

# Inhaltsverzeichnis

## Teil I Grundlagen

<b>1 Beispiele multivariater Datensätze .....</b>	<b>3</b>
<b>2 Elementare Behandlung der Daten .....</b>	<b>13</b>
2.1 Beschreibung und Darstellung univariater Datensätze .....	13
2.1.1 Beschreibung und Darstellung qualitativer Merkmale ..	15
2.1.2 Beschreibung und Darstellung quantitativer Merkmale ..	17
2.2 Beschreibung und Darstellung multivariater Datensätze .....	22
2.2.1 Beschreibung und Darstellung von Datenmatrizen quantitativer Merkmale .....	22
2.2.2 Beschreibung und Darstellung von Datenmatrizen qualitativer Merkmale .....	36
2.3 Datenbehandlung in S-PLUS.....	41
2.3.1 Univariate Datenanalyse .....	41
2.3.2 Multivariate Datenanalyse .....	51
2.4 Ergänzungen und weiterführende Literatur .....	61
2.5 Übungen .....	61
<b>3 Mehrdimensionale Zufallsvariablen .....</b>	<b>65</b>
3.1 Problemstellung .....	65
3.2 Univariate Zufallsvariablen .....	65
3.3 Zufallsmatrizen und Zufallsvektoren .....	70
3.4 Die multivariate Normalverteilung .....	81
<b>4 Ähnlichkeits- und Distanzmaße .....</b>	<b>83</b>
4.1 Problemstellung .....	83
4.2 Bestimmung der Distanzen und Ähnlichkeiten aus der Datenmatrix .....	84
4.2.1 Quantitative Merkmale.....	84
4.2.2 Binäre Merkmale .....	94
4.2.3 Qualitative Merkmale mit mehr als zwei Merkmalsausprägungen .....	98
4.2.4 Qualitative Merkmale, deren Merkmalsausprägungen geordnet sind .....	98

## XII Inhaltsverzeichnis

4.2.5	Unterschiedliche Messniveaus . . . . .	98
4.3	Distanzmaße in S-PLUS . . . . .	102
4.4	Direkte Bestimmung der Distanzen . . . . .	108
4.5	Übungen . . . . .	110
 <b>Teil II Darstellung hochdimensionaler Daten in niedrigdimensionalen Räumen</b>		
5	<b>Hauptkomponentenanalyse</b> . . . . .	115
5.1	Problemstellung . . . . .	115
5.2	Hauptkomponentenanalyse bei bekannter Varianz-Kovarianz-Matrix . . . . .	120
5.3	Hauptkomponentenanalyse bei unbekannter Varianz-Kovarianz-Matrix . . . . .	123
5.4	Praktische Aspekte . . . . .	126
5.4.1	Anzahl der Hauptkomponenten . . . . .	128
5.4.2	Überprüfung der Güte der Anpassung . . . . .	130
5.4.3	Analyse auf Basis der Varianz-Kovarianz-Matrix oder auf Basis der Korrelationsmatrix . . . . .	133
5.5	Wie geht man bei einer Hauptkomponentenanalyse vor? . . . . .	135
5.6	Hauptkomponentenanalyse in S-PLUS . . . . .	140
5.7	Ergänzungen und weiterführende Literatur . . . . .	144
5.8	Übungen . . . . .	145
6	<b>Mehrdimensionale Skalierung</b> . . . . .	149
6.1	Problemstellung . . . . .	149
6.2	Metrische mehrdimensionale Skalierung . . . . .	150
6.2.1	Theorie . . . . .	150
6.2.2	Praktische Aspekte . . . . .	165
6.2.3	Metrische mehrdimensionale Skalierung der Rangreihung der Politikerpaare . . . . .	167
6.2.4	Metrische mehrdimensionale Skalierung in S-PLUS . . . . .	169
6.3	Nichtmetrische mehrdimensionale Skalierung . . . . .	171
6.3.1	Theorie . . . . .	171
6.3.2	Nichtmetrische mehrdimensionale Skalierung in S-PLUS . . . . .	179
6.4	Ergänzungen und weiterführende Literatur . . . . .	182
6.5	Übungen . . . . .	182
7	<b>Procrustes-Analyse</b> . . . . .	185
7.1	Problemstellung und Grundlagen . . . . .	185
7.2	Illustration der Vorgehensweise . . . . .	187
7.3	Theorie . . . . .	192
7.4	Procrustes-Analyse der Reisezeiten . . . . .	194
7.5	Procrustes-Analyse in S-PLUS . . . . .	196

7.6	Ergänzungen und weiterführende Literatur . . . . .	198
7.7	Übungen . . . . .	198

## Teil III Abhängigkeitsstrukturen

<b>8</b>	<b>Lineare Regression</b> . . . . .	203
8.1	Problemstellung und Modell . . . . .	203
8.2	Schätzung der Parameter . . . . .	206
8.3	Praktische Aspekte . . . . .	211
8.3.1	Interpretation der Parameter bei mehreren erklärenden Variablen . . . . .	211
8.3.2	Die Güte der Anpassung . . . . .	215
8.3.3	Tests . . . . .	219
8.4	Lineare Regression in S-PLUS . . . . .	222
8.5	Ergänzungen und weiterführende Literatur . . . . .	224
8.6	Übungen . . . . .	224
<b>9</b>	<b>Explorative Faktorenanalyse</b> . . . . .	227
9.1	Problemstellung und Grundlagen . . . . .	227
9.2	Theorie . . . . .	235
9.2.1	Das allgemeine Modell . . . . .	235
9.2.2	Nichteindeutigkeit der Lösung . . . . .	238
9.2.3	Schätzung von $L$ und $\Psi$ . . . . .	240
9.3	Praktische Aspekte . . . . .	246
9.3.1	Bestimmung der Anzahl der Faktoren . . . . .	246
9.3.2	Rotation . . . . .	247
9.4	Faktorenanalyse in S-PLUS . . . . .	249
9.5	Ergänzungen und weiterführende Literatur . . . . .	251
9.6	Übungen . . . . .	252
<b>10</b>	<b>Hierarchische loglineare Modelle</b> . . . . .	255
10.1	Problemstellung und Grundlagen . . . . .	255
10.2	Zweidimensionale Kontingenztabellen . . . . .	265
10.2.1	Modell 0 . . . . .	265
10.2.2	Modell A . . . . .	267
10.2.3	Der IPF-Algorithmus . . . . .	268
10.2.4	Modell B . . . . .	270
10.2.5	Modell A, B . . . . .	272
10.2.6	Modell AB . . . . .	274
10.2.7	Modellselektion . . . . .	274
10.3	Dreidimensionale Kontingenztabellen . . . . .	277
10.3.1	Das Modell der totalen Unabhängigkeit . . . . .	277
10.3.2	Das Modell der Unabhängigkeit einer Variablen . . . . .	281
10.3.3	Das Modell der bedingten Unabhängigkeit . . . . .	285

## XIV Inhaltsverzeichnis

10.3.4 Das Modell ohne Drei-Faktor-Interaktion .....	288
10.3.5 Das saturierte Modell .....	290
10.3.6 Modellselektion .....	291
10.4 Loglineare Modelle in S-PLUS .....	292
10.5 Ergänzungen und weiterführende Literatur .....	298
10.6 Übungen .....	298

## Teil IV Gruppenstruktur

<b>11 Einfaktorielle Varianzanalyse .....</b>	<b>303</b>
11.1 Problemstellung .....	303
11.2 Univariate einfaktorielle Varianzanalyse .....	303
11.2.1 Theorie .....	303
11.2.2 Praktische Aspekte .....	311
11.3 Multivariate einfaktorielle Varianzanalyse .....	317
11.4 Einfaktorielle Varianzanalyse in S-PLUS .....	319
11.5 Ergänzungen und weiterführende Literatur .....	322
11.6 Übungen .....	322
<b>12 Diskriminanzanalyse .....</b>	<b>325</b>
12.1 Problemstellung und theoretische Grundlagen .....	325
12.2 Diskriminanzanalyse bei normalverteilten Grundgesamtheiten	334
12.2.1 Diskriminanzanalyse bei Normalverteilung mit bekannten Parametern .....	334
12.2.2 Diskriminanzanalyse bei Normalverteilung mit unbekannten Parametern .....	340
12.3 Fishers lineare Diskriminanzanalyse .....	343
12.4 Logistische Diskriminanzanalyse .....	348
12.5 Klassifikationsbäume .....	351
12.6 Praktische Aspekte .....	358
12.7 Diskriminanzanalyse in S-PLUS .....	362
12.8 Ergänzungen und weiterführende Literatur .....	369
12.9 Übungen .....	369
<b>13 Clusteranalyse .....</b>	<b>373</b>
13.1 Problemstellung .....	373
13.2 Hierarchische Clusteranalyse .....	374
13.2.1 Theorie .....	374
13.2.2 Verfahren der hierarchischen Clusterbildung .....	381
13.2.3 Praktische Aspekte .....	407
13.2.4 Hierarchische Clusteranalyse in S-PLUS .....	411
13.3 Partitionierende Verfahren .....	414
13.3.1 Theorie .....	414
13.3.2 Praktische Aspekte .....	417

13.3.3 Partitionierende Verfahren in S-PLUS .....	422
13.4 Clusteranalyse der Daten der Regionen .....	427
13.5 Ergänzungen und weiterführende Literatur .....	429
13.6 Übungen .....	429

## Teil V Anhänge

<b>A Mathematische Grundlagen .....</b>	435
A.1 Matrizenrechnung .....	435
A.1.1 Definitionen und spezielle Matrizen .....	436
A.1.2 Matrixverknüpfungen .....	437
A.1.3 Die inverse Matrix .....	441
A.1.4 Orthogonale Matrizen .....	442
A.1.5 Spur einer Matrix .....	443
A.1.6 Determinante einer Matrix .....	444
A.1.7 Lineare Gleichungssysteme .....	445
A.1.8 Eigenwerte und Eigenvektoren .....	447
A.1.9 Die Spektralzerlegung einer symmetrischen Matrix .....	449
A.1.10 Die Singulärwertzerlegung .....	451
A.1.11 Quadratische Formen .....	452
A.2 Extremwerte .....	453
A.2.1 Der Gradient und die Hesse-Matrix .....	454
A.2.2 Extremwerte ohne Nebenbedingungen .....	456
A.2.3 Extremwerte unter Nebenbedingungen .....	457
A.3 Matrizenrechnung in S-PLUS .....	459
<b>B S-PLUS-Funktionen .....</b>	465
B.1 Quartile .....	465
B.2 Distanzmatrix .....	465
B.3 Monotone Regression .....	466
B.4 STRESS1 .....	467
B.5 Bestimmung einer neuen Konfiguration .....	467
B.6 Kophenetische Matrix .....	468
B.7 Gamma-Koeffizient .....	469
B.8 Bestimmung der Zugehörigkeit zu Klassen .....	469
B.9 Silhouette .....	470
B.10 Zeichnen einer Silhouette .....	471
<b>C Tabellen .....</b>	473
C.1 Standardnormalverteilung .....	473
C.2 $\chi^2$ -Verteilung .....	475
C.3 $t$ -Verteilung .....	476
C.4 $F$ -Verteilung .....	477

XVI      Inhaltsverzeichnis

<b>Literaturverzeichnis</b> .....	479
<b>Index</b> .....	485