

Contents

Preface	xiii
1 Introduction	1
1.1 Introduction	1
1.2 A look at data and models	3
1.2.1 Types of data	3
1.2.2 Multidimensional scaling models	5
1.3 Proximities	8
1.3.1 Similarity/dissimilarity coefficients for mixed data	14
1.3.2 Distribution of proximity coefficients	18
1.3.3 Similarity of species populations	18
1.3.4 Transforming from similarities to dissimilarities	21
1.3.5 The metric nature of dissimilarities	21
1.3.6 Dissimilarity of variables	22
1.3.7 Similarity measures on fuzzy sets	24
1.4 Matrix results	25
1.4.1 The spectral decomposition	26
1.4.2 The singular value decomposition	26
1.4.3 The Moore-Penrose inverse	29
2 Metric multidimensional scaling	31
2.1 Introduction	31
2.2 Classical scaling	31
2.2.1 Recovery of coordinates	32
2.2.2 Dissimilarities as Euclidean distances	34
2.2.3 Classical scaling in practice	36
2.2.4 How many dimensions?	38
2.2.5 A practical algorithm for classical scaling	38
2.2.6 A grave example	39
2.2.7 Classical scaling and principal components	43

2.2.8	The additive constant problem	45
2.3	Robustness	49
2.4	Metric least squares scaling	49
2.5	Critchley's intermediate method	52
2.6	Unidimensional scaling	53
2.6.1	A classic example	55
2.7	Grouped dissimilarities	57
2.8	Inverse scaling	58
3	Nonmetric multidimensional scaling	61
3.1	Introduction	61
3.1.1	R^p space and the Minkowski metric	63
3.2	Kruskal's approach	64
3.2.1	Minimising S with respect to the disparities	65
3.2.2	A configuration with minimum stress	68
3.2.3	Kruskal's iterative technique	69
3.2.4	Nonmetric scaling of breakfast cereals	71
3.2.5	STRESS1/2, monotonicity, ties and missing data	73
3.3	The Guttman approach	75
3.4	A further look at stress	76
3.4.1	Interpretation of stress	79
3.5	How many dimensions?	88
3.6	Starting configurations	89
3.7	Interesting axes in the configuration	90
4	Further aspects of multidimensional scaling	93
4.1	Other formulations of MDS	93
4.2	MDS Diagnostics	94
4.3	Robust MDS	96
4.4	Interactive MDS	98
4.5	Dynamic MDS	99
4.6	Constrained MDS	103
4.6.1	Spherical MDS	105
4.7	Statistical inference for MDS	107
4.8	Asymmetric dissimilarities	116
5	Procrustes analysis	123
5.1	Introduction	123
5.2	Procrustes analysis	124
5.2.1	Procrustes analysis in practice	127

5.2.2	The projection case	129
5.3	Historic maps	130
5.4	Some generalizations	132
5.4.1	Weighted Procrustes rotation	132
5.4.2	Generalized Procrustes analysis	135
5.4.3	The coefficient of congruence	137
5.4.4	Oblique Procrustes problem	138
5.4.5	Perturbation analysis	139
6	Monkeys, whisky and other applications	141
6.1	Introduction	141
6.2	Monkeys	141
6.3	Whisky	143
6.4	Aeroplanes	146
6.5	Yoghurts	148
6.6	Bees	149
7	Biplots	153
7.1	Introduction	153
7.2	The classic biplot	153
7.2.1	An example	154
7.2.2	Principal component biplots	157
7.3	Another approach	159
7.4	Categorical variables	162
8	Unfolding	165
8.1	Introduction	165
8.2	Nonmetric unidimensional unfolding	166
8.3	Nonmetric multidimensional unfolding	169
8.4	Metric multidimensional unfolding	173
8.4.1	The rating of nations	177
9	Correspondence analysis	181
9.1	Introduction	181
9.2	Analysis of two-way contingency tables	181
9.2.1	Distance between rows (columns) in a contingency table	184
9.3	The theory of correspondence analysis	185
9.3.1	The cancer example	187
9.3.2	Inertia	191

9.4	Reciprocal averaging	193
9.4.1	Algorithm for solution	193
9.4.2	An example: the Münsingen data	194
9.4.3	The whisky data	195
9.4.4	The correspondence analysis connection	197
9.4.5	Two-way weighted dissimilarity coefficients	198
9.5	Multiple correspondence analysis	200
9.5.1	A three-way example	202
10	Individual differences models	205
10.1	Introduction	205
10.2	The Tucker-Messick model	205
10.3	INDSCAL	206
10.3.1	The algorithm for solution	206
10.3.2	Identifying groundwater populations	208
10.3.3	Extended INDSCAL models	210
10.4	IDIOSCAL	211
10.5	PINDIS	212
11	ALSCAL, SMACOF and Gifi	217
11.1	ALSCAL	217
11.1.1	The theory	217
11.1.2	Minimising SSTRESS	219
11.2	SMACOF	221
11.2.1	The majorization algorithm	222
11.2.2	The majorizing method for nonmetric MDS	225
11.2.3	Tunnelling for a global minimum	226
11.3	Gifi	226
11.3.1	Homogeneity	227
12	Further m-mode, n-way models	233
12.1	CANDECOMP, PARAFAC and CANDELINC	233
12.2	DEDICOM and GIPSCAL	235
12.3	The Tucker models	236
12.3.1	Relationship to other models	238
12.4	One-mode, n -way models	238
12.5	Two-mode, three-way asymmetric scaling	243
12.6	Three-way unfolding	245

Appendix: Computer programs for multidimensional scaling	247
A.1 Computer programs	247
A.2 The accompanying CD-ROM	248
A.2.1 Installation instructions	249
A.2.2 Data and output	250
A.2.3 To run the menu	250
A.2.4 Program descriptions	251
A.3 The data provided	252
A.4 To manipulate and analyse data	255
A.5 Inputting user data	259
A.5.1 Data format	259
A.6 Error messages	262
References	271
Author index	293
Subject index	299