

Nonlinear Regression Analysis and Its Applications

Douglas M. Bates

Department of Statistics
University of Wisconsin
Madison, Wisconsin

Donald G. Watts

Department of Mathematics
and Statistics
Queen's University
Kingston, Ontario, Canada



WILEY

JOHN WILEY & SONS

New York Chichester Brisbane Toronto Singapore

CONTENTS

1. Review of Linear Regression	1
1.1 The Linear Regression Model / 1	
1.1.1 The Least Squares Estimates / 4	
1.1.2 Sampling Theory Inference Results / 5	
1.1.3 Likelihood Inference Results / 6	
1.1.4 Bayesian Inference Results / 7	
1.1.5 Comments / 7	
1.2 The Geometry of Linear Least Squares / 9	
1.2.1 The Expectation Surface / 9	
1.2.2 Determining the Least Squares Estimates / 12	
1.2.3 Parameter Inference Regions / 15	
1.2.3.1 The Geometry of Sampling Theory Results / 15	
1.2.3.2 Marginal Confidence Intervals / 19	
1.2.3.3 The Geometry of Likelihood Results / 23	
1.3 Assumptions and Model Assessment / 23	
1.3.1 Assumptions and Their Implications / 24	
1.3.2 Model Assessment / 26	
1.3.2.1 Plotting Residuals / 27	
1.3.2.2 Stabilizing Variance / 28	
1.3.2.3 Lack of Fit / 29	
2. Nonlinear Regression: Iterative Estimation and Linear Approximations	32
2.1 The Nonlinear Regression Model / 32	
2.1.1 Transformably Linear Models / 34	
2.1.2 Conditionally Linear Parameters / 36	
2.1.3 The Geometry of the Expectation Surface / 36	
2.2 Determining the Least Squares Estimates / 39	

- 2.2.1 The Gauss–Newton Method / 40
 - 2.2.1.1 Step Factor / 42
- 2.2.2 The Geometry of Nonlinear Least Squares / 43
- 2.2.3 Convergence / 47
- 2.3 Nonlinear Regression Inference Using the Linear Approximation / 52
 - 2.3.1 Approximate Inference Regions for Parameters / 52
 - 2.3.2 Approximate Inference Bands for the Expected Response / 58
- 2.4 Nonlinear Least Squares via Sums of Squares / 60
 - 2.4.1 The Linear Approximation / 61
 - 2.4.2 Overshoot / 64
- 2.5 Use of the Linear Approximation / 64

3. Practical Considerations in Nonlinear Regression

67

- 3.1 Model Specification / 67
 - 3.1.1 The Expectation Function / 67
 - 3.1.2 The Disturbance Term / 69
- 3.2 Preliminary Analysis / 70
- 3.3 Starting Values / 72
 - 3.3.1 Interpreting the Expectation Function Behavior / 72
 - 3.3.2 Interpreting Derivatives of the Expectation Function / 73
 - 3.3.3 Transforming the Expectation Function / 73
 - 3.3.4 Reducing Dimensions / 76
 - 3.3.5 Conditional Linearity / 76
- 3.4 Parameter Transformations / 77
 - 3.4.1 Constrained Parameters / 77
 - 3.4.2 Facilitating Convergence / 78
- 3.5 Other Iterative Techniques / 79
 - 3.5.1 A Newton–Raphson Method / 79
 - 3.5.2 The Levenberg–Marquardt Compromise / 80
 - 3.5.3 Numerical Derivatives / 82
 - 3.5.4 Derivative-free Methods / 82
 - 3.5.5 Removing Conditionally Linear Parameters / 85
- 3.6 Obtaining Convergence / 86
- 3.7 Assessing the Fit and Modifying the Model / 90
- 3.8 Correlated Residuals / 92
- 3.9 Accumulated Data / 96
 - 3.9.1 Estimating the Parameters by Direct Integration / 98
- 3.10 Comparing Models / 103
 - 3.10.1 Nested Models / 103
 - 3.10.2 Incremental Parameters and Indicator Variables / 104
 - 3.10.3 Non-nested Models / 107

- 3.11 Parameters as Functions of Other Variables / 108
- 3.12 Presenting the Results / 109
- 3.13 Nitrite Utilization: A Case Study / 110
 - 3.13.1 Preliminary Analysis / 110
 - 3.13.2 Model Selection / 113
 - 3.13.3 Starting Values / 113
 - 3.13.4 Assessing the Fit / 113
 - 3.13.5 Modifying the Model / 114
 - 3.13.6 Assessing the Fit / 116
 - 3.13.7 Reducing the Model / 116
 - 3.13.8 Assessing the Fit / 116
 - 3.13.9 Comparing the Models / 120
 - 3.13.10 Reporting the Results / 121
- 3.14 Experimental Design / 121
 - 3.14.1 General Considerations / 122
 - 3.14.2 The Determinant Criterion / 124
 - 3.14.3 Starting Designs / 125
 - 3.14.4 Sequential designs / 127
 - 3.14.5 Subset designs / 129
 - 3.14.6 Conditionally Linear Models / 129
 - 3.14.7 Other design criteria / 131

4. Multiresponse Parameter Estimation

134

- 4.1 The Multiresponse Model / 134
 - 4.1.1 The Determinant Criterion / 138
 - 4.1.2 Inferences for Multiresponse Estimation / 139
 - 4.1.3 Dimensional Considerations in Multiresponse Estimation / 140
- 4.2 A Generalized Gauss-Newton Method / 141
 - 4.2.1 The Gradient and Hessian of the Determinant / 142
 - 4.2.2 An Approximate Hessian / 143
 - 4.2.3 Calculations for Each Iteration / 144
 - 4.2.4 A Multiresponse Convergence Criterion / 145
- 4.3 Practical Considerations / 146
 - 4.3.1 Obtaining Starting Values / 146
 - 4.3.1.1 Starting Estimates for Multiresponse Models Described by Systems of Differential Equations / 147
 - 4.3.2 Assessing the Fit / 149
 - 4.3.3 Dependencies Among Responses / 154
 - 4.3.4 Linear Combinations of Responses / 158
 - 4.3.5 Comparing Models / 162
- 4.4 Missing Data / 164

5. Models Defined by Systems of Differential Equations 168

- 5.1 Compartment Models and System Diagrams / 168
- 5.2 Estimating Parameters in Compartment Models / 172
 - 5.2.1 Solving Systems of Linear Differential Equations / 173
 - 5.2.1.1 Dead Time / 175
 - 5.2.1.2 Cessation of Infusion / 177
 - 5.2.2 Derivatives of the Expectation Function / 178
- 5.3 Practical Considerations / 179
 - 5.3.1 Parameter Transformations / 179
 - 5.3.2 Identifiability / 180
 - 5.3.3 Starting Values / 182
- 5.4 Lipoproteins: A Case Study / 182
 - 5.4.1 Preliminary Analysis / 182
 - 5.4.2 One Compartment / 182
 - 5.4.3 Two Compartments / 183
 - 5.4.4 Three Compartments / 184
 - 5.4.5 Three Compartments, Common Parameters / 187
 - 5.4.6 Conclusions / 187
- 5.5 Oil shale: A Case Study / 188
 - 5.5.1 Preliminary Analysis / 188
 - 5.5.2 Starting Values for the 673 K Data / 190
 - 5.5.3 Fitting the Individual Temperature Data / 191
 - 5.5.4 Starting Estimates for the Process Parameters / 193
 - 5.5.5 Fitting the Complete Data Set / 194
 - 5.5.6 Conclusions / 198

6. Graphical Summaries of Nonlinear Inference Regions 200

- 6.1 Likelihood Regions / 200
 - 6.1.1 Joint Parameter Likelihood Regions / 200
 - 6.1.2 Profile t Plots, Profile Traces, and Profile Pair Sketches / 205
 - 6.1.2.1 Profile Traces / 207
 - 6.1.2.2 Profile Pair Sketches / 209
 - 6.1.3 Comments / 214
- 6.2 Bayes Regions / 216
 - 6.2.1 Choice of Bayes Prior on the Parameters / 216
 - 6.2.2 Joint HPD Regions / 219
- 6.3 Exact Sampling Theory Confidence Regions / 223
- 6.4 Comparison of the Likelihood, Bayes, and Sampling Theory Approaches / 229

7. Curvature Measures of Nonlinearity	232
7.1 Velocity and Acceleration Vectors / 233	
7.1.1 Tangential and Normal Accelerations / 234	
7.1.2 The Acceleration in an Arbitrary Direction / 239	
7.2 Relative Curvatures / 241	
7.2.1 Interpreting Terms in the Curvature Arrays / 245	
7.2.1.1 Intrinsic Curvature / 245	
7.2.1.2 Parameter Effects Curvatures / 246	
7.2.2 Reparametrization / 248	
7.3 RMS Curvatures / 253	
7.3.1 Calculating RMS Curvatures / 254	
7.3.2 RMS Curvatures in Practice / 256	
7.4 Direct Assessment of the Effects of Intrinsic Nonlinearity / 256	
Appendix 1. Data Sets Used in Examples	267
1.1 PCB / 267	
1.2 Rumford / 268	
1.3 Puromycin / 269	
1.4 BOD / 270	
1.5 Isomerization / 271	
1.6 α -Pinene / 272	
1.7 Sulfisoxazole / 273	
1.8 Lubricant / 274	
1.9 Chloride / 274	
1.10 Ethyl Acrylate / 274	
1.11 Saccharin / 277	
1.12 Nitrite Utilization / 278	
1.13 s-PMMA / 280	
1.14 Tetracycline / 281	
1.15 Oil Shale / 282	
1.16 Lipoproteins / 282	
Appendix 2. QR Decompositions Using Householder Transformations	286
Appendix 3. Pseudocode for Computing Algorithms	290
3.1 Nonlinear Least Squares / 290	
3.1.1 Implementation in S / 291	
3.1.2 Implementation in GAUSS / 293	

3.1.3	Implementation in SAS/IML / 296	
3.2	Linear Summaries and Studentized Residuals / 298	
3.3	Multiresponse Estimation / 299	
3.4	Linear Systems of Differential Equations / 300	
3.5	Profile Calculations / 302	
3.5.1	Generation of τ and the Profile Traces / 302	
3.5.2	Profile Pair Plots / 303	
Appendix 4.	Data Sets Used in Problems	305
4.1	BOD Data Set 2 / 305	
4.2	Nitrendipene / 306	
4.3	Saccharin Data Set 2 / 306	
4.4	Steady State Adsorption / 306	
4.5	Leaves / 310	
4.6	α -Pinene Data Set 2 / 310	
4.7	Coal Liquefaction / 311	
4.8	Haloperidol / 315	
Appendix 5.	Evaluating Matrix Exponentials and Convolutions	316
5.1	Diagonalizable A / 317	
5.2	Nondiagonalizable A / 319	
5.3	Complex Eigenvalues / 320	
Appendix 6.	Interpolating Profile Pair Contours	323
Appendix 7.	Key to Data Sets	329
References		335
Bibliography		343
Author Index		355
Subject Index		359