

Clive Loader

Local Regression and Likelihood

With 68 Figures



Springer

Contents

Preface	v
1 The Origins of Local Regression	1
1.1 The Problem of Graduation	1
1.1.1 Graduation Using Summation Formulae	2
1.1.2 The Bias-Variance Trade-Off	7
1.2 Local Polynomial Fitting	7
1.2.1 Optimal Weights	8
1.3 Smoothing of Time Series	10
1.4 Modern Local Regression	11
1.5 Exercises	12
2 Local Regression Methods	15
2.1 The Local Regression Estimate	15
2.1.1 Interpreting the Local Regression Estimate	18
2.1.2 Multivariate Local Regression	19
2.2 The Components of Local Regression	20
2.2.1 Bandwidth	20
2.2.2 Local Polynomial Degree	22
2.2.3 The Weight Function	23
2.2.4 The Fitting Criterion	24
2.3 Diagnostics and Goodness of Fit	24
2.3.1 Residuals	25
2.3.2 Influence, Variance and Degrees of Freedom	27

2.3.3	Confidence Intervals	29
2.4	Model Comparison and Selection	30
2.4.1	Prediction and Cross Validation	30
2.4.2	Estimation Error and CP	31
2.4.3	Cross Validation Plots	32
2.5	Linear Estimation	33
2.5.1	Influence, Variance and Degrees of Freedom	35
2.5.2	Bias	37
2.6	Asymptotic Approximations	38
2.7	Exercises	42
3	Fitting with LOCFIT	45
3.1	Local Regression with LOCFIT	45
3.2	Customizing the Local Fit	47
3.3	The Computational Model	48
3.4	Diagnostics	49
3.4.1	Residuals	49
3.4.2	Cross Validation	49
3.5	Multivariate Fitting and Visualization	51
3.5.1	Additive Models	53
3.5.2	Conditionally Parametric Models	55
3.6	Exercises	57
4	Local Likelihood Estimation	59
4.1	The Local Likelihood Model	59
4.2	Local Likelihood with LOCFIT	62
4.3	Diagnostics for Local Likelihood	66
4.3.1	Deviance	66
4.3.2	Residuals for Local Likelihood	67
4.3.3	Cross Validation and AIC	68
4.3.4	Overdispersion	70
4.4	Theory for Local Likelihood Estimation	72
4.4.1	Why Maximize the Local Likelihood?	72
4.4.2	Local Likelihood Equations	72
4.4.3	Bias, Variance and Influence	74
4.5	Exercises	76
5	Density Estimation	79
5.1	Local Likelihood Density Estimation	79
5.1.1	Higher Order Kernels	81
5.1.2	Poisson Process Rate Estimation	82
5.1.3	Discrete Data	82
5.2	Density Estimation in LOCFIT	83
5.2.1	Multivariate Density Examples	86
5.3	Diagnostics for Density Estimation	87

5.3.1	Residuals for Density Estimation	88
5.3.2	Influence, Cross Validation and AIC	90
5.3.3	Squared Error Methods	92
5.3.4	Implementation	93
5.4	Some Theory for Density Estimation	95
5.4.1	Motivation for the Likelihood	95
5.4.2	Existence and Uniqueness	96
5.4.3	Asymptotic Representation	97
5.5	Exercises	98
6	Flexible Local Regression	101
6.1	Derivative Estimation	101
6.1.1	Identifiability and Derivative Estimation	102
6.1.2	Local Slope Estimation in LOCFIT	104
6.2	Angular and Periodic Data	105
6.3	One-Sided Smoothing	110
6.4	Robust Smoothing	113
6.4.1	Choice of Robustness Criterion	114
6.4.2	Choice of Scale Estimate	115
6.4.3	LOCFIT Implementation	115
6.5	Exercises	116
7	Survival and Failure Time Analysis	119
7.1	Hazard Rate Estimation	120
7.1.1	Censored Survival Data	120
7.1.2	The Local Likelihood Model	121
7.1.3	Hazard Rate Estimation in LOCFIT	122
7.1.4	Covariates	123
7.2	Censored Regression	124
7.2.1	Transformations and Estimates	126
7.2.2	Nonparametric Transformations	127
7.3	Censored Local Likelihood	129
7.3.1	Censored Local Likelihood in LOCFIT	130
7.4	Exercises	135
8	Discrimination and Classification	138
8.1	Discriminant Analysis	139
8.2	Classification with LOCFIT	140
8.2.1	Logistic Regression	140
8.2.2	Density Estimation	142
8.3	Model Selection for Classification	144
8.4	Multiple Classes	147
8.5	More on Misclassification Rates	151
8.5.1	Pointwise Misclassification	152
8.5.2	Global Misclassification	153

8.6	Exercises	155
9	Variance Estimation and Goodness of Fit	157
9.1	Variance Estimation	157
9.1.1	Other Variance Estimates	159
9.1.2	Nonhomogeneous Variance	160
9.1.3	Goodness of Fit Testing	163
9.2	Interval Estimation	165
9.2.1	Pointwise Confidence Intervals	165
9.2.2	Simultaneous Confidence Bands	166
9.2.3	Likelihood Models	169
9.2.4	Maximal Deviation Tests	170
9.3	Exercises	172
10	Bandwidth Selection	174
10.1	Approaches to Bandwidth Selection	175
10.1.1	Classical Approaches	175
10.1.2	Plug-In Approaches	176
10.2	Application of the Bandwidth Selectors	179
10.2.1	Old Faithful	180
10.2.2	The Claw Density	183
10.2.3	Australian Institute of Sport Dataset	186
10.3	Conclusions and Further Reading	188
10.4	Exercises	190
11	Adaptive Parameter Choice	192
11.1	Local Goodness of Fit	193
11.1.1	Local CP	193
11.1.2	Local Cross Validation	195
11.1.3	Intersection of Confidence Intervals	196
11.1.4	Local Likelihood	196
11.2	Fitting Locally Adaptive Models	197
11.3	Exercises	205
12	Computational Methods	206
12.1	Local Fitting at a Point	206
12.2	Evaluation Structures	208
12.2.1	Growing Adaptive Trees	209
12.2.2	Interpolation Methods	212
12.2.3	Evaluation Structures in LOCFIT	214
12.3	Influence and Variance Functions	215
12.4	Density Estimation	216
12.5	Exercises	217

13 Optimizing Local Regression	220
13.1 Optimal Rates of Convergence	220
13.2 Optimal Constants	224
13.3 Minimax Local Regression	227
13.3.1 Implementation	229
13.4 Design Adaptation and Model Indexing	231
13.5 Exercises	233
A Installing LOCFIT in R, S and S-Plus	235
A.1 Installation, S-Plus for Windows	235
A.2 Installation, S-Plus 3, UNIX	236
A.3 Installation, S-Plus 5.0	237
A.4 Installing in R	238
B Additional Features: LOCFIT in S	239
B.1 Prediction	239
B.2 Calling <code>locfit()</code>	240
B.2.1 Extracting from a Fit	240
B.2.2 Iterative Use of <code>locfit()</code>	241
B.3 Arithmetic Operators and Math Functions	243
B.4 Trellis Tricks	244
C C-LOCFIT	246
C.1 Installation	246
C.1.1 Windows 95, 98 and NT	246
C.1.2 UNIX	246
C.2 Using C-LOCFIT	247
C.2.1 Data in C-LOCFIT	248
C.3 Fitting with C-LOCFIT	250
C.4 Prediction	251
C.5 Some additional commands	251
D Plots from C-LOCFIT	252
D.1 The <code>plotdata</code> Command	253
D.2 The <code>plotfit</code> Command	253
D.3 Other Plot Options	256
References	257
Index	279