LINEAR REGRESSION ANALYSIS

G. A. F. SEBER

Professor of Statistics
University of Auckland
Auckland, New Zealand

JOHN WILEY & SONS

New York • Chichester • Brisbane • Toronto

Contents

1 VECTORS OF RANDOM VARIABLES 1

1.1 Notation 1
1.2 Linear Regression Models 2
1.3 Expectation and Covariance Operators 8
Exercises 1a, 12
1.4 Mean and Variance of Quadratic Forms 13
Exercises 1b, 16
1.5 Independence of Random Variables 17
Exercises 1c, 19
1.6 Chi-Squared Distribution 19

Miscellaneous Exercises 1, 20

2 MULTIVARIATE NORMAL DISTRIBUTION 22

2.1 Definition 22
Exercises 2a, 27
2.2 Moment Generating Function 27
Exercises 2b, 31
2.3 Independence of Normal Variables 32
Exercise 2c, 35
2.4 Quadratic Forms in Normal Variables 36
Exercises 2d, 40

Miscellaneous Exercises 2, 40

3 LINEAR REGRESSION: ESTIMATION AND DISTRIBUTION THEORY . . 42

3.1 Least Squares Estimation 42
Exercises 3a, 47

3.2 Properties of Least Squares Estimates 48
Exercises 3b, 50
3.3 Estimation of σ^2 51
Exercises 3c, 54
3.4 Distribution Theory 54
Exercises 3d, 56
3.5 Orthogonal Structure in the Design Matrix 56
Exercises 3e, 59
3.6 Generalized Least Squares 60
Exercises 3f, 64
3.7 Introducing Further Regressors 65
3.7.1 *General theory*, 65
3.7.2 *One extra variable*, 68
3.7.3 *Two-step least squares*, 69
3.7.4 *Two-step residuals*, 70
Exercises 3g, 71
3.8 Design Matrix of Less than Full Rank 72
3.8.1 *Least squares estimation*, 72
a. Reduce the model to one of full rank, 73
b. Impose identifiability constraints, 74
c. Compute a generalized inverse, 76
Exercises 3h, 78
3.8.2 *Estimable functions*, 79
Exercises 3i, 81
3.8.3 *Introducing further regressors*, 82
Exercises 3j, 84
3.9 Estimation with Linear Restrictions 84
3.9.1 *Method of Lagrange multipliers*, 84
3.9.2 *Method of orthogonal projections*, 86
Exercises 3k, 87
3.10 Other Methods of Estimation 88
3.10.1 *Biased estimation*, 88
3.10.2 *Nonnegative estimation*, 90
3.10.3 *Censored data*, 90
3.10.4 *Robust estimation*, 91
3.11 Optimal Design 92
Miscellaneous Exercises 3, 93

4 LINEAR REGRESSION: HYPOTHESIS TESTING 96

4.1 The *F*-Test 96

4.1.1 *Derivation*, 96

Exercises 4a, 99

4.1.2 *Motivation of the F-test*, 99

4.1.3 *Some examples*, 101

4.1.4 *The straight line*, 105

4.1.5 *Significant F-test*, 108

Exercises 4b, 109

4.2 Multiple Correlation Coefficient 110

Exercises 4c, 115

4.3 A Canonical Form For H 115

Exercises 4d, 117

4.4 Goodness of Fit Test 117

4.5 Design Matrix of Less than Full Rank 120

4.5.1 *The F-test and projection matrices*, 120

4.5.2 *Testable hypotheses*, 121

Exercises 4e, 122

4.6 Hypothesis Testing with Initial Constraints 123

Miscellaneous Exercises 4, 123

5 CONFIDENCE INTERVALS AND REGIONS 125

5.1 Simultaneous Interval Estimation 125

5.1.1 *The problem of simultaneous inferences*, 125

a. Bonferroni *t*-intervals, 126

b. Maximum modulus *t*-intervals, 127

c. Scheffé's *S*-method, 128

5.1.2 *Comparison of methods*, 130

5.1.3 *Hypothesis testing and confidence intervals*, 132

5.1.4 *Confidence regions*, 133

5.2 Confidence Bands for the Regression Surface 134

5.3 Prediction Intervals for the Response 136

5.4 Enlarging the Regression Matrix 137

Miscellaneous Exercises 5, 139

6 DEPARTURES FROM UNDERLYING ASSUMPTIONS 140

6.1 Bias 141

6.1.1 *Bias due to underfitting*, 141
6.1.2 *Bias due to overfitting*, 143
6.1.3 *Bias and random regressor variables*, 144
6.2 Incorrect Dispersion Matrix 144
6.2.1 *General case*, 144
6.2.2 *Diagonal dispersion matrix*, 146
a. Bartlett's test, 147
b. Cochran's test, 147
c. Hartley's F-max test, 148
6.3 Robustness of the F-Test to Nonnormality 149
6.3.1 *Effect of the regressor variables*, 149
6.3.2 *Quadratically balanced F-tests*, 151
6.4 Regressor Variables Measured with Error 155
6.4.1 *Random errors*, 155
a. Bias, 156
b. Standard errors, 157
6.4.2 *Roundoff errors*, 158
6.4.3 *Some working rules*, 159
6.5 Models with Random Regressors 160
6.6 Analysis of Residuals 162
6.6.1 *Definition and properties*, 162
6.6.2 *Residual plots*, 163
a. Normal probability plot, 164
b. Plot residual versus fitted value, 165
c. Plot residual against an omitted factor (e.g., time), 166
d. Plot residual against each regressor, 169
e. Plot x_j against x_k, 170
6.6.3 *Statistical tests based on residuals*, 170
6.6.4 *Partial residual plots*, 170
6.6.5 *Transformed residuals*, 172
6.7 Transforming the Data 173

Miscellaneous Exercises 6, 176

7 Straight-Line Regression 177
7.1 Introduction 177
7.2 Confidence Intervals and Bands 179
7.2.1 *Confidence intervals for slope and intercept*, 179
7.2.2 *Confidence interval for* $-\beta_0/\beta_1$, 180

7.2.3 *Prediction intervals and bands*, 182
7.2.4 *Prediction intervals for the response*, 186
7.2.5 *Optimal allocation of observations*, 186
7.2.6 *Inverse prediction (discrimination)*, 187
a. Alternative estimators, 189
b. Replicated observations, 190
7.3 Straight Line Through the Origin 191
7.4 Weighted Least Squares 193
7.4.1 *Known weights*, 193
7.4.2 *Unknown weights*, 194
a. Maximum likelihood method, 195
b. Least squares method, 197
7.5 Comparing Straight Lines 197
7.5.1 *General model*, 197
7.5.2 *Test for parallelism*, 198
7.5.3 *Test for coincidence*, 200
7.5.4 *Test for concurrence*, 201
a. x-coordinate known, 201
b. x-coordinate unknown, 203
7.5.5 *Use of dummy categorical regressors*, 204
7.5.6 *Test for equal variances*, 205
7.6 Two-Phase Linear Regression 205
7.6.1 *Least squares estimation with known changeover point*, 206
7.6.2 *Hypothesis test that changeover point is a given value*, 207
7.6.3 *Unknown changeover point*, 208
7.7 Random Regressor Variable 210

Miscellaneous Exercises 7, 212

8 Polynomial Regression 214

8.1 Polynomials in One Variable 214
8.1.1 *The problem of ill-conditioning*, 214
8.1.2 *Choosing the degree*, 217
8.2 Orthogonal Polynomials 218
8.2.1 *General statistical properties*, 218
8.2.2 *Generating orthogonal polynomials*, 220
8.2.3 *Weighted least squares*, 220
8.2.4 *Application of constraints*, 224

8.2.5 *Equally spaced x-values*, 225
8.3 Piecewise Polynomial Fitting 227
8.3.1 *Unsatisfactory fit*, 227
8.3.2 *Use of spline functions*, 228
8.3.3 *Multiphase polynomial regression*, 230
8.4 Optimal Allocation of Points 231
8.5 Polynomial Regression in Several Variables 234
8.5.1 *Surface fitting*, 234
8.5.2 *Response surfaces*, 234

Miscellaneous Exercises 8, 238

9 Analysis of Variance 239

9.1 One-Way Classification 239
9.1.1 *Representation as a regression model*, 239
9.1.2 *Computations*, 241
9.1.3 *Expected values*, 243
9.1.4 *Reparametrization of the model*, 244
9.1.5 *Geometrical considerations*, 245
9.1.6 *Identifiability constraints*, 247
9.1.7 *Confidence intervals*, 247
9.1.8 *Underlying assumptions*, 249
9.1.9 *Unequal numbers of observations per mean*, 250
9.2 Two-Way Classification 252
9.2.1 *Representation as a regression model*, 252
9.2.2 *Test statistics*, 254
9.2.3 *Analysis of variance table*, 256
9.2.4 *Confidence intervals*, 258
9.2.5 *Unequal numbers of observations per mean*, 259
9.3 Higher-Way Classifications with Equal Numbers per Mean 260
9.3.1 *Definition of interactions and main effects*, 260
9.3.2 *Hypothesis testing*, 262
9.4 Classifications with One Observation per Mean 265
9.4.1 *Derivation of test statistics*, 265
9.4.2 *Underlying assumptions*, 268
9.4.3 *An alternative approach*, 271
9.5 Designs with Simple Block Structure 271

Miscellaneous Exercises 9, 276

10 Analysis of Covariance and Missing Observations 279

10.1 Analysis of Covariance 279

10.1.1 *Least squares estimation*, 279
10.1.2 *Hypothesis testing*, 283
10.1.3 *Underlying assumptions*, 288

10.2 Missing Observations 290

10.2.1 *Estimation theory*, 290
10.2.2 *Hypothesis testing*, 293
10.2.3 *Analysis of covariance method*, 296
10.2.4 *Missing observations in analysis of covariance models*, 300

Miscellaneous Exercises 10, 300

11 Computational Techniques for Fitting a Specified Regression 302

11.1 Introduction 302

11.2 Full Rank Case 303

11.2.1 *Gaussian elimination*, 303
11.2.2 *Cholesky decomposition (square root) method*, 304
11.2.3 *Triangular decomposition*, 307
11.2.4 *Orthogonal-triangular decomposition*, 308
a. Gram–Schmidt orthogonalization, 311
b. Householder transformations, 312
c. Givens transformations, 314

11.3 Weighted Least Squares 317

11.3.1 *The normal equations*, 317
11.3.2 *Cholesky method*, 317
11.3.3 *Gram–Schmidt method*, 318
11.3.4 *Gentleman's method*, 318

11.4 Comparison of Methods 319

11.5 Less than Full Rank Case 322

11.5.1 *Cholesky method for a generalized inverse of $X'X$*, 322
11.5.2 *Abbreviated Doolittle method*, 324
11.5.3 *Generalized inverse of X using Householder transformations*, 324
11.5.4 *Orthogonal decomposition with identifiability constraints*, 325

11.5.5 *Singular value decomposition*, 327
11.6 Improving the Solution by Iterative Refinement 328
11.7 Centering and Scaling the Data 330
11.7.1 *Centering the data*, 330
11.7.2 *Scaling (equilibration)*, 334
11.8 Regression Updating 336
11.9 Adding or Removing a Specified Regressor 338
11.9.1 *Adding a regressor*, 338
11.9.2 *Removing a regressor*, 341
11.10 Hypothesis Testing 342
11.11 Checking Out Computer Programs 344
Miscellaneous Exercises 11, 346

12 CHOOSING THE "BEST" REGRESSION 349
12.1 Introduction 349
12.2 Generating All Possible Regressions 349
12.2.1 *Order of generation*, 350
12.2.2 *Method of generation*, 351
a. Sweeping, 351
b. Symmetric sweep, 354
c. Furnival's method, 355
d. Householder and Givens transformations, 357
12.2.3 *Comparing different equations*, 358
a. Coefficient of determination (R^2), 359
b. Adjusted coefficient of determination, 362
c. Mallows' C_p statistic, 364
d. The MSEP criterion, 369
e. Other measures, 369
12.3 Generating the Better Regressions Only 371
12.3.1 *Searching along favorable branches*, 371
12.3.2 *A t-directed search*, 374
12.4 Stepwise Regression 376
12.4.1 *Description of the method*, 376
12.4.2 *Use of sweeping*, 377
12.4.3 *Gauss–Jordan elimination*, 379
12.4.4 *Choice of F values*, 380
12.4.5 *Other one-step methods*, 380
12.5 Other Methods : 381

12.5.1 *Factor and principal component analysis*, 381
12.5.2 *A Bayesian method*, 381
12.6 General Comments 382

Miscellaneous Exercises 12, 382

APPENDIX 383

A Some Matrix Algebra 383
B Orthogonal Projections 394
C Normal Probability Paper 398
D Percentage Points of the Bonferroni t-Statistic 402
E Distribution of the Largest Absolute Value of k Student t Variates 404
F Working–Hotelling Confidence Bands for Finite Intervals . . 411

OUTLINE SOLUTIONS TO EXERCISES 413

REFERENCES 433

INDEX 459