# Contents

Preface .................................................................................. xv
Preface to the first revised printing ............................... xvii
Preface to the second revised printing ............................. xviii

## Part One—Matrices

1 Basic properties of vectors and matrices .......................... 3

   1 Introduction, 3
   2 Sets, 3
   3 Matrices: addition and multiplication, 4
   4 The transpose of a matrix, 5
   5 Square matrices, 6
   6 Linear forms and quadratic forms, 7
   7 The rank of a matrix, 8
   8 The inverse, 9
   9 The determinant, 9
  10 The trace, 10
  11 Partitioned matrices, 11
  12 Complex matrices, 13
  13 Eigenvalues and eigenvectors, 13
  14 Schur's decomposition theorem, 16
  15 The Jordan decomposition, 17
  16 The singular-value decomposition, 18
  17 Further results concerning eigenvalues, 19
  18 Positive (semi)definite matrices, 21
  19 Three further results for positive definite matrices, 23
  20 A useful result, 24
  Miscellaneous exercises, 25
  Bibliographical notes, 26

2 Kronecker products, the vec operator and the Moore–Penrose inverse ................................................. 27

   1 Introduction, 27
   2 The Kronecker product, 27
3 Eigenvalues of a Kronecker product, 28
4 The vec operator, 30
5 The Moore-Penrose (MP) inverse, 32
6 Existence and uniqueness of the MP inverse, 32
7 Some properties of the MP inverse, 33
8 Further properties, 34
9 The solution of linear equation systems, 36
Miscellaneous exercises, 38
Bibliographical notes, 39

3 Miscellaneous matrix results 40

1 Introduction, 40
2 The adjoint matrix, 40
3 Proof of Theorem 1, 41
4 Two results concerning bordered determinants, 43
5 The matrix equation AX = 0, 44
6 The Hadamard product, 45
7 The commutation matrix $K_{mn}$, 46
8 The duplication matrix $D_n$, 48
9 Relationship between $D_{n+1}$ and $D_n$, I, 50
10 Relationship between $D_{n+1}$ and $D_n$, II, 52
11 Conditions for a quadratic form to be positive (negative) subject to linear constraints, 53
12 Necessary and sufficient conditions for $r(A:B) = r(A) + r(B)$, 56
13 The bordered Gramian matrix, 57
14 The equations $X_1A + X_2B' = G_1$, $X_1B = G_2$, 60
Miscellaneous exercises, 62
Bibliographical notes, 62

Part Two—Differentials: the theory

4 Mathematical preliminaries 65

1 Introduction, 65
2 Interior points and accumulation points, 65
3 Open and closed sets, 66
4 The Bolzano-Weierstrass theorem, 69
5 Functions, 70
6 The limit of a function, 70
7 Continuous functions and compactness, 71
8 Convex sets, 72
9 Convex and concave functions, 75
Bibliographical notes, 77
Contents

5 Differentials and differentiability 78

1 Introduction, 78
2 Continuity, 78
3 Differentiability and linear approximation, 80
4 The differential of a vector function, 82
5 Uniqueness of the differential, 84
6 Continuity of differentiable functions, 84
7 Partial derivatives, 85
8 The first identification theorem, 87
9 Existence of the differential, I, 88
10 Existence of the differential, II, 89
11 Continuous differentiability, 91
12 The chain rule, 91
13 Cauchy invariance, 93
14 The mean-value theorem for real-valued functions, 93
15 Matrix functions, 94
16 Some remarks on notation, 96
Miscellaneous exercises, 98
Bibliographical notes, 98

6 The second differential 99

1 Introduction, 99
2 Second-order partial derivatives, 99
3 The Hessian matrix, 100
4 Twice differentiability and second-order approximation, I, 101
5 Definition of twice differentiability, 102
6 The second differential, 103
7 (Column) symmetry of the Hessian matrix, 105
8 The second identification theorem, 107
9 Twice differentiability and second-order approximation, II, 108
10 Chain rule for Hessian matrices, 110
11 The analogue for second differentials, 111
12 Taylor's theorem for real-valued functions, 112
13 Higher-order differentials, 113
14 Matrix functions, 114
Bibliographical notes, 115

7 Static optimization 116

1 Introduction, 116
2 Unconstrained optimization, 116
3 The existence of absolute extrema, 118
4 Necessary conditions for a local minimum, 119
Part Three—Differentials: the practice

8 Some important differentials

1 Introduction, 147
2 Fundamental rules of differential calculus, 147
3 The differential of a determinant, 149
4 The differential of an inverse, 151
5 The differential of the Moore–Penrose inverse, 152
6 The differential of the adjoint matrix, 155
7 On differentiating eigenvalues and eigenvectors, 157
8 The differential of eigenvalues and eigenvectors: the real symmetric case, 158
9 The differential of eigenvalues and eigenvectors: the general complex case, 161
10 Two alternative expressions for \( d\lambda \), 163
11 The second differential of the eigenvalue function, 166
12 Multiple eigenvalues, 167
Miscellaneous exercises, 167
Bibliographical notes, 169

9 First-order differentials and Jacobian matrices

1 Introduction, 170
2 Classification, 170
3 Bad notation, 171
4 Good notation, 173
5 Identification of Jacobian matrices, 174
6 The first identification table, 175
Contents

7 Partitioning of the derivative, 175
8 Scalar functions of a vector, 176
9 Scalar functions of a matrix, I: trace, 177
10 Scalar functions of a matrix, II: determinant, 178
11 Scalar functions of a matrix, III: eigenvalue, 180
12 Two examples of vector functions, 181
13 Matrix functions, 182
14 Kronecker products, 184
15 Some other problems, 185
Bibliographical notes, 187

10 Second-order differentials and Hessian matrices 188

1 Introduction, 188
2 The Hessian matrix of a matrix function, 188
3 Identification of Hessian matrices, 189
4 The second identification table, 190
5 An explicit formula for the Hessian matrix, 191
6 Scalar functions, 192
7 Vector functions, 194
8 Matrix functions, I, 194
9 Matrix functions, II, 195

Part Four—Inequalities

11 Inequalities 199

1 Introduction, 199
2 The Cauchy–Schwarz inequality, 199
3 Matrix analogues of the Cauchy–Schwarz inequality, 201
4 The theorem of the arithmetic and geometric means, 202
5 The Rayleigh quotient, 203
6 Concavity of $\lambda_1$, convexity of $\lambda_n$, 204
7 Variational description of eigenvalues, 205
8 Fischer's min–max theorem, 206
9 Monotonicity of the eigenvalues, 208
10 The Poincaré separation theorem, 209
11 Two corollaries of Poincaré's theorem, 210
12 Further consequences of the Poincaré theorem, 211
13 Multiplicative version, 212
14 The maximum of a bilinear form, 213
15 Hadamard's inequality, 214
16 An interlude: Karamata's inequality, 215
17 Karamata's inequality applied to eigenvalues, 217
18 An inequality concerning positive semidefinite matrices, 217
19 A representation theorem for \((Σa_i^p)^{1/p}\), 218
20 A representation theorem for \((\text{tr } A^p)^{1/p}\), 219
21 Hölder’s inequality, 220
22 Concavity of \(\log |A|\), 222
23 Minkowski’s inequality, 223
24 Quasilinear representation of \(|A|^{1/n}\), 224
25 Minkowski’s determinant theorem, 227
26 Weighted means of order \(p\), 227
27 Schlömilch’s inequality, 229
28 Curvature properties of \(M_p(x, a)\), 230
29 Least squares, 232
30 Generalized least squares, 233
31 Restricted least squares, 233
32 Restricted least squares: matrix version, 235
Miscellaneous exercises, 236
Bibliographical notes, 240

Part Five—The linear model

12 Statistical preliminaries 243

1 Introduction, 243
2 The cumulative distribution function, 243
3 The joint density function, 244
4 Expectations, 244
5 Variance and covariance, 245
6 Independence of two random variables (vectors), 247
7 Independence of \(n\) random variables (vectors), 249
8 Sampling, 249
9 The one-dimensional normal distribution, 249
10 The multivariate normal distribution, 250
11 Estimation, 252
Miscellaneous exercises, 253
Bibliographical notes, 253

13 The linear regression model 254

1 Introduction, 254
2 Affine minimum-trace unbiased estimation, 255
3 The Gauss–Markov theorem, 256
4 The method of least squares, 258
5 Aitken’s theorem, 259
6 Multicollinearity, 261
7 Estimable functions, 263
Contents

8 Linear constraints: the case $\mathcal{M}(R') \subset \mathcal{M}(X')$, 264
9 Linear constraints: the general case, 267
10 Linear constraints: the case $\mathcal{M}(R') \cap \mathcal{M}(X') = \{0\}$, 270
11 A singular variance matrix: the case $\mathcal{M}(X) \subset \mathcal{M}(V)$, 271
12 A singular variance matrix: the case $r(X'V+X) = r(X)$, 273
13 A singular variance matrix: the general case, I, 274
14 Explicit and implicit linear constraints, 275
15 The general linear model, I, 277
16 A singular variance matrix: the general case, II, 278
17 The general linear model, II, 281
18 Generalized least squares, 282
19 Restricted least squares, 283
Miscellaneous exercises, 285
Bibliographical notes, 286

14 Further topics in the linear model 287

1 Introduction, 287
2 Best quadratic unbiased estimation of $\sigma^2$, 287
3 The best quadratic and positive unbiased estimator of $\sigma^2$, 288
4 The best quadratic unbiased estimator of $\sigma^2$, 290
5 Best quadratic invariant estimation of $\sigma^2$, 292
6 The best quadratic and positive invariant estimator of $\sigma^2$, 293
7 The best quadratic invariant estimator of $\sigma^2$, 294
8 Best quadratic unbiased estimation in the multivariate normal case, 295
9 Bounds for the bias of the least squares estimator of $\sigma^2$, I, 297
10 Bounds for the bias of the least squares estimator of $\sigma^2$, II, 299
11 The prediction of disturbances, 300
12 Predictors that are best linear unbiased with scalar variance matrix (BLUS), 301
13 Predictors that are best linear unbiased with fixed variance matrix (BLUF), I, 303
14 Predictors that are best linear unbiased with fixed variance matrix (BLUF), II, 305
15 Local sensitivity of the posterior mean, 306
16 Local sensitivity of the posterior precision, 308
Bibliographical notes, 309

Part Six—Applications to maximum likelihood estimation

15 Maximum likelihood estimation 313

1 Introduction, 313
2 The method of maximum likelihood (ML), 313
3 ML estimation of the multivariate normal distribution, 314
4 Implicit versus explicit treatment of symmetry, 316
5 The treatment of positive definiteness, 317
6 The information matrix, 317
7 ML estimation of the multivariate normal distribution with distinct means, 319
8 The multivariate linear regression model, 320
9 The errors-in-variables model, 322
10 The nonlinear regression model with normal errors, 324
11 A special case: functional independence of mean parameters and variance parameters, 326
12 Generalization of Theorem 6, 327
Miscellaneous exercises, 329
Bibliographical notes, 330

16 Simultaneous equations
1 Introduction, 331
2 The simultaneous equations model, 331
3 The identification problem, 333
4 Identification with linear constraints on B and Γ only, 334
5 Identification with linear constraints on B, Γ and Σ, 335
6 Nonlinear constraints, 337
7 Full-information maximum likelihood (FIML): the information matrix (general case), 337
8 Full-information maximum likelihood (FIML): the asymptotic variance matrix (special case), 339
9 Limited-information maximum likelihood (LIML): the first-order conditions, 342
10 Limited-information maximum likelihood (LIML): the information matrix, 344
11 Limited-information maximum likelihood (LIML): the asymptotic variance matrix, 346
Bibliographical notes, 351

17 Topics in psychometrics
1 Introduction, 352
2 Population principal components, 353
3 Optimality of principal components, 353
4 A related result, 355
5 Sample principal components, 356
6 Optimality of sample principal components, 358
7 Sample analogue of Theorem 3, 358
8 One-mode component analysis, 358