Preface

This is the second volume of a text on the theory and practice of maximum penalized likelihood estimation. It is intended for graduate students in statistics, operations research, and applied mathematics, as well as researchers and practitioners in the field. The present volume was supposed to have a short chapter on nonparametric regression but was intended to deal mainly with inverse problems. However, the chapter on nonparametric regression kept growing to the point where it is now the only topic covered. Perhaps there will be a Volume III. It might even deal with inverse problems. But for now we are happy to have finished Volume II.

The emphasis in this volume is on smoothing splines of arbitrary order, but other estimators (kernels, local and global polynomials) pass review as well. We study smoothing splines and local polynomials in the context of reproducing kernel Hilbert spaces. The connection between smoothing splines and reproducing kernels is of course well-known. The new twist is that letting the inner product depend on the smoothing parameter opens up new possibilities: It leads to asymptotically equivalent reproducing kernel estimators (without qualifications) and thence, via uniform error bounds for kernel estimators, to uniform error bounds for smoothing splines and, via strong approximations, to confidence bands for the unknown regression function. It came as somewhat of a surprise that reproducing kernel Hilbert space ideas also proved useful in the study of local polynomial estimators. Throughout the text, the reproducing kernel Hilbert space approach is used as an "elementary" alternative to methods of metric entropy. It reaches its limits with least-absolute-deviations splines, where it still works, and total-variation penalization of nonparametric least-squares problems, where we miss the optimal convergence rate by a power of $\log n$ (for sample size n).

The reason for studying smoothing splines of arbitrary order is that one wants to use them for data analysis. The first question then is whether one can actually compute them. In practice, the usual scheme based on spline interpolation is useful for cubic smoothing splines only. For splines of arbitrary order, the *Kalman filter* is the bee's knees. This, in fact, is the traditional meeting ground between smoothing splines and reproducing kernel Hilbert spaces, by way of the identification of the standard smoothing problem for Gaussian processes having continuous sample paths with “generalized” smoothing spline estimation in nonparametric regression problems. We give a detailed account, culminating in the Kalman filter algorithm for spline smoothing.

The second question is how well smoothing splines of arbitrary order work. We discuss simulation results for smoothing splines and local and global polynomials for a variety of test problems. (We avoided the usual pathological examples but did include some nonsmooth examples based on the Cantor function.) We also show some results on confidence bands for the unknown regression function based on undersmoothed quintic smoothing splines with remarkably good coverage probabilities.

Acknowledgments

When we wrote the preface for Volume I, we had barely moved to our new department, Food and Resource Economics, in the College of Agriculture and Natural Resources. Having spent the last nine years here, the following assessment is time-tested: Even without the fringe benefits of easy parking, seeing the U.S. Olympic skating team practice, and enjoying (the first author, anyway) the smell of cows in the morning, we would be fortunate to be in our new surroundings. We thank the chair of the department, Tom Ilvento, for his inspiring support of the Statistics Program, and the faculty and staff of FREC for their hospitality.

As with any intellectual endeavor, we were influenced by many people and we thank them all. However, six individuals must be explicitly mentioned: first of all, Zuhair Nashed, who keeps our interest in inverse problems alive; Paul Deheuvels, for his continuing interest and encouragement in our project; Luc Devroye, despite the fact that this volume hardly mentions density estimation; David Mason, whose influence on critical parts of the manuscript speaks for itself; Randy Eubank, for his enthusiastic support of our project and subtly getting us to study the extremely effective Kalman filter; and, finally, John Kimmel, editor extraordinaire, for his continued reminders that we repeatedly promised him we would be done by next Christmas. This time around, we almost made that deadline.

Newark, Delaware *Paul Eggermont* and *Vince LaRiccia*
January 14, 2009

http://www.springer.com/978-0-387-40267-3

Maximum Penalized Likelihood Estimation
Volume II: Regression
Eggermont, P.P.; LaRiccia, V.N.
2009, XX, 572 p., Hardcover
ISBN: 978-0-387-40267-3