Preface

Getting pregnant is usually easy and fun, but the gestation and delivery may be another story; messy and painful perhaps, but instructive nevertheless. So it is with this book, which began with enthusiasm and confidence, but ten or so years later the twists and turns along the way emerge as a key part of the story.

Functional data analysis leads inevitably to dynamic systems. Ramsay and Silverman (2005) emphasized the reduction in bias and sampling variance that could be achieved by incorporating even an only approximately correct model into the penalty term by using a linear differential operator, thereby extending the more usual practice of defining roughness by the size of a high-order derivative. It was a natural next step to consider how one or more parameters that were needed to define such an operator might be estimated from data. Data associating the incidence of melanoma with solar activity became a prototype problem.

Principal differential analysis—a specific case of what we here term *gradient matching*—and its close resemblance to principal components analysis was the subject of a later chapter in Ramsay and Silverman (2005). The availability of high-resolution replicated data resulted in some rather successful applications, notably to data recording complex physical motion: handwriting and juggling. These analyses estimated low-dimensional basis systems that could be used to define the linear differential operators whose kernels were spanned by these systems which largely captured the variation in the data. This leads to the somewhat clumsy attempt, judging by limited attention that it received, to introduce the reader to Wronskians and Green's functions, not to mention such ethereal topics as reproducing kernel Hilbert spaces (we still regard the later chapters of Ramsay and Silverman (2005) as highly instructive for the so-inclined reader). In contrast, chapters on function linear regression, which did seem to appeal to readers, were for us only a partial success. Our attempts to use functions as covariates in a regression equation to approximate other functions seemed to us only marginally successful and, as a premonition of trouble to come, we were struck by how hard it was to find data to illustrate the methodologies involved.

The first serious attention given to parameter estimation for differential equations began with a collaboration with N. Heckman (Ramsay and Heckman 2000) where we noted that the large number of parameters in the smoothing function relative to the few defining the operator tended to lead to overfitting of the data and bias in the estimates of the parameters of interest. The idea of parameter cascading described in Chap. 9 came from realizing that the implicit function theorem provided a way out of this dilemma by replacing the unrestricted coefficients of the basis function expansion by a smooth function of the parameters being estimated.

We benefited enormously by a close collaboration and friendship with Kim McAuley and Jim McLellan in the Department of Chemical Engineering at Queen's University, who were able to steer us to the large literature in that field on the nonlinear least squares estimation strategy described in Chap. 7, and to pass along the nylon and refinery data used in various chapters.

We discovered again just how hard it was to find data that we could use in demonstration and test analyses in the engineering world, where data are owned by industrial concerns protected against access by competitors. When we turned to the large literature on dynamics in various fields like biology, epidemiology and physiology, we found almost no use or display of data. For example, nearly every text used the spread of disease (SIR) or the closely related Lotka-Volterra equations as a first illustration of a nonlinear system; but, if data were available at all, it was only on infected cases or predator abundances, where data fits were essentially only smooths of the data and therefore uninformative. Only recently have we discovered the invaluable archive of dynamic systems with data assembled by Klaus Schittkowski (Schittkowski 2002), which has been a great help in completing this book, and we are most grateful for his cooperation.

Why, we asked, were data-based estimations of dynamical system so hard to find? One answer seemed to be the rather restricted set of parameters yielding the solution characteristics that motivated such systems as the Lotka–Volterra, SIR, tank reactor and the FitzHugh-Nagumo, which are featured in this volume. Parameter estimation strategies, including our own, were prone to bouncing parameters into regions where they generated completely inappropriate solutions.

But another possible explanation for the paucity of data is that, in many fields, the differential equation is viewed, effectively, as data itself. That is, if there are solutions of a proposed system that exhibits the shape characteristics seen in experiments and natural settings, such as oscillations in predator-prey abundances, then these systems are considered to be demonstrations that a scientific understanding of the live system has been gained. We observed that papers on dynamics without any display of fits to data often appeared in journals like *Nature* and *Science*. Even the Hodgkin-Huxley papers, themselves exceptional examples of data-based science, were quickly followed by downgrades of their model to those like the FitzHugh–Nagumo that retained the general shape features but were not intended as data models.

Ironically, the idea of the equation as data, although rather difficult for an information scientist to warm to, dovetails beautifully with the parameter cascading algorithm that we discuss in Chaps. 9 and 10, where we allow a smooth and continuous set of compromises between data-fitting and equation-solving. This tension between data and equation is everywhere in evidence in these pages as well as the dynamical system modeling literature. It is, in fact, why we wrote this book.

Our central concept of a dynamical system as a buffer that translates sudden changes in input into smooth controlled output responses has led us to applications to data that we have previously analyzed, such as the daily Canadian weather data and the Chinese handwriting data (Ramsay 2000). We hope to have opened up entirely new opportunities for dynamical systems where none were envisaged before, which involve extensions of the functional linear model and what we call *dynamic smoothing*.

Dear reader, if you have survived to this point in this Preface, you must be wondering how much you need to know to read further. Take heart! We have worked hard to keep the technical level as low as possible, and our first goal is to bring those with little or even no exposure to differential equations as modeling objects into this fabulous data analysis landscape. Our emphasis on linear systems reflects a belief that nature is a tough place where only rugged and stable systems exhibit adaptive behavior over a wide range of environmental conditions survive.

We thank our former graduate students David Campbell and Jiguo Cao for their own versions of dynamic systems analyses. Cornell University contributed, besides the Tennebaum and Pollard team, our colleagues Steve Ellner and John Guckenheimer whose writing, experimentation and mathematical analysis continue to inform and inspire our work. We thank our Queen's collaborators Jim McClellan and Kim McAuley, as well as their students and colleagues for their guidance through the fascinating chemical engineering world and their hospitality and support.

Ottawa, Canada James Ramsay
March 2017

Springer

http://www.springer.com/978-1-4939-7188-6

Dynamic Data Analysis
Modeling Data with Differential Equations
Ramsay, J.; Hooker, G.
2017, XVII, 230 p. 84 illus., 50 illus. in color., Hardcover
ISBN: 978-1-4939-7188-6