

Preface

Among many exciting developments in statistics over the last two decades, nonlinear time series and data-analytic nonparametric methods have greatly advanced along seemingly unrelated paths. In spite of the fact that the application of nonparametric techniques in time series can be traced back to the 1940s at least, there still exists healthy and justified skepticism about the capability of nonparametric methods in time series analysis. As enthusiastic explorers of the modern nonparametric toolkit, we feel obliged to assemble together in one place the newly developed relevant techniques. The aim of this book is to advocate those modern nonparametric techniques that have proven useful for analyzing real time series data, and to provoke further research in both methodology and theory for nonparametric time series analysis.

Modern computers and the information age bring us opportunities with challenges. Technological inventions have led to the explosion in data collection (e.g., daily grocery sales, stock market trading, microarray data). The Internet makes big data warehouses readily accessible. Although classic parametric models, which postulate global structures for underlying systems, are still very useful, large data sets prompt the search for more refined structures, which leads to better understanding and approximations of the real world. Beyond postulated parametric models, there are infinite other possibilities. Nonparametric techniques provide useful exploratory tools for this venture, including the suggestion of new parametric models and the validation of existing ones.

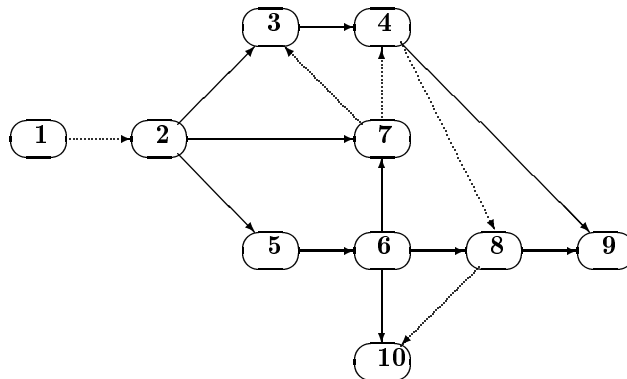
In this book, we present an up-to-date picture of techniques for analyzing time series data. Although we have tried to maintain a good balance

among methodology, theory, and numerical illustration, our primary goal is to present a comprehensive and self-contained account for each of the key methodologies. For practical relevant time series models, we aim for exposure with definition, probability properties (if possible), statistical inference methods, and numerical examples with real data sets. We also indicate where to find our (only our!) favorite computing codes to implement these statistical methods. When soliciting real-data examples, we attempt to maintain a good balance among different disciplines, although our personal interests in quantitative finance, risk management, and biology can be easily seen. It is our hope that readers can apply these techniques to their own data sets.

We trust that the book will be of interest to those coming to the area for the first time and to readers more familiar with the field. Application-oriented time series analysts will also find this book useful, as it focuses on methodology and includes several case studies with real data sets. We believe that nonparametric methods must go hand-in-hand with parametric methods in applications. In particular, parametric models provide explanatory power and concise descriptions of the underlying dynamics, which, when used sensibly, is an advantage over nonparametric models. For this reason, we have also provided a compact view of the parametric methods for both linear and selected nonlinear time series models. This will also give new comers sufficient information on the essence of the more classical approaches. We hope that this book will reflect the power of the integration of nonparametric and parametric approaches in analyzing time series data. The book has been prepared for a broad readership—the prerequisites are merely sound basic courses in probability and statistics. Although advanced mathematics has provided valuable insights into nonlinear time series, the methodological power of both nonparametric and parametric approaches can be understood without sophisticated technical details. Due to the innate nature of the subject, it is inevitable that we occasionally appeal to more advanced mathematics; such sections are marked with a “*”. Most technical arguments are collected in a “Complements” section at the end of each chapter, but key ideas are left within the body of the text.

The introduction in Chapter 1 sets the scene for the book. Chapter 2 deals with basic probabilistic properties of time series processes. The highlights include strict stationarity via ergodic Markov chains (§2.1) and mixing properties (§2.6). We also provide a generic central limit theorem for kernel-based nonparametric regression estimation for α -mixing processes. A compact view of linear ARMA models is given in Chapter 3, including Gaussian MLE (§3.3), model selection criteria (§3.4), and linear forecasting with ARIMA models (§3.7). Chapter 4 introduces three types of parametric nonlinear models. An introduction on threshold models that emphasizes developments after Tong (1990) is provided. ARCH and GARCH models are presented in detail, as they are less exposed in statistical literature. The chapter concludes with a brief account of bilinear models. Chapter 5

introduces the nonparametric kernel density estimation. This is arguably the simplest problem for understanding nonparametric techniques. The relation between “localization” for nonparametric problems and “whitening” for time series data is elucidated in §5.3. Applications of nonparametric techniques for estimating time trends and univariate autoregressive functions can be found in Chapter 6. The ideas in Chapter 5 and §6.3 provide a foundation for the nonparametric techniques introduced in the rest of the book. Chapter 7 introduces spectral density estimation and nonparametric procedures for testing whether a series is white noise. Various high-order autoregressive models are highlighted in Chapter 8. In particular, techniques for estimating nonparametric functions in FAR models are introduced in §8.3. The additive autoregressive model is exposed in §8.5, and methods for estimating conditional variance or volatility functions are detailed in §8.7. Chapter 9 outlines approaches to testing a parametric family of models against a family of structured nonparametric models. The wide applicability of the generalized likelihood ratio test is emphasized. Chapter 10 deals with nonlinear prediction. It highlights the features that distinguish nonlinear prediction from linear prediction. It also introduces nonparametric estimation for conditional predictive distribution functions and conditional minimum volume predictive intervals.



The interdependence of the chapters is depicted above, where solid directed lines indicate prerequisites and dotted lines indicate weak associations. For lengthy chapters, the dependence among sections is not very strong. For example, the sections in Chapter 4 are fairly independent, and so are those in Chapter 8 (except that §8.4 depends on §8.3, and §8.7 depends on the rest). They can be read independently. Chapter 5 and §6.3 provide a useful background for nonparametric techniques. With an understanding of this material, readers can jump directly to sections in Chapters 8 and 9. For readers who wish to obtain an overall impression of the book, we suggest reading Chapter 1, §2.1, §2.2, Chapter 3, §4.1, §4.2, Chapter 5,

§6.3, §8.3, §8.5, §8.7, §9.1, §9.2, §9.4, §9.5 and §10.1. These core materials may serve as the text for a graduate course on nonlinear time series.

Although the scope of the book is wide, we have not achieved completeness. The nonparametric methods are mostly centered around kernel/local polynomial based smoothing. Nonparametric hypothesis testing with structured nonparametric alternatives is mainly confined to the generalized likelihood ratio test. In fact, many techniques that are introduced in this book have not been formally explored mathematically. State-space models are only mentioned briefly within the discussion on bilinear models and stochastic volatility models. Multivariate time series analysis is untouched. Another noticeable gap is the lack of exposure of the variety of parametric nonlinear time series models listed in Chapter 3 of Tong (1990). This is undoubtedly a shortcoming. In spite of the important initial progress, we feel that the methods and theory of statistical inference for some of those models are not as well-established as, for example, ARCH/GARCH models or threshold models. Their potential applications should be further explored.

Extensive effort was expended in the composition of the reference list, which, together with the bibliographical notes, should guide readers to a wealth of available materials. Although our reference list is long, it merely reflects our immediate interests. Many important papers that do not fit our presentation have been omitted. Other omissions and discrepancies are inevitable. We apologize for their occurrence.

Although we both share the responsibility for the whole book, Jianqing Fan was the lead author for Chapters 1 and 5–9 and Qiwei Yao for Chapters 2–4 and 10.

Many people have been of great help to our work on this book. In particular, we would like to thank Hong-Zhi An, Peter Bickel, Peter Brockwell, Yuzhi Cai, Zongwu Cai, Kung-Sik Chan, Cees Diks, Rainer Dahlhaus, Liudas Giraitis, Peter Hall, Wai-Keung Li, Jianzhong Lin, Heng Peng, Liang Peng, Stathis Paparoditis, Wolfgang Polonik, John Rice, Peter Robinson, Richard Smith, Howell Tong, Yingcun Xia, Chongqi Zhang, Wenyang Zhang, and anonymous reviewers. Thanks also go to *Biometrika* for permission to reproduce Figure 6.10, to Blackwell Publishers Ltd. for permission to reproduce Figures 8.8, 8.15, 8.16, to *Journal of American Statistical Association* for permission to reproduce Figures 8.2 – 8.5, 9.1, 9.2, 9.5, and 10.4 – 10.12, and to World Scientific Publishing Co, Inc. for permission to reproduce Figures 10.2 and 10.3.

Jianqing Fan's research was partially supported by the National Science Foundation and National Institutes of Health of the USA and the Research Grant Council of the Hong Kong Special Administrative Region. Qiwei Yao's work was partially supported by the Engineering and Physical Sciences Research Council and the Biotechnology and Biological Sciences Research Council of the UK. This book was written while Jianqing Fan was employed by the University of California at Los Angeles, the University of

North Carolina at Chapel Hill, and the Chinese University of Hong Kong, and while Qiwei Yao was employed by the University of Kent at Canterbury and the London School of Economics and Political Science. We acknowledge the generous support and inspiration of our colleagues. Last but not least, we would like to take this opportunity to express our gratitude to all our collaborators for their friendly and stimulating collaboration. Many of their ideas and efforts have been reflected in this book.

December 2002

Jianqing Fan
Qiwei Yao

Contents

Preface	vii
1 Introduction	1
1.1 Examples of Time Series	1
1.2 Objectives of Time Series Analysis	9
1.3 Linear Time Series Models	10
1.3.1 White Noise Processes	10
1.3.2 AR Models	10
1.3.3 MA Models	12
1.3.4 ARMA Models	12
1.3.5 ARIMA Models	13
1.4 What Is a Nonlinear Time Series?	14
1.5 Nonlinear Time Series Models	16
1.5.1 A Simple Example	16
1.5.2 ARCH Models	17
1.5.3 Threshold Models	18
1.5.4 Nonparametric Autoregressive Models	18
1.6 From Linear to Nonlinear Models	20
1.6.1 Local Linear Modeling	20
1.6.2 Global Spline Approximation	23
1.6.3 Goodness-of-Fit Tests	24
1.7 Further Reading	25
1.8 Software Implementations	27

2	Characteristics of Time Series	29
2.1	Stationarity	29
2.1.1	Definition	29
2.1.2	Stationary ARMA Processes	30
2.1.3	Stationary Gaussian Processes	32
2.1.4	Ergodic Nonlinear Models*	33
2.1.5	Stationary ARCH Processes	37
2.2	Autocorrelation	38
2.2.1	Autocovariance and Autocorrelation	39
2.2.2	Estimation of ACVF and ACF	41
2.2.3	Partial Autocorrelation	43
2.2.4	ACF Plots, PACF Plots, and Examples	45
2.3	Spectral Distributions	48
2.3.1	Periodic Processes	49
2.3.2	Spectral Densities	51
2.3.3	Linear Filters	55
2.4	Periodogram	60
2.4.1	Discrete Fourier Transforms	60
2.4.2	Periodogram	62
2.5	Long-Memory Processes*	64
2.5.1	Fractionally Integrated Noise	65
2.5.2	Fractionally Integrated ARMA processes	66
2.6	Mixing*	67
2.6.1	Mixing Conditions	68
2.6.2	Inequalities	71
2.6.3	Limit Theorems for α -Mixing Processes	74
2.6.4	A Central Limit Theorem for Nonparametric Regression	76
2.7	Complements	78
2.7.1	Proof of Theorem 2.5(i)	78
2.7.2	Proof of Proposition 2.3(i)	79
2.7.3	Proof of Theorem 2.9	79
2.7.4	Proof of Theorem 2.10	80
2.7.5	Proof of Theorem 2.13	81
2.7.6	Proof of Theorem 2.14	81
2.7.7	Proof of Theorem 2.22	84
2.8	Additional Bibliographical Notes	87
3	ARMA Modeling and Forecasting	89
3.1	Models and Background	89
3.2	The Best Linear Prediction—Prewhitening	91
3.3	Maximum Likelihood Estimation	93
3.3.1	Estimators	93
3.3.2	Asymptotic Properties	97
3.3.3	Confidence Intervals	99

3.4	Order Determination	99
3.4.1	Akaike Information Criterion	100
3.4.2	FPE Criterion for AR Modeling	102
3.4.3	Bayesian Information Criterion	103
3.4.4	Model Identification	104
3.5	Diagnostic Checking	110
3.5.1	Standardized Residuals	110
3.5.2	Visual Diagnostic	110
3.5.3	Tests for Whiteness	111
3.6	A Real Data Example—Analyzing German Egg Prices	113
3.7	Linear Forecasting	117
3.7.1	The Least Squares Predictors	117
3.7.2	Forecasting in AR Processes	118
3.7.3	Mean Squared Predictive Errors for AR Processes	119
3.7.4	Forecasting in ARMA Processes	120
4	Parametric Nonlinear Time Series Models	125
4.1	Threshold Models	125
4.1.1	Threshold Autoregressive Models	126
4.1.2	Estimation and Model Identification	131
4.1.3	Tests for Linearity	134
4.1.4	Case Studies with Canadian Lynx Data	136
4.2	ARCH and GARCH Models	143
4.2.1	Basic Properties of ARCH Processes	143
4.2.2	Basic Properties of GARCH Processes	147
4.2.3	Estimation	156
4.2.4	Asymptotic Properties of Conditional MLEs*	161
4.2.5	Bootstrap Confidence Intervals	163
4.2.6	Testing for the ARCH Effect	165
4.2.7	ARCH Modeling of Financial Data	168
4.2.8	A Numerical Example: Modeling S&P 500 Index Returns	171
4.2.9	Stochastic Volatility Models	179
4.3	Bilinear Models	181
4.3.1	A Simple Example	182
4.3.2	Markovian Representation	184
4.3.3	Probabilistic Properties*	185
4.3.4	Maximum Likelihood Estimation	189
4.3.5	Bispectrum	189
4.4	Additional Bibliographical notes	191
5	Nonparametric Density Estimation	193
5.1	Introduction	193
5.2	Kernel Density Estimation	194
5.3	Windowing and Whitening	197

5.4	Bandwidth Selection	199
5.5	Boundary Correction	202
5.6	Asymptotic Results*	204
5.7	Complements—Proof of Theorem 5.3	211
5.8	Bibliographical Notes	212
6	Smoothing in Time Series	215
6.1	Introduction	215
6.2	Smoothing in the Time Domain	215
6.2.1	Trend and Seasonal Components	215
6.2.2	Moving Averages	217
6.2.3	Kernel Smoothing	218
6.2.4	Variations of Kernel Smoothers	220
6.2.5	Filtering	221
6.2.6	Local Linear Smoothing	222
6.2.7	Other Smoothing Methods	224
6.2.8	Seasonal Adjustments	224
6.2.9	Theoretical Aspects*	225
6.3	Smoothing in the State Domain	228
6.3.1	Nonparametric Autoregression	228
6.3.2	Local Polynomial Fitting	230
6.3.3	Properties of the Local Polynomial Estimator	234
6.3.4	Standard Errors and Estimated Bias	241
6.3.5	Bandwidth Selection	243
6.4	Spline Methods	246
6.4.1	Polynomial Splines	247
6.4.2	Nonquadratic Penalized Splines	249
6.4.3	Smoothing Splines	251
6.5	Estimation of Conditional Densities	253
6.5.1	Methods of Estimation	253
6.5.2	Asymptotic Properties*	256
6.6	Complements	257
6.6.1	Proof of Theorem 6.1	257
6.6.2	Conditions and Proof of Theorem 6.3	260
6.6.3	Proof of Lemma 6.1	266
6.6.4	Proof of Theorem 6.5	268
6.6.5	Proof for Theorems 6.6 and 6.7	269
6.7	Bibliographical Notes	271
7	Spectral Density Estimation and Its Applications	275
7.1	Introduction	275
7.2	Tapering, Kernel Estimation, and Prewhitening	276
7.2.1	Tapering	277
7.2.2	Smoothing the Periodogram	281
7.2.3	Prewhitening and Bias Reduction	282

7.3	Automatic Estimation of Spectral Density	283
7.3.1	Least-Squares Estimators and Bandwidth Selection	284
7.3.2	Local Maximum Likelihood Estimator	286
7.3.3	Confidence Intervals	289
7.4	Tests for White Noise	296
7.4.1	Fisher's Test	296
7.4.2	Generalized Likelihood Ratio Test	298
7.4.3	χ^2 -Test and the Adaptive Neyman Test	300
7.4.4	Other Smoothing-Based Tests	302
7.4.5	Numerical Examples	303
7.5	Complements	304
7.5.1	Conditions for Theorems 7.1—7.3	304
7.5.2	Lemmas	305
7.5.3	Proof of Theorem 7.1	306
7.5.4	Proof of Theorem 7.2	307
7.5.5	Proof of Theorem 7.3	307
7.6	Bibliographical Notes	310
8	Nonparametric Models	313
8.1	Introduction	313
8.2	Multivariate Local Polynomial Regression	314
8.2.1	Multivariate Kernel Functions	314
8.2.2	Multivariate Local Linear Regression	316
8.2.3	Multivariate Local Quadratic Regression	317
8.3	Functional-Coefficient Autoregressive Model	318
8.3.1	The Model	318
8.3.2	Relation to Stochastic Regression	318
8.3.3	Ergodicity*	319
8.3.4	Estimation of Coefficient Functions	321
8.3.5	Selection of Bandwidth and Model-Dependent Variable	322
8.3.6	Prediction	324
8.3.7	Examples	324
8.3.8	Sampling Properties*	332
8.4	Adaptive Functional-Coefficient Autoregressive Models	333
8.4.1	The Models	334
8.4.2	Existence and Identifiability	335
8.4.3	Profile Least-Squares Estimation	337
8.4.4	Bandwidth Selection	340
8.4.5	Variable Selection	340
8.4.6	Implementation	341
8.4.7	Examples	343
8.4.8	Extensions	349
8.5	Additive Models	349
8.5.1	The Models	349
8.5.2	The Backfitting Algorithm	350

8.5.3	Projections and Average Surface Estimators	352
8.5.4	Estimability of Coefficient Functions	354
8.5.5	Bandwidth Selection	355
8.5.6	Examples	356
8.6	Other Nonparametric Models	364
8.6.1	Two-Term Interaction Models	365
8.6.2	Partially Linear Models	366
8.6.3	Single-Index Models	367
8.6.4	Multiple-Index Models	368
8.6.5	An Analysis of Environmental Data	371
8.7	Modeling Conditional Variance	374
8.7.1	Methods of Estimating Conditional Variance	375
8.7.2	Univariate Setting	376
8.7.3	Functional-Coefficient Models	382
8.7.4	Additive Models	382
8.7.5	Product Models	384
8.7.6	Other Nonparametric Models	384
8.8	Complements	384
8.8.1	Proof of Theorem 8.1	384
8.8.2	Technical Conditions for Theorems 8.2 and 8.3	386
8.8.3	Preliminaries to the Proof of Theorem 8.3	387
8.8.4	Proof of Theorem 8.3	390
8.8.5	Proof of Theorem 8.4	392
8.8.6	Conditions of Theorem 8.5	394
8.8.7	Proof of Theorem 8.5	395
8.9	Bibliographical Notes	399
9	Model Validation	405
9.1	Introduction	405
9.2	Generalized Likelihood Ratio Tests	406
9.2.1	Introduction	406
9.2.2	Generalized Likelihood Ratio Test	408
9.2.3	Null Distributions and the Bootstrap	409
9.2.4	Power of the GLR Test	414
9.2.5	Bias Reduction	414
9.2.6	Nonparametric versus Nonparametric Models	415
9.2.7	Choice of Bandwidth	416
9.2.8	A Numerical Example	417
9.3	Tests on Spectral Densities	419
9.3.1	Relation with Nonparametric Regression	421
9.3.2	Generalized Likelihood Ratio Tests	421
9.3.3	Other Nonparametric Methods	425
9.3.4	Tests Based on Rescaled Periodogram	427
9.4	Autoregressive versus Nonparametric Models	430
9.4.1	Functional-Coefficient Alternatives	430

9.4.2	Additive Alternatives	434
9.5	Threshold Models versus Varying-Coefficient Models	437
9.6	Bibliographical Notes	439
10	Nonlinear Prediction	441
10.1	Features of Nonlinear Prediction	441
10.1.1	Decomposition for Mean Square Predictive Errors	441
10.1.2	Noise Amplification	444
10.1.3	Sensitivity to Initial Values	445
10.1.4	Multiple-Step Prediction versus a One-Step Plug-in Method	447
10.1.5	Nonlinear versus Linear Prediction	448
10.2	Point Prediction	450
10.2.1	Local Linear Predictors	450
10.2.2	An Example	451
10.3	Estimating Predictive Distributions	454
10.3.1	Local Logistic Estimator	455
10.3.2	Adjusted Nadaraya–Watson Estimator	456
10.3.3	Bootstrap Bandwidth Selection	457
10.3.4	Numerical Examples	458
10.3.5	Asymptotic Properties	463
10.3.6	Sensitivity to Initial Values: A Conditional Distribu- tion Approach	466
10.4	Interval Predictors and Predictive Sets	470
10.4.1	Minimum-Length Predictive Sets	471
10.4.2	Estimation of Minimum-Length Predictors	474
10.4.3	Numerical Examples	476
10.5	Complements	482
10.6	Additional Bibliographical Notes	485
	References	487
	Author index	537
	Subject index	545