

Contents

1	Introduction	1
1.1	Why Numerical Ecology?	1
1.2	Why R?	2
1.3	Readership and Structure of the Book	2
1.4	How to Use This Book	3
1.5	The Data Sets	4
1.5.1	The Doubs Fish Data	5
1.5.2	The Oribatid Mite Data	7
1.6	A Quick Reminder About Help Sources	7
1.7	Now It Is Time	9
2	Exploratory Data Analysis	11
2.1	Objectives	11
2.2	Data Exploration	11
2.2.1	Data Extraction	11
2.2.2	Species Data: First Contact	12
2.2.3	Species Data: A Closer Look	14
2.2.4	Ecological Data Transformation	21
2.2.5	Environmental Data	28
2.3	Conclusion	34
3	Association Measures and Matrices	35
3.1	Objectives	35
3.2	The Main Categories of Association Measures (Short Overview)	35
3.2.1	Q Mode and R Mode	36
3.2.2	Symmetrical or Asymmetrical Coefficients in Q Mode: The Double-Zero Problem	36
3.2.3	Association Measures for Qualitative or Quantitative Data	37
3.2.4	To Summarize	37

3.3	Q Mode: Computing Dissimilarity Matrices Among Objects	38
3.3.1	Q Mode: Quantitative Species Data	39
3.3.2	Q Mode: Binary (Presence-Absence) Species Data	42
3.3.3	Q Mode: Quantitative Data (Excluding Species Abundances)	46
3.3.4	Q Mode: Binary Data (Excluding Species Presence-Absence Data)	48
3.3.5	Q Mode: Mixed Types Including Categorical (Qualitative Multiclass) Variables	49
3.4	R Mode: Computing Dependence Matrices Among Variables . . .	51
3.4.1	R Mode: Species Abundance Data	52
3.4.2	R Mode: Species Presence-Absence Data	52
3.4.3	R Mode: Quantitative and Ordinal Data (Other than Species Abundances)	53
3.4.4	R Mode: Binary Data (Other than Species Abundance Data)	55
3.5	Pre-transformations for Species Data	55
3.6	Conclusion	57
4	Cluster Analysis	59
4.1	Objectives	59
4.2	Clustering Overview	59
4.3	Hierarchical Clustering Based on Links	62
4.3.1	Single Linkage Agglomerative Clustering	62
4.3.2	Complete Linkage Agglomerative Clustering	64
4.4	Average Agglomerative Clustering	65
4.5	Ward's Minimum Variance Clustering	68
4.6	Flexible Clustering	69
4.7	Interpreting and Comparing Hierarchical Clustering Results . . .	70
4.7.1	Introduction	70
4.7.2	Cophenetic Correlation	71
4.7.3	Looking for Interpretable Clusters	74
4.8	Non-hierarchical Clustering	96
4.8.1	<i>k</i> -means Partitioning	96
4.8.2	Partitioning Around Medoids (PAM)	103
4.9	Comparison with Environmental Data	107
4.9.1	Comparing a Typology with External Data (ANOVA Approach)	107
4.9.2	Comparing Two Typologies (Contingency Table Approach)	111
4.10	Species Assemblages	111
4.10.1	Simple Statistics on Group Contents	111
4.10.2	Kendall's <i>W</i> Coefficient of Concordance	112
4.10.3	Species Assemblages in Presence-Absence Data	115
4.10.4	Species Co-occurrence Network	117

4.11	Indicator Species	120
4.11.1	Introduction	120
4.11.2	IndVal: Species Indicator Values	120
4.11.3	Correlation-Type Indices	125
4.12	Multivariate Regression Trees (MRT):	
	Constrained Clustering	126
4.12.1	Introduction	126
4.12.2	Computation (Principle)	127
4.12.3	Application Using Packages <code>mvpart</code> and <code>MVPARTwrap</code>	129
4.12.4	Combining MRT and IndVal	134
4.13	MRT as a Monothetic Clustering Method	135
4.14	Sequential Clustering	138
4.15	A Very Different Approach: Fuzzy Clustering	141
4.15.1	Fuzzy <i>c</i> -means Using Package <code>cluster</code> 's Function <code>fanny()</code>	141
4.15.2	Noise Clustering Using the <code>vegclust()</code> Function	146
4.16	Conclusion	150
5	Unconstrained Ordination	151
5.1	Objectives	151
5.2	Ordination Overview	151
5.2.1	Multidimensional Space	151
5.2.2	Ordination in Reduced Space	152
5.3	Principal Component Analysis (PCA)	153
5.3.1	Overview	153
5.3.2	PCA of the Environmental Variables of the Doubs River Data Using <code>rda()</code>	154
5.3.3	PCA on Transformed Species Data	166
5.3.4	Domain of Application of PCA	169
5.3.5	PCA Using Function <code>PCA.newr()</code>	170
5.3.6	Imputation of Missing Values in PCA	171
5.4	Correspondence Analysis (CA)	175
5.4.1	Introduction	175
5.4.2	CA Using Function <code>cca()</code> of Package <code>vegan</code>	176
5.4.3	CA Using Function <code>CA.newr()</code>	181
5.4.4	Arch Effect and Detrended Correspondence Analysis (DCA)	182
5.4.5	Multiple Correspondence Analysis (MCA)	183
5.5	Principal Coordinate Analysis (PCoA)	187
5.5.1	Introduction	187
5.5.2	Application of PCoA to the Doubs Data Set Using <code>cmdscale()</code> and <code>vegan</code>	188
5.5.3	Application of PCoA to the Doubs Data Set Using <code>pcoa()</code>	190

5.6	Nonmetric Multidimensional Scaling (NMDS)	193
5.6.1	Introduction	193
5.6.2	Application to the Doubs Fish Data	193
5.6.3	PCoA or NMDS?	196
5.7	Hand-Written PCA Ordination Function	198
6	Canonical Ordination	203
6.1	Objectives	203
6.2	Canonical Ordination Overview	204
6.3	Redundancy Analysis (RDA)	204
6.3.1	Introduction	204
6.3.2	RDA of the Doubs River Data	206
6.3.3	Distance-Based Redundancy Analysis (db-RDA)	249
6.3.4	A Hand-Written RDA Function	253
6.4	Canonical Correspondence Analysis (CCA)	256
6.4.1	Introduction	256
6.4.2	CCA of the Doubs River Data	257
6.5	Linear Discriminant Analysis (LDA)	263
6.5.1	Introduction	263
6.5.2	Discriminant Analysis Using <code>lda()</code>	264
6.6	Other Asymmetric Analyses	268
6.6.1	Principal Response Curves (PRC)	268
6.6.2	Co-correspondence Analysis (CoCA)	271
6.7	Symmetric Analysis of Two (or More) Data Sets	274
6.8	Canonical Correlation Analysis (CCorA)	275
6.8.1	Introduction	275
6.8.2	Canonical Correlation Analysis Using <code>CCorA()</code>	275
6.9	Co-inertia Analysis (CoIA)	277
6.9.1	Introduction	277
6.9.2	Co-inertia Analysis Using Function <code>coinertia()</code> of <code>ade4</code>	278
6.10	Multiple Factor Analysis (MFA)	282
6.10.1	Introduction	282
6.10.2	Multiple Factor Analysis Using <code>FactoMineR</code>	283
6.11	Relating Species Traits and Environment	287
6.11.1	The Fourth-Corner Method	288
6.11.2	RLQ Analysis	290
6.11.3	Application in R	291
6.12	Conclusion	296
7	Spatial Analysis of Ecological Data	299
7.1	Objectives	299
7.2	Spatial Structures and Spatial Analysis: A Short Overview	300
7.2.1	Introduction	300
7.2.2	Induced Spatial Dependence and Spatial Autocorrelation	301
7.2.3	Spatial Scale	302

7.2.4	Spatial Heterogeneity	303
7.2.5	Spatial Correlation or Autocorrelation Functions and Spatial Correlograms	303
7.2.6	Testing for the Presence of Spatial Correlation: Conditions	308
7.2.7	Modelling Spatial Structures	309
7.3	Multivariate Trend-Surface Analysis	309
7.3.1	Introduction	309
7.3.2	Trend-Surface Analysis in Practice	310
7.4	Eigenvector-Based Spatial Variables and Spatial Modelling	314
7.4.1	Introduction	314
7.4.2	Distance-Based Moran's Eigenvector Maps (dbMEM) and Principal Coordinates of Neighbour Matrices (PCNM)	315
7.4.3	MEM in a Wider Context: Weights Other than Geographic Distances	333
7.4.4	MEM with Positive or Negative Spatial Correlation: Which Ones should Be Used?	348
7.4.5	Asymmetric Eigenvector Maps (AEM): When Directionality Matters	348
7.5	Another Way to Look at Spatial Structures: Multiscale Ordination (MSO)	355
7.5.1	Principle	355
7.5.2	Application to the Mite Data – Exploratory Approach	356
7.5.3	Application to the Detrended Mite and Environmental Data	359
7.6	Space-Time Interaction Test in Multivariate ANOVA, Without Replicates	361
7.6.1	Introduction	361
7.6.2	Testing the Space-Time Interaction with the <i>sti</i> Functions	364
7.7	Conclusion	367
8	Community Diversity	369
8.1	Objectives	369
8.2	The Multiple Facets of Diversity	370
8.2.1	Introduction	370
8.2.2	Species Diversity Measured by a Single Number	370
8.2.3	Taxonomic Diversity Indices in Practice	374
8.3	When Space Matters: Alpha, Beta and Gamma Diversities	379
8.4	Beta Diversity	379
8.4.1	Beta Diversity Measured by a Single Number	379
8.4.2	Beta Diversity as the Variance of the Community Composition Table: SCBD and LCBF Indices	382

8.4.3	Partitioning Beta Diversity into Replacement, Richness Difference and Nestedness Components	388
8.5	Functional Diversity, Functional Composition and Phylogenetic Diversity of Communities	404
8.5.1	Alpha Functional Diversity	404
8.5.2	Beta Taxonomic, Phylogenetic and Functional Diversities	408
8.6	Conclusion	412
	Bibliography	413
	Index	427