

Inhaltsverzeichnis

Teil I Grundlagen

1	Beispiele multivariater Datensätze	3
2	Elementare Behandlung der Daten	15
2.1	Beschreibung und Darstellung univariater Datensätze	15
2.1.1	Beschreibung und Darstellung qualitativer Merkmale	17
2.1.2	Beschreibung und Darstellung quantitativer Merkmale	19
2.2	Beschreibung und Darstellung multivariater Datensätze	25
2.2.1	Beschreibung und Darstellung von Datenmatrizen quantitativer Merkmale	25
2.2.2	Datenmatrizen qualitativer Merkmale	40
2.3	Datenbehandlung in R	44
2.3.1	R als mächtiger Taschenrechner	44
2.3.2	Univariate Datenanalyse	46
2.3.3	Multivariate Datenanalyse	57
2.4	Ergänzungen und weiterführende Literatur	67
2.5	Übungen	68
3	Mehrdimensionale Zufallsvariablen	71
3.1	Problemstellung	71
3.2	Univariate Zufallsvariablen	71
3.3	Zufallsmatrizen und Zufallsvektoren	76
3.4	Multivariate Normalverteilung	88
4	Ähnlichkeits- und Distanzmaße	91
4.1	Problemstellung	91
4.2	Bestimmung der Distanzen und Ähnlichkeiten aus der Datenmatrix	92
4.2.1	Quantitative Merkmale	92
4.2.2	Binäre Merkmale	103
4.2.3	Qualitative Merkmale mit mehr als zwei Merkmalsausprägungen	107

4.2.4	Qualitative Merkmale, deren Merkmalsausprägungen geordnet sind	107
4.2.5	Unterschiedliche Messniveaus	107
4.3	Distanzmaße in R	110
4.4	Direkte Bestimmung der Distanzen	116
4.5	Übungen	117

Teil II Darstellung hochdimensionaler Daten in niedrigdimensionalen Räumen

5	Hauptkomponentenanalyse	121
5.1	Problemstellung	121
5.2	Hauptkomponentenanalyse bei bekannter Varianz-Kovarianz-Matrix . . .	126
5.3	Hauptkomponentenanalyse bei unbekannter Varianz-Kovarianz-Matrix .	129
5.4	Praktische Aspekte	133
5.4.1	Anzahl der Hauptkomponenten	135
5.4.2	Überprüfung der Güte der Anpassung	139
5.4.3	Analyse auf Basis der Varianz-Kovarianz-Matrix oder auf Basis der Korrelationsmatrix	142
5.5	Wie geht man bei einer Hauptkomponentenanalyse vor?	145
5.6	Hauptkomponentenanalyse in R	150
5.7	Ergänzungen und weiterführende Literatur	155
5.8	Übungen	155
6	Mehrdimensionale Skalierung	159
6.1	Problemstellung	159
6.2	Metrische mehrdimensionale Skalierung	161
6.2.1	Theorie	161
6.2.2	Praktische Aspekte	178
6.2.3	Metrische mehrdimensionale Skalierung der Rangreihung von Politikerpaaren	180
6.2.4	Metrische mehrdimensionale Skalierung in R	182
6.3	Nichtmetrische mehrdimensionale Skalierung	185
6.3.1	Theorie	185
6.3.2	Nichtmetrische mehrdimensionale Skalierung in R	194
6.4	Ergänzungen und weiterführende Literatur	196
6.5	Übungen	197
7	Procrustes-Analyse	199
7.1	Problemstellung und Grundlagen	199
7.2	Illustration der Vorgehensweise	202
7.3	Theorie	208
7.4	Procrustes-Analyse der Reisezeiten	210

7.5 Procrustes-Analyse in R 212

7.6 Ergänzungen und weiterführende Literatur 215

7.7 Übungen 215

Teil III Abhängigkeitsstrukturen

8 Lineare Regression 219

8.1 Problemstellung und Modell 219

8.2 Schätzung der Parameter 222

8.3 Praktische Aspekte 230

8.3.1 Interpretation der Parameter bei mehreren erklärenden Variablen . 230

8.3.2 Güte der Anpassung 235

8.3.3 Tests 238

8.4 Lineare Regression in R 243

8.5 Ergänzungen und weiterführende Literatur 246

8.6 Übungen 246

9 Explorative Faktorenanalyse 249

9.1 Problemstellung und Grundlagen 249

9.2 Theorie 257

9.2.1 Allgemeines Modell 257

9.2.2 Nichteindeutigkeit der Lösung 261

9.2.3 Schätzung von L und Ψ 264

9.3 Praktische Aspekte 270

9.3.1 Bestimmung der Anzahl der Faktoren 270

9.3.2 Rotation 270

9.4 Faktorenanalyse in R 272

9.5 Ergänzungen und weiterführende Literatur 275

9.6 Übungen 275

10 Hierarchische loglineare Modelle 279

10.1 Problemstellung und Grundlagen 279

10.2 Zweidimensionale Kontingenztabellen 289

10.2.1 Modell 0 290

10.2.2 Modell A 291

10.2.3 IPF-Algorithmus 293

10.2.4 Modell B 295

10.2.5 Modell A, B 296

10.2.6 Modell AB 298

10.2.7 Modellselektion 299

10.3 Dreidimensionale Kontingenztabellen 302

10.3.1 Modell der totalen Unabhängigkeit 302

10.3.2 Modell der Unabhängigkeit einer Variablen	306
10.3.3 Modell der bedingten Unabhängigkeit	309
10.3.4 Modell ohne Drei-Faktor-Interaktion	312
10.3.5 Saturiertes Modell	315
10.3.6 Modellselektion	315
10.4 Loglineare Modelle in R	317
10.5 Ergänzungen und weiterführende Literatur	323
10.6 Übungen	323

Teil IV Gruppenstrukturen

11 Einfaktorielle Varianzanalyse	329
11.1 Problemstellung	329
11.2 Univariate einfaktorielle Varianzanalyse	330
11.2.1 Theorie	330
11.2.2 Praktische Aspekte	339
11.3 Multivariate einfaktorielle Varianzanalyse	345
11.4 Jonckheere-Test	348
11.5 Einfaktorielle Varianzanalyse in R	354
11.6 Ergänzungen und weiterführende Literatur	360
11.7 Übungen	360
12 Diskriminanzanalyse	363
12.1 Problemstellung und theoretische Grundlagen	363
12.2 Diskriminanzanalyse bei normalverteilten Grundgesamtheiten	372
12.2.1 Diskriminanzanalyse bei Normalverteilung mit bekannten Parametern	372
12.2.2 Diskriminanzanalyse bei Normalverteilung mit unbekannten Parametern	378
12.3 Fishers lineare Diskriminanzanalyse	383
12.4 Logistische Diskriminanzanalyse	388
12.5 Klassifikationsbäume	391
12.6 Praktische Aspekte	399
12.7 Diskriminanzanalyse in R	405
12.8 Ergänzungen und weiterführende Literatur	411
12.9 Übungen	412
13 Clusteranalyse	415
13.1 Problemstellung	415
13.2 Hierarchische Clusteranalyse	416
13.2.1 Theorie	416
13.2.2 Verfahren der hierarchischen Clusterbildung	422

13.2.3 Praktische Aspekte	449
13.2.4 Hierarchische Clusteranalyse in R	454
13.3 Partitionierende Verfahren	457
13.3.1 Theorie	457
13.3.2 Praktische Aspekte	460
13.3.3 Partitionierende Verfahren in R	466
13.4 Clusteranalyse von Daten der Regionen	471
13.5 Ergänzungen und weiterführende Literatur	474
13.6 Übungen	474
Anhang A Mathematische Grundlagen	477
A.1 Matrizenrechnung	477
A.1.1 Definitionen und spezielle Matrizen	478
A.1.2 Matrixverknüpfungen	479
A.1.3 Inverse Matrix	484
A.1.4 Orthogonale Matrizen	485
A.1.5 Spur einer Matrix	486
A.1.6 Determinante einer Matrix	487
A.1.7 Lineare Gleichungssysteme	488
A.1.8 Eigenwerte und Eigenvektoren	490
A.1.9 Spektralzerlegung einer symmetrischen Matrix	493
A.1.10 Singulärwertzerlegung	495
A.1.11 Quadratische Formen	496
A.2 Extremwerte	497
A.2.1 Gradient und Hesse-Matrix	498
A.2.2 Extremwerte ohne Nebenbedingungen	500
A.2.3 Extremwerte unter Nebenbedingungen	501
A.3 Matrizenrechnung in R	503
Anhang B Eigene R-Funktionen	509
B.1 Quartile	509
B.2 Monotone Regression	509
B.3 STRESS1	510
B.4 Bestimmung einer neuen Konfiguration	510
B.5 Kophenetische Matrix	511
B.6 Gamma-Koeffizient	512
B.7 Bestimmung der Zugehörigkeit zu Klassen	513
B.8 Silhouette	513
B.9 Zeichnen einer Silhouette	514
Anhang C Tabellen	517
C.1 Standardnormalverteilung	517

C.2	χ^2 -Verteilung	519
C.3	t -Verteilung	520
C.4	F -Verteilung	521
Literatur		523
Sachverzeichnis		529