archive ouverte UNIGE
http://archive-ouverte.unige.ch

Book

Nonparametric and Semiparametric Models

HÄRDLE, Wolfgang, *et al*.

Reference

HÄRDLE, Wolfgang, *et al*. *Nonparametric and Semiparametric Models*. Springer, 2004

Available at:
http://archive-ouverte.unige.ch/unige:111629

Disclaimer: layout of this document may differ from the published version.

Contents

Preface 7

Notation 11

1 Introduction 19

1.1 Density Estimation 19

1.2 Regression 20

1.2.1 Parametric Regression 22

1.2.2 Nonparametric Regression 25

1.2.3 Semiparametric Regression 27

Summary 37

I Nonparametric Models 39

2 Histogram 41

2.1 Motivation and Derivation 41

2.1.1 Construction 41

2.1.2 Derivation 43

2.1.3 Varying the Binwidth 44

2.2 Statistical Properties 45

2.2.1 Bias 46

2.2.2 Variance 47

2.2.3 Mean Squared Error 48

2.2.4 Mean Integrated Squared Error 49

2.2.5 Optimal Binwidth 50
2.3 Dependence of the Histogram on the Origin 51
2.4 Averaged Shifted Histogram 52
Bibliographic Notes 56
Exercises 57
Summary 59

3 Nonparametric Density Estimation **61**
3.1 Motivation and Derivation 61
3.1.1 Introduction 61
3.1.2 Derivation 62
3.1.3 Varying the Bandwidth 64
3.1.4 Varying the Kernel Function 64
3.1.5 Kernel Density Estimation as a Sum of Bumps 67
3.2 Statistical Properties 69
3.2.1 Bias 69
3.2.2 Variance 71
3.2.3 Mean Squared Error 72
3.2.4 Mean Integrated Squared Error 73
3.3 Smoothing Parameter Selection 74
3.3.1 Silverman's Rule of Thumb 74
3.3.2 Cross-Validation 76
3.3.3 Refined Plug-in Methods 79
3.3.4 An Optimal Bandwidth Selector?! 80
3.4 Choosing the Kernel 81
3.4.1 Canonical Kernels and Bandwidths 81
3.4.2 Adjusting Bandwidths across Kernels 83
3.4.3 Optimizing the Kernel 84
3.5 Confidence Intervals and Confidence Bands 85
3.6 Multivariate Kernel Density Estimation 90
3.6.1 Bias, Variance and Asymptotics 95

3.6.2 Bandwidth selection 97
3.6.3 Computation and Graphical Representation 100
Bibliographic Notes 104
Exercises 105
Summary 107

4 Nonparametric Regression 109
4.1 Univariate Kernel Regression 109
4.1.1 Introduction 109
4.1.2 Kernel Regression 113
4.1.3 Local Polynomial Regression and Derivative Estimation 119
4.2 Other Smoothers 124
4.2.1 Nearest-Neighbor Estimator 124
4.2.2 Median Smoothing 126
4.2.3 Smoothing Splines 126
4.2.4 Orthogonal Series 130
4.3 Smoothing Parameter Selection 133
4.3.1 A Closer Look at the Averaged Squared Error 136
4.3.2 Cross-Validation 140
4.3.3 Penalizing Functions 141
4.4 Confidence Regions and Tests 145
4.4.1 Pointwise Confidence Intervals 145
4.4.2 Confidence Bands 147
4.4.3 Hypothesis Testing 150
4.5 Multivariate Kernel Regression 156
4.5.1 Statistical Properties 157
4.5.2 Practical Aspects 160
Bibliographic Notes 163
Exercises 165
Summary 167

II Semiparametric Models **171**

5 Semiparametric and Generalized Regression Models **173**

5.1 Dimension Reduction 173
5.1.1 Variable Selection in Nonparametric Regression 176
5.1.2 Nonparametric Link Function 177
5.1.3 Semi- or Nonparametric Index 177
5.2 Generalized Linear Models 180
5.2.1 Exponential Families 180
5.2.2 Link Functions 182
5.2.3 Iteratively Reweighted Least Squares Algorithm 183
Bibliographic Notes 191
Exercises 193
Summary 194

6 Single Index Models **195**

6.1 Identification 196
6.2 Estimation 198
6.2.1 Semiparametric Least Squares 200
6.2.2 Pseudo Likelihood Estimation 203
6.2.3 Weighted Average Derivative Estimation 207
6.3 Testing the SIM 212
Bibliographic Notes 215
Exercises 216
Summary 217

7 Generalized Partial Linear Models **219**

7.1 Partial Linear Models 219
7.2 Estimation Algorithms for PLM and GPLM 221
7.2.1 Profile Likelihood 222
7.2.2 Generalized Speckman Estimator 225
7.2.3 Backfitting 228

7.2.4 Computational Issues 230
7.3 Testing the GPLM 234
7.3.1 A LR Test with Approximate Degrees of Freedom . . . 234
7.3.2 A Modified LR Test 236
Bibliographic Notes 239
Exercises 240
Summary 241

8 Additive Models and Marginal Effects 243
8.1 Backfitting 244
8.1.1 Classical Backfitting 245
8.1.2 Modified Backfitting 252
8.1.3 Consistent Backfitting 253
8.2 Marginal Integration Estimator 255
8.2.1 Estimation of Marginal Effects 256
8.2.2 Derivative Estimation for the Marginal Effects 258
8.2.3 Interaction Terms 260
8.3 Finite Sample Behavior 267
8.3.1 Bandwidth Choice 269
8.3.2 MASE in Finite Samples 270
8.3.3 Equivalent Kernel Weights 272
Bibliographic Notes 280
Exercises 281
Summary 283

9 Generalized Additive Models 285
9.1 Additive Partial Linear Models 286
9.2 Additive Models with Known Link 292
9.2.1 GAM using Backfitting 293
9.2.2 GAM using Marginal Integration 295
9.3 Generalized Additive Partial Linear Models 297

9.3.1 GAPLM using Backfitting 297
9.3.2 GAPLM using Marginal Integration 297
9.4 Testing in Additive Models, GAM, and GAPLM 301
Bibliographic Notes 308
Exercises 309
Summary 310

Bibliography **313**

Author Index **327**

Index **331**