that they must first learn some infrastructure material on *numeric* Python, specifically the **NumPy** and **SciPy** packages, which are presented clearly in the early chapters. R programmers will find Chapter 2 well within their comfort zone, with lots of similarities.

Some readers may be a bit startled to see that Chapter 3 then covers an unusual topic, symbolic computing, but this reviewer was delighted to see it. Here the **SymPy** package is presented in a fair amount of detail, and even those with no background in symbolic math will find themselves able to manipulate algebraic expressions, including matrix-valued ones, with rather little effort.

Chapter 4, *Plotting and Visualization*, uses **Matplotlib**. My own preference would be for R's extensive – and more statistics-oriented – graphical facilities, but the **Matplotlib** APIs are clean and easy to apply.

Chapters 5–11 are devoted to classical numerical methods, such as the solution of ODEs, but many *JSS* readers will find Chapter 10, *Sparse Matrices and Graphs*, to be of interest. There is a nice section at the end on analysis and display of random graphs.

Chapters 14–16 are titled *Statistical Modeling*, *Machine Learning* and *Bayesian Statistics*, respectively. Recently there has been a push in the Python world to develop more statistical capabilities in the language, and here I must disclose my own bias, which is that generally the quality and depth of this software does not rise to the level of R, which I always like to describe as "Written *by* statisticians *for* statisticians." For instance, in an example presented on linear regression modeling, we find phrasing such as "...data whose true value is $y = 1+2x_1+3x_2+4x_1x_2$," an oversimplified if not outright misleading characterization. However, the presentation is clear and engaging, with nice examples and graphics.

Chapter 17, *Signal Processing*, has material on Fourier analysis, again sure to be of interest to some *JSS* readers.

Chapter 18, *Data Input and Output*, is overly modest in its title, as it goes beyond the usual topics such as reading CSV files. Its section on HDF5 is particularly interesting.

Chapter 19, *Code Optimization*, brings up a question quite familiar to us R folks: How does one achieve better performance from an interpreted language?

In summary, this book is chock full of interesting and useful material, even for those who are not primarily Python users. It would make a fine addition to any data analyst's bookshelf, and I am glad it is on mine.

Reviewer:

Norman Matloff
University of California, Davis
Department of Computer Science
Davis, CA 95616, United States of America
E-mail: matloff@cs.ucdavis.edu

Journal of Statistical Software http://www.jstatsoft.org/
published by the Foundation for Open Access Statistics http://www.foastat.org/
April 2016, Volume 70, Book Review 4 *Published:* 2016-04-15
doi:10.18637/jss.v070.b04