O. Berke
Department of Population Medicine
University of Guelph
Guelph, Canada
and
Department of Biometry, Epidemiology,
and Information Processing
School of Veterinary Medicine
Hannover, Germany

RUPPERT, D., WAND, M. P., and CARROLL, R. J. **Semiparametric Regression.** Cambridge University Press, Cambridge, 2003. xvi + 386 pp. US$45.00/£29.95 (paperback), US$100.00/£70.00 (hardcover). ISBN 0-521-78516-2.

The preface of this book starts with "The primary aim of this book is to guide researchers needing to incorporate nonlinear relationships into their regression analyses. (. . .) Almost all existing regression texts treat either parametric or nonparametric regression exclusively. The level of exposition between books of either type differs alarmingly."

Based on my own experience in collaborative research I completely concur with this statement: It often took a lot of time and work to convince colleagues from medicine, biology, economics, or social science that modern non- and semiparametric regression has a lot to offer to analyze substantive research problems. While scientists in these fields are often quite familiar with linear models, they hesitate to apply state-of-the-art flexible regression. Moreover, sometimes it is easier to convince them to use "black box" data mining tools. I think this situation is really alarming for statistics as a discipline.

This great book is the first one to remove barriers and to close gaps between advanced statistical methodology and applied research in various fields. Many applications from the authors' collaborative research demonstrate the usefulness of flexible semiparametric regression and motivate readers from other disciplines. The level of exposition keeps methodological technicalities to a minimum, so that the book is suitable even for students or working scientists with only a moderate background in regression. It is perhaps most useful for statistically oriented scientists who are familiar with linear regression and are confronted with complex problems where more flexible semiparametric models are needed.

There is even enough new material to attract experts interested in methodological development. As a particular feature, the approach to semiparametric regression in this book is based on penalized regression splines and (generalized) mixed model technology (including pioneering work by the authors). An obvious advantage of this approach is that semiparametric regression can be treated as a rather natural extension of linear regression. For other models and methods of smoothing and nonparametric regression readers are referred to the relevant literature in bibliographic notes at the end of chapters.

After an introductory chapter with motivating examples, Chapter 2 contains the essentials of linear parametric regression and some extensions. Starting from polynomial regression, scatter plot smoothers based on linear (penalized) splines, represented in the easily understandable truncated power series basis, are introduced in Chapter 3. Extensions and other bases, such as B-splines or radial basis functions, are covered in short sections.

Chapter 4 introduces linear mixed models and corresponding inference. While these models are of interest on their own, they are also needed as a prerequisite for the mixed model representation of penalized regression splines, which first appears in Section 4.9. Using this basic concept, mixed model methodology can be applied for automatic scatterplot smoothing in Chapters 5 and 6. Chapters 7, 8, and 9 extend the mixed model approach for penalized splines to additive and additive mixed models. Generalized linear and mixed models are introduced in Chapter 10, and Chapter 11 extends the mixed model approach for penalized splines to generalized additive models.

Models with various forms of interactions are reviewed in the expository Chapter 12. More details, including quite recent methodological research on bivariate smoothers for continuous-by-continuous interactions and related spatial smoothers, are given in Chapter 13. Several concepts such as tensor product splines, radial basis functions, and kriging are treated from a unifying point of view. This chapter is of particular value for researchers in the fields of spatial statistics and surface smoothing. Variance function estimation and measurement error models are briefly addressed in Chapters 14 and 15. Chapter 16 gives some insight into the rapidly developing fully Bayesian approaches to semiparametric regression using MCMC, while Chapter 17 focuses on a spatially adaptive version of penalized splines developed by the authors. Chapter 18 describes several applications, and the book concludes with appendices on matrix algebra, probability and statistical inference, and computational issues.

I highly recommend this book, which definitely reaches the aims stated at the beginning. It provides a very readable access to modern semiparametric regression, demonstrates its potential in various applications, and is an inspiring source for new ideas. I enjoyed reading this book.

L. Fahrmeir
Ludwig Maximilian University
München, Germany

ZIMMERMANN, K.-H. **An Introduction to Protein Informatics.** Kluwer Academic Publishers, Boston, 2003. xviii + 280 pp. US$135.00/£84.00/EUR122.00. ISBN 1-40207578-2.

This book will be useful to readers, whether professional researchers or graduate students, interested in the mathematical apparatus required for the computer simulation of protein structures.

The de novo protein design constitutes the main goal of protein research. That goal may be attained with appropriate computer programs, tested in the modeling of known protein structures.

The computer simulation of the tertiary structure for a given amino acid sequence may be carried out on the basis of sequence homology, or by build up from either fragments of known secondary structure or directly from the amino acid residues. The build up may then be followed by an optimization procedure (usually restricted to variation of