

Linear Models

S. R. SEARLE

*Professor of Biological Statistics
Biometrics Unit
N.Y. State College of Agriculture
Cornell University, Ithaca, N.Y.*

John Wiley & Sons, Inc.
New York · London · Sydney · Toronto

List of Chapters

1. Generalized Inverse Matrices	1
2. Distributions and Quadratic Forms	31
3. Regression, or the Full Rank Model	75
4. Introducing Linear Models: Regression on Dummy Variables	135
5. Models Not of Full Rank	164
6. Two Elementary Models	226
7. The 2-Way Crossed Classification	261
8. Some Other Analyses	332
9. Introduction to Variance Components	376
10. Methods of Estimating Variance Components from Unbalanced Data	421
11. Variance Component Estimation from Unbalanced Data: Formulae	473
Literature Cited	515
Statistical Tables	523
Index	529

Contents

1. Generalized Inverse Matrices	1
1. Introduction	1
<i>a. Definition and existence, 1</i>	
<i>b. An algorithm, 4</i>	
2. Solving linear equations	7
<i>a. Consistent equations, 7</i>	
<i>b. Obtaining solutions, 8</i>	
<i>c. Properties of solutions, 11</i>	
3. The Penrose inverse	16
4. Other definitions	19
5. Symmetric matrices	20
<i>a. Properties of a generalized inverse, 20</i>	
<i>b. Two methods of derivation, 21</i>	
6. Arbitrariness in a generalized inverse	23
7. Other results	27
8. Exercises	28
2. Distributions and Quadratic Forms	31
1. Introduction	31
2. Symmetric matrices	33
3. Positive definiteness	34
4. Distributions	38
<i>a. Multivariate density functions, 38</i>	
<i>b. Moments, 39</i>	
<i>c. Linear transformations, 40</i>	
<i>d. Moment generating functions, 41</i>	
<i>e. Univariate normal, 42</i>	
<i>f. Multivariate normal, 43</i>	
(i) Density function, 43	
(ii) Aitken's integral, 43	
(iii) Moment generating function, 44	
(iv) Marginal distributions, 45	
(v) Conditional distributions, 46	
(vi) Independence, 47	
<i>g. Central χ^2, F and t, 47</i>	

2. Distributions and Quadratic Forms (Continued)

- h. Non-central χ^2 , 49
- i. Non-central F, 51
- j. Other non-central distributions, 53
- 5. Distribution of quadratic forms 54
 - a. Cumulants, 55
 - b. Distributions, 57
 - c. Independence, 59
- 6. Bilinear forms 64
- 7. The singular normal distribution 66
- 8. Exercises 72

3. Regression, or the Full Rank Model 75

- 1. Introduction 75
 - a. The model, 75
 - b. Observations, 76
 - c. Estimation, 77
 - d. Example, 78
 - e. The general case of k x -variables, 79
 - f. Example (continued), 81
 - g. Intercept and no-intercept models, 82
 - h. Example (continued), 83
- 2. Deviations from means 83
- 3. Four methods of estimation 86
 - a. Ordinary least squares, 87
 - b. Generalized least squares, 87
 - c. Maximum likelihood, 87
 - d. The best linear unbiased estimator (b.l.u.e.), 88
- 4. Consequences of estimation 89
 - a. Unbiasedness, 89
 - b. Variances, 90
 - c. Estimating $E(y)$, 90
 - d. Residual error sum of squares, 92
 - e. Estimating the residual error variance, 93
 - f. Partitioning the total sum of squares, 93
 - g. Multiple correlation, 95
 - h. Example (continued), 96
- 5. Distributional properties 99
 - a. y is normal, 99
 - b. \hat{b} is normal, 99
 - c. \hat{b} and $\hat{\sigma}^2$ are independent, 99
 - d. SSE/σ^2 has a χ^2 -distribution, 99
 - e. Non-central χ^2 's, 100
 - f. F-distributions, 101
 - g. Analyses of variance, 101

3 (Continued)

h. Pure error, 103	
i. Tests of hypotheses, 104	
j. Example (continued), 105	
k. Confidence intervals, 107	
l. Example (continued), 109	
6. The general linear hypothesis	110
a. Testing linear hypotheses, 110	
b. Estimation under the null hypothesis, 112	
c. Four common hypotheses, 113	
(i) $H: \mathbf{b} = \mathbf{0}$, 114	
(ii) $H: \mathbf{b} = \mathbf{b}_0$, 114	
(iii) $H: \boldsymbol{\lambda}'\mathbf{b} = m$, 114	
(iv) $H: \mathbf{b}_q = \mathbf{0}$, 115	
d. Reduced models, 116	
(i) $\mathbf{K}'\mathbf{b} = \mathbf{m}$, 116	
(ii) $\mathbf{K}'\mathbf{b} = \mathbf{0}$, 118	
(iii) $\mathbf{b}_q = \mathbf{0}$, 120	
7. Related topics	124
a. The likelihood ratio test, 124	
b. Type I and II errors, 125	
c. The power of a test, 128	
d. Examining residuals, 129	
8. Summary of regression calculations	130
9. Exercises	132
4. Introducing Linear Models: Regression on Dummy Variables	135
1. Regression on allocated codes	135
a. Allocated codes, 135	
b. Difficulties and criticism, 136	
c. Grouped variables, 137	
d. Unbalanced data, 138	
2. Regression on dummy (0, 1) variables	140
a. Factors and levels, 140	
b. The regression, 141	
3. Describing linear models	145
a. A 1-way classification, 145	
b. A 2-way classification, 147	
c. A 3-way classification, 148	
d. Main effects and interactions, 149	
(i) Main effects, 149	
(ii) Interactions, 151	
e. Nested and crossed classifications, 155	
4. The normal equations	159
5. Exercises	162

5. Models Not of Full Rank	164
1. The normal equations	164
<i>a. The equations, 165</i>	
<i>b. Example, 165</i>	
<i>c. Solutions, 168</i>	
2. Consequences of a solution	169
<i>a. Expected values, 169</i>	
<i>b. Variances, 169</i>	
<i>c. Estimating $E(y)$, 170</i>	
<i>d. Residual error sum of squares, 170</i>	
<i>e. Estimating the residual error variance, 170</i>	
<i>f. Partitioning the total sum of squares, 171</i>	
<i>g. Coefficient of determination, 172</i>	
<i>h. Example (continued), 172</i>	
3. Distributional properties	174
<i>a. y is normal, 174</i>	
<i>b. \mathbf{b}^0 is normal, 174</i>	
<i>c. \mathbf{b}^0 and $\hat{\sigma}^2$ are independent, 174</i>	
<i>d. $SSE \sigma^2$ is χ^2, 174</i>	
<i>e. Non-central χ^2's, 175</i>	
<i>f. F-distributions, 176</i>	
<i>g. Analyses of variance, 176</i>	
<i>h. Tests of hypotheses, 178</i>	
<i>i. Example (continued), 180</i>	
4. Estimable functions	180
<i>a. Definition, 180</i>	
<i>b. Properties, 181</i>	
(i) The expected value of any observation is estimable, 181	
(ii) Linear combinations of estimable functions, 181	
(iii) The form of an estimable function, 181	
(iv) Invariance to the solution \mathbf{b}^0 , 181	
(v) The b.l.u.e., 181	
<i>c. Confidence intervals, 183</i>	
<i>d. Example (continued), 183</i>	
<i>e. What functions are estimable? 184</i>	
<i>f. Linearly independent estimable functions, 184</i>	
<i>g. Testing for estimability, 185</i>	
<i>h. General expressions, 186</i>	
<i>i. Example (continued), 187</i>	
5. The general linear hypothesis	188
<i>a. Testable hypotheses, 189</i>	
<i>b. Testing testable hypotheses, 190</i>	
<i>c. The hypothesis $\mathbf{K}'\mathbf{b} = \mathbf{0}$, 191</i>	
<i>d. Non-testable hypotheses, 193</i>	

5 (Continued)

- e. *Checking for testability*, 195
- f. *Example (continued)*, 196
- g. *Independent and orthogonal contrasts*, 199
- h. *Example (continued)*, 201

6. Restricted models	204
a. <i>Restrictions involving estimable functions</i> , 206	
b. <i>Restrictions involving non-estimable functions</i> , 208	
7. The "usual constraints"	209
a. <i>Limitations on constraints</i> , 212	
b. <i>Constraints of the form $b_i^0 = 0$</i> , 213	
c. <i>Procedure for deriving \mathbf{b}^0 and \mathbf{G}</i> , 215	
d. <i>Restrictions on the model</i> , 215	
e. <i>Example (continued)</i> , 217	
8. Generalizations	220
a. <i>Non-singular \mathbf{V}</i> , 220	
b. <i>Singular \mathbf{V}</i> , 221	
9. Summary	224
10. Exercises	224

6. Two Elementary Models 226

1. Summary of general results	227
2. The 1-way classification	229
a. <i>Model</i> , 229	
b. <i>Normal equations</i> , 232	
c. <i>Solving the normal equations</i> , 232	
d. <i>Analysis of variance</i> , 234	
e. <i>Estimable functions</i> , 235	
(i) μ is not estimable, 237	
(ii) α_i is not estimable, 237	
(iii) $(\sum \lambda_i)\mu + \sum \lambda_i \alpha_i$ is estimable for any λ_i , 237	
(iv) $\sum \lambda_i \alpha_i$ for $\sum \lambda_i = 0$ is estimable, 238	
(v) $\alpha_i - \alpha_k$ for $i \neq k$ is estimable, 238	
f. <i>Tests of linear hypotheses</i> , 238	
(i) <i>General hypotheses</i> , 238	
(ii) <i>The test based on $F(M)$</i> , 239	
(iii) <i>The test based on $F(R_m)$</i> , 240	
g. <i>Independent and orthogonal contrasts</i> , 242	
h. <i>Models that include restrictions</i> , 243	
i. <i>Balanced data</i> , 245	
3. Reductions in sums of squares	246
a. <i>The $R(\)$ notation</i> , 246	
b. <i>Analyses of variance</i> , 247	
c. <i>Tests of hypotheses</i> , 248	

6. Two Elementary Models (Continued)	
4. The 2-way nested classification	249
a. Model, 250	
b. Normal equations, 250	
c. Solving the normal equations, 251	
d. Analysis of variance, 252	
e. Estimable functions, 254	
f. Tests of hypotheses, 255	
g. Models that include restrictions, 256	
h. Balanced data, 257	
5. Normal equations for design models	257
6. Exercises	259
7. The 2-Way Crossed Classification	261
1. The 2-way classification without interaction	261
a. Model, 261	
b. Normal equations, 264	
c. Solving the normal equations, 264	
d. Absorbing equations, 266	
e. Analyses of variance, 270	
(i) Basic calculations, 270	
(ii) Fitting the model, 270	
(iii) Fitting rows before columns, 271	
(iv) Fitting columns before rows, 273	
(v) Ignoring and/or adjusting for effects, 274	
(vi) Interpretation of results, 276	
f. Estimable functions, 279	
g. Tests of hypotheses, 280	
h. Models that include restrictions, 283	
i. Balanced data, 284	
2. The 2-way classification with interaction	286
a. Model, 286	
b. Normal equations, 288	
c. Solving the normal equations, 291	
d. Analysis of variance, 292	
(i) Basic calculations, 292	
(ii) Fitting different models, 295	
(iii) Computational alternatives, 297	
(iv) Interpretation of results, 300	
(v) Fitting main effects before interactions, 300	
e. Estimable functions, 301	
f. Tests of hypotheses, 305	
(i) The general hypothesis, 305	
(ii) The hypothesis for $F(M)$, 306	
(iii) Hypotheses for $F(\alpha \mu)$ and $F(\beta \mu)$, 307	
(iv) Hypotheses for $F(\alpha \mu, \beta)$ and $F(\beta \mu, \alpha)$, 308	

7 (Continued)

- (v) Hypotheses for $F(\gamma \mid \mu, \alpha, \beta)$, 311
- (vi) Reduction to the no-interaction model, 312
- (vii) Independence properties, 313
- g. Models that include restrictions, 313*
- h. All cells filled, 314*
- i. Balanced data, 315*
- 3. Interpretation of hypotheses 316
- 4. Connectedness 318
- 5. μ_{ij} -models 324
- 6. Exercises 327

8. Some Other Analyses 332

- 1. Large-scale survey-type data 332
 - a. Example, 332*
 - b. Fitting a linear model, 333*
 - c. Main-effects-only models, 335*
 - d. Stepwise fitting, 337*
 - e. Connectedness, 337*
 - f. μ_{ij} -models, 338*
- 2. Covariance 340
 - a. A general formulation, 341*
 - (i) The model, 341
 - (ii) Solving the normal equations, 341
 - (iii) Estimability, 342
 - (iv) A model for handling the covariates, 342
 - (v) Analyses of variance, 343
 - (vi) Tests of hypotheses, 345
 - (vii) Summary, 347
 - b. The 1-way classification, 348*
 - (i) A single regression, 348
 - (ii) Example, 353
 - (iii) The intra-class regression model, 355
 - (iv) Example (continued), 358
 - (v) Another example, 359
 - c. The 2-way classification (with interaction), 360*
- 3. Data having all cells filled 361
 - a. Estimating missing observations, 362*
 - b. Setting data aside, 364*
 - c. Analyses of means, 365*
 - (i) Unweighted means analysis, 365
 - (ii) Example, 367
 - (iii) Weighted squares of means, 369
 - (iv) Example (continued), 371
 - d. Separate analyses, 372*
- 4. Exercises 373

9. Introduction to Variance Components	376
1. Fixed and random models	376
<i>a. A fixed effects model, 377</i>	
<i>b. A random effects model, 377</i>	
<i>c. Other examples, 379</i>	
(i) Of treatments and varieties, 379	
(ii) Of mice and men, 379	
(iii) Of cows and bulls, 380	
2. Mixed models	380
(i) Of mice and diets, 381	
(ii) Of treatments and crosses, 381	
(iii) On measuring shell velocities, 381	
3. Fixed or random?	382
4. Finite populations	383
5. Introduction to estimation	384
<i>a. Variance matrix structures, 384</i>	
<i>b. Analyses of variance, 385</i>	
<i>c. Estimation, 388</i>	
6. Rules for balanced data	389
<i>a. Establishing analysis of variance tables, 389</i>	
(i) Factors and levels, 389	
(ii) Lines in the analysis of variance table, 390	
(iii) Interactions, 390	
(iv) Degrees of freedom, 391	
(v) Sums of squares, 391	
<i>b. Calculating sums of squares, 392</i>	
<i>c. Expected values of mean squares, $E(MS)$, 393</i>	
(i) Completely random models, 393	
(ii) Fixed effects and mixed models, 393	
7. The 2-way classification	394
<i>a. The fixed effects model, 397</i>	
<i>b. The random effects model, 400</i>	
<i>c. The mixed model, 400</i>	
8. Estimating variance components from balanced data	404
<i>a. Unbiasedness and minimum variance, 405</i>	
<i>b. Negative estimates, 406</i>	
9. Normality assumptions	408
<i>a. Distribution of mean squares, 409</i>	
<i>b. Distribution of estimators, 410</i>	
<i>c. Tests of hypotheses, 411</i>	
<i>d. Confidence intervals, 413</i>	
<i>e. Probability of negative estimates, 415</i>	
<i>f. Sampling variances of estimators, 415</i>	
(i) Derivation, 416	
(ii) Covariance matrix, 416	
(iii) Unbiased estimation, 417	
<i>g. Maximum likelihood estimation, 418</i>	
10. Exercises	419

10. Methods of Estimating Variance Components from Unbalanced Data	421
1. Expectations of quadratic forms	421
<i>a. Fixed effects models, 422</i>	
<i>b. Mixed models, 423</i>	
<i>c. Random effects models, 424</i>	
<i>d. Applications, 424</i>	
2. Analysis of variance method (Henderson's method 1)	424
<i>a. Model and notation, 425</i>	
<i>b. Analogous sums of squares, 425</i>	
(i) Empty cells, 426	
(ii) Balanced data, 426	
(iii) A negative "sum of squares," 426	
(iv) Uncorrected sums of squares, 426	
<i>c. Expectations, 427</i>	
(i) An example, 427	
(ii) Mixed models, 429	
(iii) General results, 431	
(iv) Calculation by "synthesis", 432	
<i>d. Sampling variances of estimators, 433</i>	
(i) Derivation, 434	
(ii) Estimation, 436	
(iii) Calculation by "synthesis", 439	
3. Adjusting for bias in mixed models	441
<i>a. General method, 441</i>	
<i>b. A simplification, 442</i>	
<i>c. A special case: Henderson's method 2, 442</i>	
4. Fitting constants method (Henderson's method 3)	443
<i>a. General properties, 443</i>	
<i>b. The 2-way classification, 446</i>	
(i) Expected values, 446	
(ii) Estimation, 447	
(iii) Calculation, 447	
<i>c. Too many equations, 448</i>	
<i>d. Mixed models, 450</i>	
<i>e. Sampling variances of estimators, 451</i>	
5. Analysis of means methods	451
6. Symmetric sums methods	452
7. Infinitely many quadratics	455
8. Maximum likelihood for mixed models	458
<i>a. Estimating fixed effects, 459</i>	
<i>b. Fixed effects and variance components, 462</i>	
<i>c. Large sample variances, 464</i>	
9. Mixed models having one random factor	465
10. Best quadratic unbiased estimation	470
11. Exercises	472

11. Variance Component Estimation from Unbalanced Data: Formulae	473
1. The 1-way classification	473
a. Model, 473	
b. Analysis of variance estimators, 474	
c. Variances of analysis of variance estimators, 474	
d. Variances of large sample Maximum likelihood estimators, 474	
e. Symmetric sums estimators, 474	
f. BQUE's for the model with $\mu = 0$, 474	
g. BQUE's for the model with $\mu \neq 0$, 475	
2. The 2-way nested classification	475
a. Model, 475	
b. Analysis of variance estimators, 475	
c. Variances of analysis of variance estimators, 476	
d. Variances of large sample maximum likelihood estimators, 477	
e. Symmetric sums estimators, 477	
3. The 3-way nested classification	477
a. Model, 477	
b. Analysis of variance estimators, 478	
c. Variances of analysis of variance estimators, 478	
4. The 2-way classification with interaction, random model	480
a. Model, 480	
b. Analysis of variance estimators, 480	
c. Variances of analysis of variance estimators, 481	
d. Symmetric sums estimators, 483	
e. Fitting constants method estimators, 483	
f. Analysis of means estimators, 484	
5. The 2-way classification with interaction, mixed model	486
a. Model, 486	
b. Fitting constants method estimators, 486	
c. Fixed effects estimators, 486	
6. The 2-way classification without interaction, random model	487
a. Model, 487	
b. Analysis of variance estimators, 487	
c. Variances of analysis of variance estimators, 487	
d. Symmetric sums estimators, 488	
e. Fitting constants method estimators, 488	
f. Variances of fitting constants method estimators, 489	
7. Mixed models with one random factor	489
a. Model, 489	
b. Fitting constants method estimators, 489	
c. An iterative procedure, 490	
8. The 2-way classification without interaction, mixed model	490
a. Model, 490	
b. Fitting constants method estimators, 490	
c. An iterative procedure, 490	
d. Fixed effects estimators, 491	

11 (*Continued*)

9. The 3-way classification, random model 491
- a. Model, 491*
 - b. Analysis of variance estimators, 491*
 - c. Variances of analysis of variance estimators, 493*

Literature Cited 515

Statistical Tables 523

Index 529