

Contents

Preface	vii
1 Multivariate Data and Multivariate Analysis	1
1.1 Introduction	1
1.2 A brief history of the development of multivariate analysis	3
1.3 Types of variables and the possible problem of missing values .	4
1.3.1 Missing values	5
1.4 Some multivariate data sets.....	7
1.5 Covariances, correlations, and distances	12
1.5.1 Covariances.....	12
1.5.2 Correlations	14
1.5.3 Distances.....	14
1.6 The multivariate normal density function	15
1.7 Summary.....	23
1.8 Exercises	23
2 Looking at Multivariate Data: Visualisation	25
2.1 Introduction	25
2.2 The scatterplot.....	26
2.2.1 The bivariate boxplot	28
2.2.2 The convex hull of bivariate data	32
2.2.3 The chi-plot	34
2.3 The bubble and other glyph plots	34
2.4 The scatterplot matrix	39
2.5 Enhancing the scatterplot with estimated bivariate densities ..	42
2.5.1 Kernel density estimators	42
2.6 Three-dimensional plots	47
2.7 Trellis graphics.....	50
2.8 Stalactite plots.....	53
2.9 Summary.....	56
2.10 Exercises	60

3	Principal Components Analysis	61
3.1	Introduction	61
3.2	Principal components analysis (PCA)	61
3.3	Finding the sample principal components	63
3.4	Should principal components be extracted from the covariance or the correlation matrix?	65
3.5	Principal components of bivariate data with correlation coefficient r	68
3.6	Rescaling the principal components	70
3.7	How the principal components predict the observed covariance matrix	70
3.8	Choosing the number of components	71
3.9	Calculating principal components scores	72
3.10	Some examples of the application of principal components analysis	74
3.10.1	Head lengths of first and second sons	74
3.10.2	Olympic heptathlon results	78
3.10.3	Air pollution in US cities	86
3.11	The biplot	92
3.12	Sample size for principal components analysis	93
3.13	Canonical correlation analysis	94
3.13.1	Head measurements	96
3.13.2	Health and personality	99
3.14	Summary	101
3.15	Exercises	102
4	Multidimensional Scaling	105
4.1	Introduction	105
4.2	Models for proximity data	105
4.3	Spatial models for proximities: Multidimensional scaling	106
4.4	Classical multidimensional scaling	106
4.4.1	Classical multidimensional scaling: Technical details	107
4.4.2	Examples of classical multidimensional scaling	110
4.5	Non-metric multidimensional scaling	121
4.5.1	House of Representatives voting	123
4.5.2	Judgements of World War II leaders	124
4.6	Correspondence analysis	127
4.6.1	Teenage relationships	130
4.7	Summary	131
4.8	Exercises	132
5	Exploratory Factor Analysis	135
5.1	Introduction	135
5.2	A simple example of a factor analysis model	136
5.3	The k -factor analysis model	137

5.4	Scale invariance of the k -factor model	138
5.5	Estimating the parameters in the k -factor analysis model	139
5.5.1	Principal factor analysis	141
5.5.2	Maximum likelihood factor analysis	142
5.6	Estimating the number of factors	142
5.7	Factor rotation	143
5.8	Estimating factor scores	147
5.9	Two examples of exploratory factor analysis	148
5.9.1	Expectations of life	148
5.9.2	Drug use by American college students	151
5.10	Factor analysis and principal components analysis compared ..	157
5.11	Summary	159
5.12	Exercises	159
6	Cluster Analysis	163
6.1	Introduction	163
6.2	Cluster analysis	165
6.3	Agglomerative hierarchical clustering	166
6.3.1	Clustering jet fighters	171
6.4	K -means clustering	175
6.4.1	Clustering the states of the USA on the basis of their crime rate profiles	176
6.4.2	Clustering Romano-British pottery	180
6.5	Model-based clustering	183
6.5.1	Finite mixture densities	186
6.5.2	Maximum likelihood estimation in a finite mixture density with multivariate normal components	187
6.6	Displaying clustering solutions graphically	191
6.7	Summary	197
6.8	Exercises	200
7	Confirmatory Factor Analysis and Structural Equation Models	201
7.1	Introduction	201
7.2	Estimation, identification, and assessing fit for confirmatory factor and structural equation models	202
7.2.1	Estimation	202
7.2.2	Identification	203
7.2.3	Assessing the fit of a model	204
7.3	Confirmatory factor analysis models	206
7.3.1	Ability and aspiration	206
7.3.2	A confirmatory factor analysis model for drug use	211
7.4	Structural equation models	216
7.4.1	Stability of alienation	216
7.5	Summary	222

7.6	Exercises	223
8	The Analysis of Repeated Measures Data	225
8.1	Introduction	225
8.2	Linear mixed-effects models for repeated measures data	232
8.2.1	Random intercept and random intercept and slope models for the timber slippage data	233
8.2.2	Applying the random intercept and the random intercept and slope models to the timber slippage data ..	235
8.2.3	Fitting random-effect models to the glucose challenge data	240
8.3	Prediction of random effects	247
8.4	Dropouts in longitudinal data	248
8.5	Summary	257
8.6	Exercises	257
	References	259
	Index	271