

Daniel Borcard • François Gillet
Pierre Legendre

Numerical Ecology with R

 Springer

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.....	4
1.5.2	The Oribatid Mite Data.....	5
1.6	A Quick Reminder about Help Sources.....	7
1.7	Now It Is Time.....	7
2	Exploratory Data Analysis	9
2.1	Objectives.....	9
2.2	Data Exploration.....	9
2.2.1	Data Extraction.....	9
2.2.2	Species Data: First Contact.....	10
2.2.3	Species Data: A Closer Look.....	12
2.2.4	Species Data Transformation.....	18
2.2.5	Environmental Data.....	24
2.3	Conclusion.....	30
3	Association Measures and Matrices	31
3.1	Objectives.....	31
3.2	The Main Categories of Association Measures (Short Overview).....	31
3.2.1	Q Mode and R Mode.....	32
3.2.2	Symmetrical or Asymmetrical Coefficients in Q Mode: The Double-Zero Problem.....	32

3.2.3	Association Measures for Qualitative or Quantitative Data.....	33
3.2.4	To Summarize.....	33
3.3	Q Mode: Computing Distance Matrices Among Objects.....	34
3.3.1	Q Mode: Quantitative Species Data.....	34
3.3.2	Q Mode: Binary (Presence–Absence) Species Data.....	36
3.3.3	Q Mode: Quantitative Data (Excluding Species Abundances).....	41
3.3.4	Q Mode: Binary Data (Excluding Species Presence–Absence Data).....	43
3.3.5	Q Mode: Mixed Types, Including Categorical (Qualitative Multiclass) Variables.....	44
3.4	R Mode: Computing Dependence Matrices Among Variables.....	46
3.4.1	R Mode: Species Abundance Data.....	46
3.4.2	R Mode: Species Presence–Absence Data.....	47
3.4.3	R Mode: Quantitative and Ordinal Data (Other than Species Abundances).....	47
3.4.4	R Mode: Binary Data (Other than Species Abundance Data).....	49
3.5	Pre-transformations for Species Data.....	50
3.6	Conclusion.....	50
4	Cluster Analysis.....	53
4.1	Objectives.....	53
4.2	Clustering Overview.....	53
4.3	Hierarchical Clustering Based on Links.....	56
4.3.1	Single Linkage Agglomerative Clustering.....	56
4.3.2	Complete Linkage Agglomerative Clustering.....	57
4.4	Average Agglomerative Clustering.....	59
4.5	Ward’s Minimum Variance Clustering.....	61
4.6	Flexible Clustering.....	63
4.7	Interpreting and Comparing Hierarchical Clustering Results.....	63
4.7.1	Introduction.....	63
4.7.2	Cophenetic Correlation.....	63
4.7.3	Looking for Interpretable Clusters.....	65
4.8	Non-hierarchical Clustering.....	79
4.8.1	<i>k</i> -Means Partitioning.....	80
4.8.2	Partitioning Around Medoids.....	84
4.9	Comparison with Environmental Data.....	87
4.9.1	Comparing a Typology with External Data (ANOVA Approach).....	88
4.9.2	Comparing Two Typologies (Contingency Table Approach).....	91

4.10	Species Assemblages	91
4.10.1	Simple Statistics on Group Contents	91
4.10.2	Kendall's <i>W</i> Coefficient of Concordance	92
4.10.3	Species Assemblages in Presence–Absence Data	95
4.10.4	IndVal: Species Indicator Values	97
4.11	Multivariate Regression Trees: Constrained Clustering	99
4.11.1	Introduction	99
4.11.2	Computation (Principle)	99
4.11.3	Application Using Packages <i>mvpart</i> and <i>MVPARTwrap</i>	102
4.11.4	Combining MRT and IndVal	107
4.11.5	MRT as a “Chronological” Clustering Method	108
4.12	A Very Different Approach: Fuzzy Clustering	110
4.12.1	Fuzzy <i>c</i> -Means Clustering Using <i>cluster</i> 's Function <i>fanny()</i>	110
4.13	Conclusion	114
5	Unconstrained Ordination	115
5.1	Objectives	115
5.2	Ordination Overview	115
5.2.1	Multidimensional Space	115
5.2.2	Ordination in Reduced Space	116
5.3	Principal Component Analysis	117
5.3.1	Overview	117
5.3.2	PCA on the Environmental Variables of the Doubs Data Set Using <i>rda()</i>	118
5.3.3	PCA on Transformed Species Data	128
5.3.4	Domain of Application of PCA	130
5.3.5	PCA Using Function <i>PCA()</i>	131
5.4	Correspondence Analysis	132
5.4.1	Introduction	132
5.4.2	CA Using Function <i>cca()</i> of Package <i>vegan</i>	133
5.4.3	CA Using Function <i>CA()</i>	138
5.4.4	Arch Effect and Detrended Correspondence Analysis	139
5.4.5	Multiple Correspondence Analysis	140
5.5	Principal Coordinate Analysis	140
5.5.1	Introduction	140
5.5.2	Application to the Doubs Data Set Using <i>cmdscale</i> and <i>vegan</i>	141
5.5.3	Application to the Doubs Data Set Using <i>pcoa()</i>	143
5.6	Nonmetric Multidimensional Scaling	145
5.6.1	Introduction	145
5.6.2	Application to the Fish Data	146
5.7	Handwritten Ordination Function	149

6	Canonical Ordination	153
6.1	Objectives	153
6.2	Canonical Ordination Overview	154
6.3	Redundancy Analysis	154
6.3.1	Introduction	154
6.3.2	RDA of the Doubs River Data.....	156
6.3.3	A Hand-Written RDA Function	195
6.4	Canonical Correspondence Analysis	198
6.4.1	Introduction	198
6.4.2	CCA of the Doubs Data.....	199
6.5	Linear Discriminant Analysis	207
6.5.1	Introduction	207
6.5.2	Discriminant Analysis Using <code>lda()</code>	208
6.6	Other Asymmetrical Analyses.....	210
6.7	Symmetrical Analysis of Two (or More) Data Sets.....	211
6.8	Canonical Correlation Analysis.....	211
6.8.1	Introduction	211
6.8.2	Canonical Correlation Analysis using <code>CCorA</code>	212
6.9	Co-inertia Analysis	214
6.9.1	Introduction	214
6.9.2	Co-inertia Analysis Using <code>ade4</code>	215
6.10	Multiple Factor Analysis	218
6.10.1	Introduction	218
6.10.2	Multiple Factor Analysis Using <code>FactoMineR</code>	219
6.11	Conclusion	224
7	Spatial Analysis of Ecological Data	227
7.1	Objectives	227
7.2	Spatial Structures and Spatial Analysis: A Short Overview.....	228
7.2.1	Introduction	228
7.2.2	Induced Spatial Dependence and Spatial Autocorrelation	229
7.2.3	Spatial Scale	230
7.2.4	Spatial Heterogeneity	231
7.2.5	Spatial Correlation or Autocorrelation Functions and Spatial Correlograms	232
7.2.6	Testing for the Presence of Spatial Correlation: Conditions	236
7.2.7	Modelling Spatial Structures.....	238
7.3	Multivariate Trend-Surface Analysis.....	238
7.3.1	Introduction	238
7.3.2	Trend-Surface Analysis in Practice	239
7.4	Eigenvector-Based Spatial Variables and Spatial Modelling.....	243
7.4.1	Introduction	243

7.4.2	Classical Distance-Based MEM, Formerly Called Principal Coordinates of Neighbour Matrices.....	244
7.4.3	MEM in a Wider Context: Weights Other than Geographic Distances.....	263
7.4.4	MEM with Positive or Negative Spatial Correlation: Which Ones Should Be Used?	278
7.4.5	Asymmetric Eigenvector Maps : When Directionality Matters.....	279
7.5	Another Way to Look at Spatial Structures: Multiscale Ordination.....	285
7.5.1	Principle	285
7.5.2	Application to the Mite Data: Exploratory Approach.....	286
7.5.3	Application to the Detrended Mite and Environmental Data	289
7.6	Conclusion	292
	Bibliographical References.....	293
	Index.....	301