Contents

NOTATION xix

1 PRELIMINARIES 1

1.1 Notation 1
1.2 What Is Multivariate Analysis? 3
1.3 Expectation and Covariance Operators 5
1.4 Sample Data 8
1.5 Mahalanobis Distances and Angles 10
1.6 Simultaneous Inference 11
1.6.1 *Simultaneous tests*, 11
1.6.2 *Union–intersection principle*, 13
1.7 Likelihood Ratio Tests 14

Exercises 1, 14

2 MULTIVARIATE DISTRIBUTIONS 17

2.1 Introduction 17
2.2 Multivariate Normal Distribution 17
2.3 Wishart Distribution 20
2.3.1 *Definition and properties*, 20
2.3.2 *Generalized quadratics*, 22
2.3.3 *Noncentral Wishart distribution*, 26
2.3.4 *Eigenvalues of a Wishart matrix*, 27
2.3.5 *Determinant of a Wishart matrix*, 27
2.4 Hotelling's T^2 Distribution 28
2.4.1 *Central distribution*, 28
2.4.2 *Noncentral distribution*, 32
2.5. Multivariate Beta Distributions 32
2.5.1 *Derivation*, 32
2.5.2 *Multivariate beta eigenvalues*, 35
2.5.3 *Two trace statistics*, 38
a. Lawley–Hotelling statistic, 38
b. Pillai's trace statistic, 39

2.5.4 U-*Distribution*, 40
2.5.5 *Summary of special distributions*, 42
a. Hotelling's T^2, 42
b. *U*-statistic, 43
c. Maximum root statistic, 43
d. Trace statistics, 43
e. Equivalence of statistics when $m_H = 1$, 43
2.5.6 *Factorizations of* U, 45
a. Product of beta variables, 45
b. Product of two *U*-statistics, 48
2.6 Rao's Distribution 50
2.7 Multivariate Skewness and Kurtosis 54

Exercises 2, 55

3 Inference for the Multivariate Normal 59

3.1 Introduction 59
3.2 Estimation 59
3.2.1 *Maximum likelihood estimation*, 59
3.2.2 *Distribution theory*, 63
3.3 Testing for the Mean 63
3.3.1 *Hotelling's T^2 test*, 63
3.3.2 *Power of the test*, 69
3.3.3 *Robustness of the test*, 69
3.3.4 *Step-down test procedure*, 70
3.4 Linear Constraints on the Mean 71
3.4.1 *Generalization of the paired comparison test*, 71
3.4.2 *Some examples*, 72
a. Repeated-measurement designs, 72
b. Testing specified contrasts, 76
c. Test for symmetry, 76
d. Testing for a polynomial growth trend, 77
e. Independent estimate of Σ, 77
3.4.3 *Minimization technique for the test statistic*, 77
3.4.4 *Confidence intervals*, 81
3.4.5 *Functional relationships (errors in variables regression)*, 85
3.5 Inference for the Dispersion Matrix 86
3.5.1 *Introduction*, 86
3.5.2 *Blockwise independence: Two blocks*, 87
a. Likelihood ratio test, 88
b. Maximum root test, 89
3.5.3 *Blockwise independence:* b *Blocks*, 90

3.5.4 *Diagonal dispersion matrix*, 92
3.5.5 *Equal diagonal blocks*, 94
3.5.6 *Equal correlations and equal variances*, 95
3.5.7 *Simultaneous confidence intervals for correlations*, 96
3.5.8 *Large-sample inferences*, 98
3.5.9 *More general covariance structures*, 102
3.6 Comparing Two Normal Populations 102
3.6.1 *Tests for equal dispersion matrices*, 102
a. Likelihood ratio test, 103
b. Union–intersection test, 105
c. Robustness of tests, 106
d. Robust large-sample tests, 107
3.6.2 *Test for equal means assuming equal dispersion matrices*, 108
a. Hotelling's T^2 test, 108
b. Effect of unequal dispersion matrices, 111
c. Effect of nonnormality, 112
3.6.3 *Test for equal means assuming unequal dispersion matrices*, 114
3.6.4 *Profile analysis: Two populations*, 117

Exercises 3, 124

4 Graphical and Data - Oriented Techniques 127

4.1 Multivariate Graphical Displays 127
4.2 Transforming to Normality 138
4.2.1 *Univariate transformations*, 138
4.2.2 *Multivariate transformations*, 140
4.3 Distributional Tests and Plots 141
4.3.1 *Investigating marginal distributions*, 141
4.3.2 *Tests and plots for multivariate normality*, 148
4.4 Robust Estimation 156
4.4.1 *Why robust estimates?*, 156
4.4.2 *Estimation of location*, 156
a Univariate methods, 156
b. Multivariate methods, 162
4.4.3 *Estimation of dispersion and covariance*, 162
a. Univariate methods, 162
b. Multivariate methods, 165
4.5 Outlying Observations 169

Exercises 4, 173

5 DIMENSION REDUCTION AND ORDINATION 175

5.1 Introduction 175
5.2 Principal Components 176
5.2.1 *Definition*, 176
5.2.2 *Dimension reduction properties*, 176
5.2.3 *Further properties*, 181
5.2.4 *Sample principal components*, 184
5.2.5 *Inference for sample components*, 197
5.2.6 *Preliminary selection of variables*, 200
5.2.7 *General applications*, 200
5.2.8 *Generalized principal components analysis*, 203
5.3 Biplots and *h*-Plots 204
5.4 Factor Analysis 213
5.4.1 *Underlying model*, 213
5.4.2 *Estimation procedures*, 216
a. Method of maximum likelihood, 216
b. Principal factor analysis, 219
c. Estimating factor scores, 220
5.4.3 *Some difficulties*, 222
a. Principal factor analysis, 225
b. Maximum likelihood factor analysis, 230
c. Conclusions, 231
5.5 Multidimensional Scaling 235
5.5.1 *Classical (metric) solution*, 235
5.5.2 *Nonmetric scaling*, 241
5.6 Procrustes Analysis (Matching Configurations) 253
5.7 Canonical Correlations and Variates 256
5.7.1 *Population correlations*, 256
5.7.2 *Sample canonical correlations*, 260
5.7.3 *Inference*, 264
5.7.4 *More than two sets of variables*, 267
5.8 Discriminant Coordinates 269
5.9 Assessing Two-Dimensional Representations 273
5.10 A Brief Comparison of Methods 274

Exercises 5, 275

6 DISCRIMINANT ANALYSIS 279

6.1 Introduction 279
6.2 Two Groups: Known Distributions 280
6.2.1 *Misclassification errors*, 280
6.2.2 *Some allocation principles*, 281
a. Minimize total probability of misclassification, 281

b. Likelihood ratio method, 285
c. Minimize total cost of misclassification, 285
d. Maximize the posterior probability, 285
e. Minimax allocation, 286
f. Summary, 287
6.3 Two Groups: Known Distributions With Unknown Parameters 287
6.3.1 *General methods*, 287
6.3.2 *Normal populations*, 293
a. Linear discriminant function, 293
b. Quadratic discriminant function, 297
c. Robustness of LDF and QDF, 297
d. Missing values, 300
e. Predictive discriminant, 301
6.3.3 *Multivariate discrete distributions*, 303
a. Independent binary variables, 303
b. Correlated binary variables, 304
c. Correlated discrete variables, 306
6.3.4 *Multivariate discrete–continuous distributions*, 306
6.4 Two Groups: Logistic Discriminant 308
6.4.1 *General model*, 308
6.4.2 *Sampling designs*, 309
a. Conditional sampling, 309
b. Mixture sampling, 310
c. Separate sampling, 310
6.4.3 *Computations*, 312
6.4.4 *LGD versus LDF*, 317
6.4.5 *Predictive logistic model*, 318
6.4.6 *Quadratic discrimination*, 319
6.5 Two Groups: Unknown Distributions 320
6.5.1 *Kernel method*, 320
a. Continuous data, 320
b. Binary data, 322
c. Continuous and discrete data, 323
6.5.2 *Other nonparametric methods*, 323
a Nearest neighbor techniques, 323
b. Partitioning methods, 324
c. Distance methods, 324
d. Rank procedures, 324
e. Sequential discrimination on the variables, 326
6.6 Utilizing Unclassified Observations 327
6.7 All Observations Unclassified (Method of Mixtures) 328
6.8 Cases of Doubt 329
6.9 More Than Two Groups 330
6.10 Selection of Variables 337

6.10.1 *Two groups*, 337
6.10.2 *More than two groups*, 341
6.11 Some Conclusions 342

Exercises 6, 343

7 CLUSTER ANALYSIS 347

7.1 Introduction 347
7.2 Proximity data 351
7.2.1 *Dissimilarities*, 351
7.2.2 *Similarities*, 356
7.3 Hierarchical Clustering: Agglomerative Techniques 359
7.3.1 *Some commonly used methods*, 360
a. Single linkage (nearest neighbor) method, 360
b. Complete linkage (farthest neighbor) method, 361
c. Centroid method, 362
d. Incremental sum of squares method, 363
e. Median method, 363
f. Group average method, 363
g. Lance and Williams flexible method, 364
h. Information measures, 365
i. Other methods, 366
7.3.2 *Comparison of methods*, 368
a. Monotonicity, 368
b. Spatial properties, 373
c. Computational effort, 375
7.4 Hierarchical Clustering: Divisive Techniques 376
7.4.1 *Monothetic methods*, 377
7.4.2 *Polythetic methods*, 378
7.5. Partitioning Methods 379
7.5.1 *Number of partitions*, 379
7.5.2 *Starting the process*, 379
7.5.3 *Reassigning the objects*, 380
a. Minimize trace W, 382
b. Minimize $|W|$, 382
c. Maximize trace BW^{-1}, 383
7.5.4 *Other techniques*, 386
7.6 Overlapping Clusters (Clumping) 387
7.7 Choosing the Number of Clusters 388
7.8 General Comments 390

Exercises 7, 392

8 MULTIVARIATE LINEAR MODELS 395

8.1 Least Squares Estimation 395
8.2 Properties of Least Squares Estimates 399

8.3 Least Squares With Linear Constraints 403
8.4 Distribution Theory 405
8.5 Analysis of Residuals 408
8.6 Hypothesis Testing 409
8.6.1 *Extension of univariate theory*, 409
8.6.2 *Test procedures*, 411
a. Union–intersection test, 411
b. Likelihood ratio test, 412
c. Other test statistics, 413
d. Comparison of four test statistics, 414
e. Mardia's permutation test, 416
8.6.3 *Simultaneous confidence intervals*, 417
8.6.4 *Comparing two populations*, 419
8.6.5 *Canonical form*, 421
8.6.6 *Missing observations*, 422
8.6.7 *Other topics*, 423
8.7 A Generalized Linear Hypothesis 423
8.7.1 *Theory*, 423
8.7.2 *Profile analysis for* K *populations*, 424
8.7.3 *Tests for mean of multivariate normal*, 425
8.8 Step-Down Procedures 426
8.9 Multiple Design Models 428

Exercises 8, 431

9 Multivariate Analysis of Variance and Covariance 433

9.1 Introduction 433
9.2 One-Way Classification 433
9.2.1 *Hypothesis testing*, 433
9.2.2 *Multiple comparisons*, 437
9.2.3 *Comparison of test statistics*, 440
9.2.4 *Robust tests for equal means*, 443
a. Permutation test, 443
b. James' test, 445
9.2.5 *Reparameterization*, 447
9.2.6 *Comparing dispersion matrices*, 448
a. Test for equal dispersion matrices, 448
b. Graphical comparisons, 451
9.2.7 *Exact procedures for means assuming unequal dispersion matrices*, 452
a. Hypothesis testing, 452
b. Multiple confidence intervals, 454
9.3 Randomized Block Design 454
9.3.1 *Hypothesis testing and confidence intervals*, 454
9.3.2 *Underlying assumptions*, 458

9.4 Two-Way Classification With Equal Observations per Mean . . . 458
9.5 Analysis of Covariance 463
9.5.1 *Univariate theory*, 463
9.5.2 *Multivariate theory*, 465
9.5.3 *Test for additional information*, 471
9.6 Multivariate Cochran's Theorem on Quadratics 472
9.7 Growth Curve Analysis 474
9.7.1 *Examples*, 474
a. Single growth curve, 474
b. Two growth curves, 475
c. Single growth curve for a randomized block design, 476
d. Two-dimensional growth curve, 477
9.7.2 *General theory*, 478
a. Potthoff and Roy's method, 479
b. Rao–Khatri analysis of covariance method, 480
c. Choice of method, 483
9.7.3 *Single growth curve*, 484
9.7.4 *Test for internal adequacy*, 486
9.7.5 *A case study*, 487
9.7.6 *Further topics*, 492

Exercises 9, 494

10 Special Topics 496

10.1 Computational Techniques 496
10.1.1 *Solving the normal equations*, 496
a. Cholesky decomposition, 496
b. *QR*-algorithm, 497
10.1.2 *Hypothesis matrix*, 498
10.1.3 *Calculating Hotelling's* T^2, 499
10.1.4 *Generalized symmetric eigenproblem*, 500
a. Multivariate linear hypothesis, 501
b. One-way classification, 503
c. Discriminant coordinates, 504
10.1.5 *Singular value decomposition*, 504
a. Definition, 504
b. Solution of normal equations, 505
c. Principal components, 506
d. Canonical correlations, 506
10.1.6 *Selecting the best subset*, 507
a. Response variables, 507
b. Regressor variables, 510

10.2 Log-Linear Models for Binary Data 512
10.3 Incomplete Data 514

APPENDIX 517

A Some Matrix Algebra 517
A1 *Trace and eigenvalues*, 517
A2 *Rank*, 518
A3 *Patterned matrices*, 519
A4 *Positive semidefinite matrices*, 521
A5 *Positive definite matrices*, 521
A6 *Idempotent matrices*, 522
A7 *Optimization and inequalities*, 523
A8 *Vector and matrix differentiation*, 530
A9 *Jacobians and transformations*, 531
A10 *Asymptotic normality*, 532
B Orthogonal Projections 533
B1 *Orthogonal decomposition of vectors*, 533
B2 *Orthogonal complements*, 534
B3 *Projections on subspaces*, 535
C Order Statistics and Probability Plotting 539
C1 *Sample distribution functions*, 539
C2 *Gamma distribution*, 540
C3 *Beta distribution*, 541
C4 *Probability plotting*, 542
D Statistical Tables 545
D1 *Bonferroni* t*-percentage points*, 546
D2 *Maximum likelihood estimates for the gamma distribution*, 550
D3 *Upper tail percentage points for* $\sqrt{b_1}$, 551
D4 *Coefficients in a normalizing transformation of* $\sqrt{b_1}$, 551
D5 *Simulation percentiles for* b_2, 553
D6 *Charts for the percentiles of* b_2, 554
D7 *Coefficients for the Wilk–Shapiro* (W) *test*, 556
D8 *Percentiles for the Wilk–Shapiro* (W) *test*, 558
D9 *D'Agostino's test for normality*, 558
D10 *Anderson–Darling* (A_n^2) *test for normality*, 560
D11 *Discordancy test for single gamma outlier*, 561
D12 *Discordancy test for single multivariate normal outlier*, 562
D13 *Wilks' likelihood ratio test*, 562
D14 *Roy's maximum root statistic*, 563
D15 *Lawley–Hotelling trace statistic*, 563
D16 *Pillai's trace statistic*, 564

D17 *Test for mutual independence*, 564
D18 *Test for equal dispersion matrices with equal sample sizes*, 564

OUTLINE SOLUTIONS TO EXERCISES 615

REFERENCES 626

INDEX 671