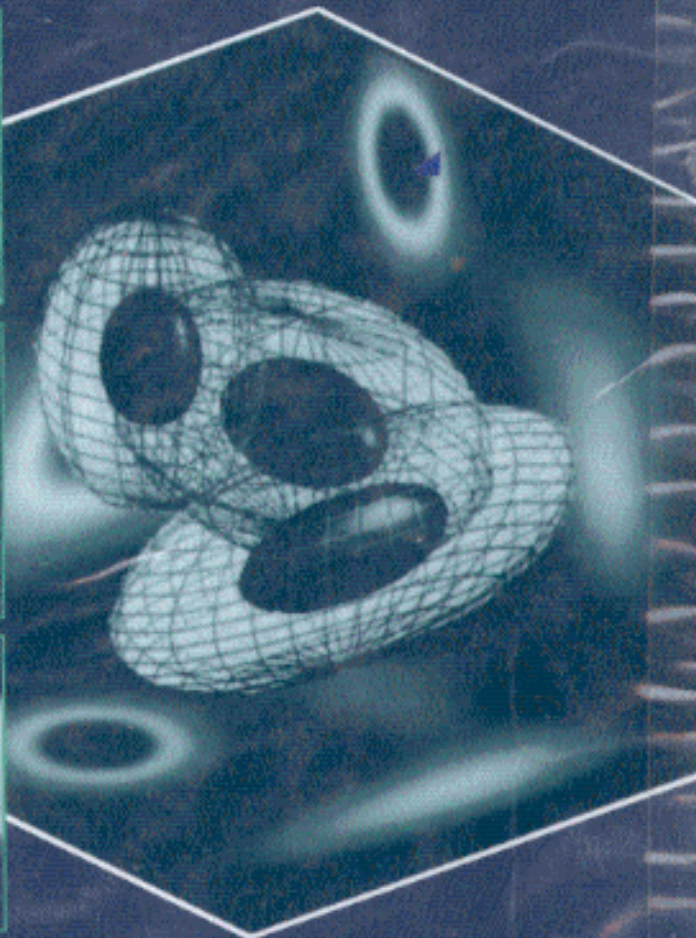


Multivariate Statistics for Wildlife and Ecology Research



Kevin McGarigal
Sam Cushman
Susan Stafford

Contents

Preface	v
1 Introduction and Overview	1
1.1 Objectives	1
1.2 Multivariate Statistics: An Ecological Perspective	2
1.3 Multivariate Description and Inference	9
1.4 Multivariate Confusion!	10
1.5 Types of Multivariate Techniques	14
1.5.1 Ordination	14
1.5.2 Cluster Analysis	15
1.5.3 Discriminant Analysis	15
1.5.4 Canonical Correlation Analysis	16
Bibliography	16
2 Ordination: Principal Components Analysis	19
2.1 Objectives	19
2.2 Conceptual Overview	20
2.2.1 Ordination	20
2.2.2 Principal Components Analysis (PCA)	23
2.3 Geometric Overview	24
2.4 The Data Set	25
2.5 Assumptions	27
2.5.1 Multivariate Normality	28
2.5.2 Independent Random Sample and the Effects of Outliers	31
2.5.3 Linearity	33

2.6	Sample Size Requirements	34
2.6.1	General Rules	34
2.6.2	Specific Rules	35
2.7	Deriving the Principal Components	35
2.7.1	The Use of Correlation and Covariance Matrices	35
2.7.2	Eigenvalues and Associated Statistics	37
2.7.3	Eigenvectors and Scoring Coefficients	39
2.8	Assessing the Importance of the Principal Components	41
2.8.1	Latent Root Criterion	41
2.8.2	Scree Plot Criterion	41
2.8.3	Broken Stick Criterion	43
2.8.4	Relative Percent Variance Criterion	43
2.8.5	Significance Tests	45
2.9	Interpreting the Principal Components	50
2.9.1	Principal Component Structure	50
2.9.2	Significance of Principal Component Loadings	51
2.9.3	Interpreting the Principal Component Structure	53
2.9.4	Communality	56
2.9.5	Principal Component Scores and Associated Plots ...	57
2.10	Rotating the Principal Components	58
2.11	Limitations of Principal Components Analysis	61
2.12	R-Factor Versus Q-Factor Ordination	61
2.13	Other Ordination Techniques	63
2.13.1	Polar Ordination	63
2.13.2	Factor Analysis	64
2.13.3	Nonmetric Multidimensional Scaling	66
2.13.4	Reciprocal Averaging	67
2.13.5	Detrended Correspondence Analysis	68
2.13.6	Canonical Correspondence Analysis	69
	Bibliography	73
	Appendix 2.1	78

3	Cluster Analysis	81
3.1	Objectives	81
3.2	Conceptual Overview	82
3.3	The Definition of Cluster	85
3.4	The Data Set	86
3.5	Clustering Techniques	89
3.6	Nonhierarchical Clustering	91
3.6.1	Polythetic Agglomerative Nonhierarchical Clustering. .	92
3.6.2	Polythetic Divisive Nonhierarchical Clustering	93
3.7	Hierarchical Clustering	94
3.7.1	Polythetic Agglomerative Hierarchical Clustering ...	95

3.7.2	Polythetic Divisive Hierarchical Clustering	120
3.8	Evaluating the Stability of the Cluster Solution	121
3.9	Complementary Use of Ordination and Cluster Analysis	122
3.10	Limitations of Cluster Analysis	123
	Bibliography	124
	Appendix 3.1	127
4	Discriminant Analysis	129
4.1	Objectives	129
4.2	Conceptual Overview	130
4.2.1	Overview of Canonical Analysis of Discriminance. . .	133
4.2.2	Overview of Classification.	134
4.2.3	Analogy with Multiple Regression Analysis and Multivariate Analysis of Variance	136
4.3	Geometric Overview	137
4.4	The Data Set	138
4.5	Assumptions	141
4.5.1	Equality of Variance-Covariance Matrices	141
4.5.2	Multivariate Normality	144
4.5.3	Singularities and Multicollinearity	146
4.5.4	Independent Random Sample and the Effects of Outliers	151
4.5.5	Prior Probabilities Are Identifiable	152
4.5.6	Linearity.	153
4.6	Sample Size Requirements	153
4.6.1	General Rules.	153
4.6.2	Specific Rules	154
4.7	Deriving the Canonical Functions.	155
4.7.1	Stepwise Selection of Variables.	155
4.7.2	Eigenvalues and Associated Statistics	158
4.7.3	Eigenvectors and Canonical Coefficients	159
4.8	Assessing the Importance of the Canonical Functions	161
4.8.1	Relative Percent Variance Criterion.	161
4.8.2	Canonical Correlation Criterion	162
4.8.3	Classification Accuracy	163
4.8.4	Significance Tests	167
4.8.5	Canonical Scores and Associated Plots	169
4.9	Interpreting the Canonical Functions	169
4.9.1	Standardized Canonical Coefficients	171
4.9.2	Total Structure Coefficients	171
4.9.3	Covariance-Controlled Partial <i>F</i> -Ratios	173
4.9.4	Significance Tests Based on Resampling Procedures ..	175
4.9.5	Potency Index	175

4.10	Validating the Canonical Functions	176
4.10.1	Split-Sample Validation	177
4.10.2	Validation Using Resampling Procedures	178
4.11	Limitations of Discriminant Analysis	179
	Bibliography	180
	Appendix 4.1	185
5	Canonical Correlation Analysis	189
5.1	Objectives	189
5.2	Conceptual Overview	190
5.3	Geometric Overview	195
5.4	The Data Set	196
5.5	Assumptions	198
5.5.1	Multivariate Normality	198
5.5.2	Singularities and Multicollinearity	199
5.5.3	Independent Random Sample and the Effects of Outliers	202
5.5.4	Linearity	203
5.6	Sample Size Requirements	204
5.6.1	General Rules	204
5.6.2	Specific Rules	204
5.7	Deriving the Canonical Variates	205
5.7.1	The Use of Covariance and Correlation Matrices	205
5.7.2	Eigenvalues and Associated Statistics	206
5.7.3	Eigenvectors and Canonical Coefficients	208
5.8	Assessing the Importance of the Canonical Variates	209
5.8.1	Canonical Correlation Criterion	209
5.8.2	Canonical Redundancy Criterion	212
5.8.3	Significance Tests	216
5.8.4	Canonical Scores and Associated Plots	218
5.9	Interpreting the Canonical Variates	220
5.9.1	Standardized Canonical Coefficients	220
5.9.2	Structure Coefficients	221
5.9.3	Canonical Cross-Loadings	223
5.9.4	Significance Tests Based on Resampling Procedures	225
5.10	Validating the Canonical Variates	225
5.10.1	Split-Sample Validation	226
5.10.2	Validation Using Resampling Procedures	227
5.11	Limitations of Canonical Correlation Analysis	227
	Bibliography	228
	Appendix 5.1	230

6	Summary and Comparison	233
6.1	Objectives	233
6.2	Relationship Among Techniques	234
6.2.1	Purpose and Source of Variation Emphasized	234
6.2.2	Statistical Procedure	236
6.2.3	Type of Statistical Technique and Variable Set Characteristics	237
6.2.4	Data Structure	238
6.2.5	Sampling Design	239
6.3	Complementary Use of Techniques	241
	Appendix: Acronyms Used in This Book	249
	Glossary	251
	Index	279