

Contents

Part I Descriptive Techniques

1	Comparison of Batches	3
1.1	Boxplots	4
1.2	Histograms	11
1.3	Kernel Densities	14
1.4	Scatterplots	18
1.5	Chernoff-Flury Faces	21
1.6	Andrews' Curves	25
1.7	Parallel Coordinate Plots	28
1.8	Hexagon Plots	33
1.9	Boston Housing	35
1.10	Exercises	43

Part II Multivariate Random Variables

2	A Short Excursion into Matrix Algebra	49
2.1	Elementary Operations	49
2.2	Spectral Decompositions	55
2.3	Quadratic Forms	57
2.4	Derivatives	60
2.5	Partitioned Matrices	61
2.6	Geometrical Aspects	63
2.7	Exercises	70
3	Moving to Higher Dimensions	73
3.1	Covariance	74
3.2	Correlation	78
3.3	Summary Statistics	83
3.4	Linear Model for Two Variables	86
3.5	Simple Analysis of Variance	93
3.6	Multiple Linear Model	97

3.7	Boston Housing	101
3.8	Exercises	104
4	Multivariate Distributions	107
4.1	Distribution and Density Function	108
4.2	Moments and Characteristic Functions	113
4.3	Transformations	123
4.4	The Multinormal Distribution	125
4.5	Sampling Distributions and Limit Theorems	129
4.6	Heavy-Tailed Distributions	136
4.7	Copulae	150
4.8	Bootstrap	160
4.9	Exercises	163
5	Theory of the Multinormal	167
5.1	Elementary Properties of the Multinormal	167
5.2	The Wishart Distribution	174
5.3	Hotelling's T^2 -Distribution	176
5.4	Spherical and Elliptical Distributions	178
5.5	Exercises	180
6	Theory of Estimation	183
6.1	The Likelihood Function	184
6.2	The Cramer-Rao Lower Bound	188
6.3	Exercises	191
7	Hypothesis Testing	193
7.1	Likelihood Ratio Test	194
7.2	Linear Hypothesis	203
7.3	Boston Housing	220
7.4	Exercises	223

Part III Multivariate Techniques

8	Regression Models	229
8.1	General ANOVA and ANCOVA Models	231
8.1.1	ANOVA Models	231
8.1.2	ANCOVA Models	236
8.1.3	Boston Housing	237
8.2	Categorical Responses	238
8.2.1	Multinomial Sampling and Contingency Tables	238
8.2.2	Log-linear Models for Contingency Tables	239
8.2.3	Testing Issues with Count Data	243
8.2.4	Logit Models	246
8.3	Exercises	252
9	Decomposition of Data Matrices by Factors	255
9.1	The Geometric Point of View	256

9.2	Fitting the p -dimensional Point Cloud	257
9.3	Fitting the n -dimensional Point Cloud	260
9.4	Relations Between Subspaces	262
9.5	Practical Computation	263
9.6	Exercises	266
10	Principal Components Analysis	269
10.1	Standardized Linear Combination	270
10.2	Principal Components in Practice	274
10.3	Interpretation of the PCs	276
10.4	Asymptotic Properties of the PCs	280
10.5	Normalized Principal Components Analysis	283
10.6	Principal Components as a Factorial Method	284
10.7	Common Principal Components	290
10.8	Boston Housing	293
10.9	More Examples	294
10.10	Exercises	302
11	Factor Analysis	307
11.1	The Orthogonal Factor Model	308
11.2	Estimation of the Factor Model	314
11.3	Factor Scores and Strategies	322
11.4	Boston Housing	324
11.5	Exercises	329
12	Cluster Analysis	331
12.1	The Problem	332
12.2	The Proximity Between Objects	332
12.3	Cluster Algorithms	337
12.4	Boston Housing	345
12.5	Exercises	347
13	Discriminant Analysis	351
13.1	Allocation Rules for Known Distributions	351
13.2	Discrimination Rules in Practice	358
13.3	Boston Housing	364
13.4	Exercises	365
14	Correspondence Analysis	367
14.1	Motivation	367
14.2	Chi-square Decomposition	370
14.3	Correspondence Analysis in Practice	373
14.4	Exercises	383
15	Canonical Correlation Analysis	385
15.1	Most Interesting Linear Combination	385
15.2	Canonical Correlation in Practice	390
15.3	Exercises	395

16	Multidimensional Scaling	397
16.1	The Problem	397
16.2	Metric Multidimensional Scaling	402
16.3	Nonmetric Multidimensional Scaling	406
16.4	Exercises	412
17	Conjoint Measurement Analysis	413
17.1	Introduction	413
17.2	Design of Data Generation	415
17.3	Estimation of Preference Orderings	417
17.4	Exercises	424
18	Applications in Finance	427
18.1	Portfolio Choice	427
18.2	Efficient Portfolio	428
18.3	Efficient Portfolios in Practice	434
18.4	The Capital Pricing Model (CAPM)	436
18.5	Exercises	437
19	Computationally Intensive Techniques	439
19.1	Simplicial Depth	440
19.2	Projection Pursuit	443
19.3	Sliced Inverse Regression	448
19.4	Support Vector Machines	455
19.5	Classification and Regression Trees	469
19.6	Boston Housing	486
19.7	Exercises	486

Part IV Appendix

Appendix A	Symbols and Notations	493
Appendix B	Data	497
B.1	Boston Housing Data	497
B.2	Swiss Bank Notes	497
B.3	Car Data	498
B.4	Classic Blue Pullovers Data	498
B.5	U.S. Companies Data	499
B.6	French Food Data	499
B.7	Car Marks	499
B.8	French Baccalauréat Frequencies	499
B.9	Journaux Data	500
B.10	U.S. Crime Data	500
B.11	Plasma Data	501
B.12	WAIS Data	501
B.13	ANOVA Data	502
B.14	Timebudget Data	502
B.15	Geopol Data	503

B.16 U.S. Health Data 504

B.17 Vocabulary Data 504

B.18 Athletic Records Data 505

B.19 Unemployment Data 505

B.20 Annual Population Data 505

B.21 Bankruptcy Data I 505

B.22 Bankruptcy Data II 505

Bibliography 509

Index 513