

Contents

Preface	vii
1 Some Regression Examples	1
1.1 Influence and Outliers	1
1.2 Three Examples	2
1.2.1 Forbes' Data	2
1.2.2 Multiple Regression Data	5
1.2.3 Wool Data	9
1.3 Checking and Building Models	14
2 Regression and the Forward Search	16
2.1 Least Squares	16
2.1.1 Parameter Estimates	16
2.1.2 Residuals and Leverage	18
2.1.3 Formal Tests	19
2.2 Added Variables	20
2.3 Deletion Diagnostics	22
2.3.1 The Algebra of Deletion	22
2.3.2 Deletion Residuals	23
2.3.3 Cook's Distance	24
2.4 The Mean Shift Outlier Model	26
2.5 Simulation Envelopes	27
2.6 The Forward Search	28
2.6.1 General Principles	28

2.6.2	Step 1: Choice of the Initial Subset	31
2.6.3	Step 2: Adding Observations During the Forward Search	32
2.6.4	Step 3: Monitoring the Search	33
2.6.5	Forward Deletion Formulae	34
2.7	Further Reading	35
2.8	Exercises	36
2.9	Solutions	37
3	Regression	43
3.1	Hawkins' Data	43
3.2	Stack Loss Data	50
3.3	Salinity Data	62
3.4	Ozone Data	67
3.5	Exercises	73
3.6	Solutions	74
4	Transformations to Normality	81
4.1	Background	81
4.2	Transformations in Regression	82
4.2.1	Transformation of the Response	82
4.2.2	Graphics for Transformations	86
4.2.3	Transformation of an Explanatory Variable	87
4.3	Wool Data	88
4.4	Poison Data	95
4.5	Modified Poison Data	98
4.6	Doubly Modified Poison Data: An Example of Masking	101
4.7	Multiply Modified Poison Data—More Masking	104
4.7.1	A Diagnostic Analysis	104
4.7.2	A Forward Analysis	106
4.7.3	Other Graphics for Transformations	108
4.8	Ozone Data	110
4.9	Stack Loss Data	111
4.10	Mussels' Muscles: Transformation of the Response	116
4.11	Transforming Both Sides of a Model	121
4.12	Shortleaf Pine	124
4.13	Other Transformations and Further Reading	127
4.14	Exercises	128
4.15	Solutions	129
5	Nonlinear Least Squares	136
5.1	Background	137
5.1.1	Nonlinear Models	137
5.1.2	Curvature	141

5.2	The Forward Search	148
	5.2.1 Parameter Estimation	148
	5.2.2 Monitoring the Forward Search	150
5.3	Radioactivity and Molar Concentration of Nifedipene	151
5.4	Enzyme Kinetics	154
5.5	Calcium Uptake	159
5.6	Nitrogen in Lakes	164
5.7	Isomerization of <i>n</i> -Pentane	170
5.8	Related Literature	173
5.9	Exercises	174
5.10	Solutions	176
6	Generalized Linear Models	179
6.1	Background	180
	6.1.1 British Train Accidents	180
	6.1.2 Bliss's Beetle Data	181
	6.1.3 The Link Function	181
6.2	The Exponential Family	185
6.3	Mean, Variance, and Likelihood	185
	6.3.1 One Observation	185
	6.3.2 The Variance Function	186
	6.3.3 Canonical Parameterization	188
	6.3.4 The Likelihood	188
6.4	Maximum Likelihood Estimation	189
	6.4.1 Least Squares	189
	6.4.2 Weighted Least Squares	190
	6.4.3 Newton's Method for Solving Equations	190
	6.4.4 Fisher Scoring	191
	6.4.5 The Algorithm	192
6.5	Inference	194
	6.5.1 The Deviance	194
	6.5.2 Estimation of the Dispersion Parameter	197
	6.5.3 Inference About Parameters	197
6.6	Checking Generalized Linear Models	198
	6.6.1 The Hat Matrix	198
	6.6.2 Residuals	198
	6.6.3 Cook's Distance	200
	6.6.4 A Goodness of Link Test	200
	6.6.5 Monitoring the Forward Search	201
6.7	Gamma Models	202
6.8	Car Insurance Data	204
6.9	Dielectric Breakdown Strength	209
6.10	Poisson Models	221
6.11	British Train Accidents	222
6.12	Cellular Differentiation Data	226

6.13	Binomial Models	230
6.14	Bliss's Beetle Data	232
6.15	Mice with Convulsions	234
6.16	Toxoplasmosis and Rainfall	238
	6.16.1 A Forward Analysis	238
	6.16.2 Comparison with Backwards Methods	245
6.17	Binary Data	246
	6.17.1 Introduction: Vasoconstriction Data	246
	6.17.2 The Deviance	248
	6.17.3 The Forward Search for Binary Data	249
	6.17.4 Perfect Fit	250
6.18	Theory: The Effect of Perfect Fit and the Arcsine Link	253
6.19	Vasoconstriction Data and Perfect Fit	256
6.20	Chapman Data	259
6.21	Developments and Further Reading	265
6.22	Exercises	267
6.23	Solutions	268
A Data		277
Bibliography		311
Author Index		319
Subject Index		323