## Numerical Methods of Statistics

John R. Monahan. New York:Cambridge University Press, 2001.
*xiii* + 428pp. $xx.

The suitability of a book as a textbook for any course depends on the depth and breadth of material covered. In the preface, the author states that the book grew out of notes for a graduate level course in statistical computing that he has taught for 20 years. This sounds familiar in that Kennedy and Gentle (1980) make a similar comment in the preface to their book. It seems fitting that this book appears roughly 20 years later. It is for the user to decide if this is a suitable replacement for the earlier book that now appears to have become a classic. Lange (1999) in the preface to his book comments on the difficulty of finding a suitable text for teaching a graduate level course on statistical computing. Thus any effort at fulfilling this need is to be commended.

One reason for this dearth is that the area of statistical computing has been constantly changing during the last two decades. Because statistical methodology developed during this period involves the heavy use of computers and computational algorithms, more and more statisticians are expected to gain expertise in a variety of numerical techniques. Modern applications range from Markov Chain Monte Carlo in Bayesian applications to computations associated with nonparametric function estimation, such as those using wavelets. Perhaps this is an expected natural progression of knowledge in an area of study in statistics. However, the effect of this has been that the interface between statistics and computing has now become blurred to such an extent that almost anyone being trained as a statistician has to become an expert in one or more aspects of computing. There is not a more clear example of this than in the area of biostatistics. One might claim that some biostatisticians today have more expertise in the use of certain computational algorithms than mainstream statisticians.

So it is inevitable that many statistics departments attempt to accommodate these changes by strengthening statistical computing instruction in their curricula. This requires that adequate material (preferably in the form of textbooks) are available for instructors. Lange(1999) in prefacing his book documents that the classical book by Kennedy and Gentle (1980) has become outdated and needs updating. The appearance of Thisted (1988) may have fulfilled this need only partially, mainly because it does not examine some computational methods that are now used extensively (e.g., Markov Chain Monte Carlo). Monahan's goal, however, is not too different: as he explains, it is "to prepare doctoral students with the computing tools needed for statistical research." Thus the intended target of this book is clear.

Chapters 1 and 2 consist of the obligatory introductions to algorithms, programming languages, and computer arithmetic. The author, by recounting a personal experience where an algorithm for the iterative computation of the Hodges-Lehmann estimator was stuck in a loop, emphasizes the need for care in constructing iterative procedures. The discussion of floating point computations is an essential part of any text on numerical computing, and the understanding

of this material is extremely important in this day of point-and-click computing. Starting with a nice introduction to number systems, the author retraces the fundamental material in a matter of a few pages. Although the discussion is somewhat similar to that in Thisted (1988), I found the presentation here more useful for someone new to numerical computation using a language like Fortran. The author ends the discussion on computer arithmetic by offering some prescriptions for limitations on accuracy imposed by floating point representation of numbers.

Chapters 3 to 6 address computations associated with linear algebra. The standard algorithms for solving either systems of linear equations or eigen problems are presented using a rather traditional approach. In computational statistics, for understandable reasons linear algebra computations have made the most advances over the years. Many statisticians may question the need for detailed knowledge of the implementation of such algorithms (e.g., QR decomposition), particularly since these are rather well implemented in software systems such as S-Plus. Also, FORTRAN implementations of LINPACK/LAPACK routines are available through popular libraries such as IMSL or NAG. The need, therefore, one might argue, is to emphasize their applications in statistics rather than implementation, particularly for those training to be statisticians.

The author attempts to balance the two perspectives, illustrating with examples of how various algorithms work as well as their applications, and including discussions on topics such as accuracy and conditioning. I found the discussion here to be a bit broader than in Thisted (1988) who covers three topics in a single large chapter. I found it a little strange that conjugate gradients methods were covered in Chapter 5 as an application to the least squares problem rather than a general optimization algorithm with other applications in statistics in Chapter 8. As in Thisted (1988), iterative methods (e.g., Gauss-Seidel) are fittingly discussed in Chapter 4. The treatment here, however, leaves something to be desired. Thisted, for example, discusses several possible applications of these methods in statistics. Such a discussion could be more useful to encourage possible future applications of these algorithms in statistics. I believe the author takes such an approach in presenting the complex singular value decomposition in Chapter 6.

As in the earlier chapters, Monahan provides quite a number of useful Fortran programs to illustrate the algorithms discussed (e.g., sweep). As the author himself states, many of the presentations follow those in the texts by Stewart (1973) and Golub and Van Loan (1984, 1989, 1990). This might be a blessing for those who have never been formally introduced to expert presentations in these superb texts on algorithms associated with matrix algebra.

Chapter 7 presents an excellent exposition of function interpolation, smoothing and approximation. Monahan covers this material separately from numerical integration, whereas these are presented in earlier texts in sections introducing numerical quadrature (e.g., Kennedy and Gentle (1980 ), Thisted (1988) ). Perhaps this is warranted since function approximation has now become a more specialized research area in statistics. This chapter ends with a section on computing probability functions, perhaps duplicated in other books. The difference

here is that Fortran routines are again provided for testing and comparing the algorithms presented.

Chapter 8 presents the all-important topic of optimization with maximum likelihood and associated topics deferred until Chapter 9. However, the presentation in this chapter is too short and the approach too cursory in comparison with those in Thisted (1988). Even the older Kennedy and Gentle (1980) is far more informative to those uninitiated on this subject. The introduction, where the mean value theorems are presented, would have been an ideal place to introduce the fixed point algorithm, and later would have allowed the author to discuss a method like Aitken acceleration as a supplement to the discussion on convergence rates. The materials though are clearly and succinctly presented and the discussions on stopping rules and numerical differentiation are useful to the practitioner. These are again supplemented by very useful test/demonstration code. Important material omitted from this section includes extensions to Newton-Raphson and modified secant methods. These are treated as special cases of the nonlinear regression problem in Chapter 9.

In Chapter 9, following an introduction that includes some theoretical results on maximum likelihood, an extended example is presented. This example allows the author to illustrate the variations of Newton-Raphson, including the method of scoring and the use of numerical derivatives. However, including a few more statistical examples, as in Thisted (1988), would have helped elaborate on various aspects of this important numerical method in statistics. These include choice of initial values, convergence tests, convergence rates of the different methods, and related these to the prescriptions offered in Chapter 8. A discussion on concentrated likelihood leads to a presentation on the EM algorithm that in my opinion is highly inadequate. At least an entire section would have been needed to do justice to this important topic that finds many applications in statistical research. Since the appearance of Thisted (1988), many advances have been made in EM and a few pages of new material could have updated the content in that text.

The section on nonlinear regression begins with a conventional introduction to the Gauss-Newton method. Modifications to Newton-Raphson (Levenberg-Marquardt) and quasi-Newton methods (e.g., Broyden-Fletcher-Goldfarb-Shanno) are discussed here in the context of these iterations. From the statistician's point of view this appears to be sufficient, but it may lead to some misunderstanding of the techniques involved, because these methods are more general. In the discussion on software that incorporates these algorithms (including NL2SOL of Dennis, Gay and Welsch (1981) ), Monahan is excellent. The author even provides a program *nllsq*, for implementing one algorithm along with several test problems from the literature. Overall, the code provided for this Chapter is quite extensive and useful for both the novice and the expert alike.

Chapters 10 through 12 cover numerical integration and Monte Carlo methods, while Markov Chain Monte Carlo (MCMC) is covered entirely in Chapter 13. The author emphasizes the important point that, "Monte Carlo should be viewed as just another way to compute an integral; numerical integration should be viewed as just another way to sample points in space". Monahan begins with

a simple simulation example in statistics to compare two estimators and extends it to motivate a discussion on methods available for examining the posterior in the Bayesian context. This leads to the introduction of both one-dimensional and high dimensional quadrature, which is applied to the computation of the posterior mean and variance of an example introduced earlier in Chapters 8 and 9. Here the author alludes to Chapter 7 where function approximation was first discussed and uses it in constructing Gauss quadrature formulas. This kind of development of material in terms of practical problems is sprinkled throughout and I find this aspect to be the most appealing feature of the book. The discussion on quadrature is followed by one on uniform random number generators. This section is quite well-written and provides an excellent summary of the current state of the literature, supplemented by code for 15 generators. My only complaint is that the discussion on variance reduction methods could have been developed more fully. These methods find many applications in real problems, thus deserving a broader treatment, although importance sampling gets an extended discussion in Chapter 12 in another context. An interesting example is used to illustrate the computation of the posterior mean vector and covariance matrix in a 3-dimensional variance component problem, using a Korobov scheme to perform quasi-Monte Carlo integration. Code for this problem is provided.

Chapter 11 presents a discussion of basic methods for generating random samples from distributions and I did not find this treatment too different from other sources, say, for e.g., Gentle (1999). Material discussed in Chapter 12 purports to be statistical tools for examining output from Monte Carlo experiments; however, it includes several sections on specialized numerical integration methods. One feels that material in Chapters 10 and 12 could have been better organized; maybe one of those chapters could have exclusively reserved for material on Bayesian computation. The coverage of MCMC as a separate chapter (Chapter 13) is a step in the right direction. However, instructors may find that the author's organization of the material leads to some confusion. The introductory discussion on Markov Chains should ideally lead directly to Metropolis-Hastings so that the idea that the algorithm is really the construction of a Markov Chain satisfying the reversibility condition is solidified. Also, the author emphasizes the application of MCMC methods to Bayesian problems; the merit of this approach is arguable. It would be helpful to present MCMC as a general algorithm for generating random samples from some target distribution. This chapter includes excellent discussions on Gibbs Sampling, adaptive acceptance/rejection methods, and diagnostics for samples generated from MCMC methods. Again this chapter is accompanied by excellent code for demonstrating Metropolis-Hastings and an adaptive acceptance/rejection algorithm.

Chapter 14 is somewhat of an outlier in that it contains a collection of numerical methods that are useful but seem to have no relation to the rest of the material in the book. For example, discussion on sorting algorithms and Fourier transform are easily accessible from many specialized texts.

All of the chapters conclude with a nice set of exercises that vary in difficulty

level, followed by a set of references prefaced by the author's comments on the texts and software referenced. Some of the exercises not only require the use of the routines provided by the author, but also may require additional programming. In closing, I found this to be an extremely readable book. The writing style is informal and concise and the use of mathematics is kept to a minimum, with some of the proofs left as exercises. A determined attempt is made to suggest a statistical application for every method or algorithm discussed. This would be an excellent book for a graduate level course in statistical computing that requires knowledge of Fortran programming as a prerequisite. As expected in a book based on a course taught by the author, the selection of topics and the coverage is not complete. Thus, such a course may need to be supplemented by other complementary texts such as Thisted (1988) and Lange (1999). The author must be congratulated for putting together (and making available in a easily accessible form and free of charge) the excellent software component of the text.

Mervyn G. Marasinghe

*Iowa State University*

## REFERENCES

Dennis, Jr. J. E., Gay, D. M., and Welsch, R. E.(1981). " An Adaptive Nonlinear Least-squares Algorithm," *ACM Transactions on Mathematical Software*, 7:348-383

Gentle, J. E. (1998), *Random Number Generation and Monte Carlo Methods*, New York:Springer.

Golub, G. H. and Van Loan, C. (1984, 1989, 1990), *Matrix Computations*, Baltimore:Johns Hopkins University Press.

Kennedy, W. J. and Gentle, J. E. (1980), *Statistical Computing*, New York: Marcel Dekker.

Lange, K. (1999), *Numerical Analysis for Statisticians*, New York:Springer.

Stewart, G. W. (1973), *Inroduction to Matrix Computations*, New York:Academic Press.

Thisted, R. A. (1988), *Elements of Statistical Computing:Numerical Computation*, New York:Chapman and Hall.