

2. DATA STRUCTURES FOR COMPUTATIONAL STATISTICS. Sigbert Klinke, Physica-Verlag, Springer, Heidelberg, 1997. No. of pages: viii + 284. Price: \$60.40. ISBN 3-7908-0982-9

This book is essentially the author's doctoral thesis wherein computational structures (data and programs) are described as implemented by the author. The book is perhaps mistitled for it suggests that one might have found a substantive review and exploration of the novel computational structures which have appeared in the statistical computing literature. This is not the case. Instead, the book's principal computational focus is on the statistical system XploRE [1] and the data structures which are described are routine.

A fair bit of ground related to exploratory data analysis is hastily surveyed in the early chapters. A chapter on standard statistical graphics for exploratory data analysis is provided but the hasty treatment provides few pointers to the original literature and occasionally makes naive interpretative errors. This chapter is followed by one which discusses a few selected methods for computationally uncovering structure in data, namely some cluster analysis methods and some modern computationally intensive regression methods in fairly short order (for example, alternating conditional expectations, isotonic regression, projection pursuit regression etc.). These chapters are meant to provide motivation for the computational structures to follow but the cursory treatment means they will not be of much value as introduction to these topics.

Fully one-quarter of the book is devoted to a single chapter entitled 'Exploratory projection pursuit'. Here the projection pursuit methodology is illustrated in conjunction with interactive graphics to explore the structure of a data set. Some well known data sets are examined. Much discussion is given to the selection and calculation of appropriate indices of 'interestingness' and it is here where the greatest attention is given to detail. Here also is where one finds the author's novel work on the development of multivariate indices.

If the early discussion was to motivate interesting data structures, then the next chapter (called 'Data structures') does not live up to expectations. The putative decomposition of graphical objects is remarkably disappointing – consisting as it does of the general three step recommendation:

1. Do the mathematical computation;
2. Create graphical objects;
3. Show the graphical objects in a plot to the user.

For the majority of the examples then provided at least one of the steps in this 'decomposition' is 'do nothing'; the 3-d scatter plot for example has this as its first two steps! The chapter concludes with a brief summary of data structures and procedures for a variety of well-known statistical systems (for example, vectors:  $y$  or  $n$ , matrices:  $y$  or  $n$  etc.). It is surprisingly bereft of any information on the various rich data structures which have been proposed and implemented for statistical graphics in a variety of North American research dating back to the 1980s. The last substantive chapter of the book, entitled 'Implementation in XploRE', essentially lists the XploRE commands which implement the statistical and graphical methods described earlier in the book. This would be the principal value of the book to users of XploRE.

Computational procedures are described at levels of detail which vary throughout the book – from program data structures, to pseudo code algorithms, to general discussions. Attention is mostly directed toward describing algorithms rather than details of their numerical properties (for example, one algorithm for weighted least squares suggests forming the  $(X^T W X)^{-1}$  directly and multiplying by  $X^T W y$  rather than describing a QR decomposition and back-solving; forming the cross-product matrix is notorious for introducing ill conditioning and can only add unnecessary operation counts). The treatment of object-oriented programming also seems naive given the importance the book attributes to the methodology.

Relatively minor blemishes of the book include the absence of an index, occasionally awkward English, and typographical errors.

In summary, the book is not what one might expect from its title. In depth of coverage and understanding, it is uneven across the topics it covers. On the topics of statistical graphics and data structures, it is less than authoritative. I expect the book will be of some interest to users of XploRE and possibly to those interested in indices for exploratory projection pursuit.

R. W. OLDFORD

*Department of Statistics and Actuarial Science  
University of Waterloo  
Waterloo, Ontario N2L 3G1, Canada*

#### REFERENCE

1. Härdle W, Klinke S, Turluch B.A. *XploRE: An Interactive Statistical Computing Environment*. (Statistics and Computing series). Springer-Verlag: New York, 1995.