

# Stichwortverzeichnis

- Abhängigkeit
  - , lineare 45
  - , nichtlineare 201
- Abstand 25
  - (→ Distanz)
- Abweichungsindex 100
- Abzählverfahren 161, 278
  - , reduzierte 171, 175
- adaptive Verfahren 279ff.
- Ähnlichkeit 14, 20, 24ff.
  - , mittlere 82
  - , Transformation in Distanz 77f.
  - , zwischen Klassen 81ff.
  - , zwischen Objekten 24ff.
- Ähnlichkeitsmaß 25
  - , äquivalente -e 30, 55
  - , axiomatische Forderungen 25, 49
  - , dichotomes 210
  - , empirische Festlegung 29
    - für Rasterbilder 50
  - , induzierte Präordnung 30
  - , invariantes 26, 50
  - , probabilistisches 63, 70, 72, 75, 231
  - , Vergleich von -en 32ff., 55
  - , vertauschungsinvariantes 50, 51, 60, 68
  - von Watanabe 62
  - , Wahl eines -es 26, 49f.
- Ähnlichkeitsmatrix 25
  - , positiv definite 40, 75, 77f.
  - -reihenfolge (→ Präordnung) 21, 212
  - -methoden 20
- Ähnlichkeitsschranke
  - , empirische Festlegung 228, 231, 331
  - , Variation einer 229, 307, 419
  - als Signifikanzschranke 28, 205, 229, 307
- Alternativproblem 88
- Anfangsklassifikation
  - , iterative Verbesserung 219ff., 232, 332ff., 351ff., 382
  - , Wahl der 173, 223, 338, 353
- Anhang eines Objekts 344
- Anordnung 30
- Anwendungen, Hinweise auf – (vgl. auch Bibliographie!) 13–17, 50, 226, 316, 339f., 420
- Anziehung eines Objekts 339
- a-posteriori-Wahrscheinlichkeit 118, 131, 139
- Approximation
  - einer Relation 211
  - optimaler Gruppierungen 161, 171ff., 218ff., 382f.
  - , stochastische 284ff.
  - , ultrametrische 372ff., 427f.
- a-priori-Verteilung 116f.
  - , konjugierte 130, 139
- äquivalente| Ähnlichkeits- (Distanz-)maße 30, 55, 306, 311, 376, 391
  - Objekte 25, 362
- Äquivalenz|klasse 26, 108
  - -relation 108, 211
- association-analysis 418
- Assoziationskoeffizient 62
- Assoziationsmaße 59ff.
- Ausreißer 107, 229
- Austauschverfahren 220, 223
- Average-Linkage-Methode 231, 402f.
  
- Baum 303, 310
- Bayesrisiko 117
- Bayesverfahren 118
  - , empirische 291
  - für Diskriminationsproblem 251
  - für Gruppierungsprobleme 116–158
- Bellsche Zahlen  $B_N$  110
- Beobachtungsvektor 18f.
- Binomialverteilung 52
  
- C\*-Kriterium 205, 207, 211, 221
- Clique 318
  - , disjunkte -n 330
  - , maximale 318ff., 423ff.
  - der Stufe  $d$  318, 394
  - , Konstruktion von -n 320, 324ff.
  - , Kombination von -n 330, 424
- Cluster-Analyse (→ Klassifikation, automatische)
- Complete-Linkage-Methode 315, 381, 392ff., 400
- cophenetic correlation 33, 382
- $\chi^2$ -Verteilung 36, 84, 98, 434

- Daten** (binäre, quantitative, qualitative) 20  
 –, fehlende 37, 75ff.  
 –, gemischte 20, 64, 74f., 409, 414, 417, 418  
**Datenmatrix** 18  
**Datenreduktion** 13  
**Dendrogramm** 357, 362  
**Determinantenkriterium** 138, 141, 155, 185ff., 220, 222  
 – als ML-Verfahren 138  
 –, modifiziertes 194  
 –, Optimalität 141  
**Differenz zweier Mengen** 431  
 –, symmetrische 34, 431  
 dilatierendes Verfahren 391  
**disjunkte Mengen** 430  
 – Gruppierung 15, 22, 105ff.  
**Diskretisierung** 67  
**Diskriminanzanalyse** 23, 170, 251  
 –, sequentielle 293  
 –, Bayesverfahren 251  
**Diskriminanzfunktion** 283  
**Diskriminationsebene, optimale** 173  
**Diskriminationsproblem** 88  
**Diskriminationsverfahren** 251  
**Distanz** 24  
 –, euklidische 35, 71, 83  
 –, empirische Festlegung 21  
 –,  $k$ -ultrametrische 427f.  
 –,  $L_p$ - 39  
 –, -matrix 20, 25  
 –, mittlere 92  
 –, rekursiv definierte 404f.  
 –, Transformation in Ähnlichkeiten 77ff.  
 –, ultrametrische 366ff.  
 – zwischen Klassen 81, 88ff.  
 – zwischen Objekten 24ff.  
**Distanzmaß**  
 –, allgemeine Methoden und Gesichtspunkte 26ff., 88f.  
 –, äquivalente -e, 30, 34, 306, 311, 312, 377  
 –, empirische -e, 47f.  
 – für Rasterbilder 50  
 –, induzierte Präordnung 29  
 –, invariantes 27, 84  
 –, metrisches 26, 37, 39  
 –, monotonen 386, 405  
 –, nicht monotonen 386, 401  
 –, skaleninvarianten 27, 36, 42, 46f., 84  
 –, translationsinvarianten 27, 35, 40, 42, 47, 84  
 –, ( $k$ -)ultrametrische Approximation 372ff., 427f.,  
 –, Vergleich von -en, 32ff.  
 –, Wahl eines -es, 26  
**Distanzreihenfolge** 29, 307, 377  
 (→ Präordnung)  
**Distanzschranke**  
 –, empirische Festlegung 228ff., 230, 331  
 –, Variation einer 229, 307, 329, 419f.  
 – als Signifikanzschranke 28, 205, 229, 307, 331  
**dissimilarity analysis** 414  
**Divergenz** 89  
**Dokumentation** 17, 339, 359  
**dominierte Ultrametrik** 372  
**Dreiecksungleichung** 26, 366  
 **$d$ -Umgebung** 298  
**Durchmesser** 92, 318, 362  
 – als Index 380, 393f.  
**Durchschnitt** 430  
 **$d$ -verbunden** 422  
 $\delta^+$ ,  $\delta^-$  (→ Ultrametrik)  
  
**Entropie** 93ff., 135, 136, 193, 242  
**Erwartungswert** 433  
**Evolutionsprozesse** 359  
**Expansionsbedingung** 378, 390  
**Exponentialzahlen  $B_N$**  110  
**Extremalproblem, kontinuierliches** 164ff., 178, 195, 281  
 –, stationäre Lösung 168, 283  
  
**Faktoranalyse** 247ff.  
**Fehlerwahrscheinlichkeit**  
 – 1. und 2. Art 88  
 – eines Gruppierungsverfahrens 117  
 –, totale 118, 149, 251  
**flexible strategy** 406  
**Fourier-Transformierte** 295  
  
**Gamma-Verteilung** 130  
 –, inverse 130  
 **$G(d)$  (→ Graph)**  
**Gewichtung von Merkmalen** 29, 65, 69, 76  
**Gradientenvektor** 273, 282  
**Gradientenverfahren** 272ff.  
**Graph** 300  
 –, bewerteter 301  
 – einer Relation (Präordnung) 31f., 213  
 –  $G(d)$  278, 301ff., 319, 374, 389, 421  
 –, gerichteter 308f.

- , Verbindbarkeit in -en, 301
- , vollständiger 301
- , zusammenhängender 303
- GR-Gruppe 339, 349ff.
- Gruppe 105
  - (auch → Klasse, Klassen)
  - absolut hoher Punktdichte 278, 422
  - , axiomatische Festlegung 298ff., 308ff., 317, 318ff.
  - der Stufe  $d$  299ff.
  - , durch Hyperebene charakterisiert 195ff.
  - relativ hoher Punktdichte 272
  - , runde, kompakte 101, 219, 227
  - , verzweigte, langgestreckte 101, 231, 275, 300, 306, 330, 424
- Gruppierung (als Mengensystem)
  - , disjunkte 15, 22, 105ff.
  - , optimale 159–217, 307
  - , lokal 172, 220
  - der Stufe  $d$  279, 300ff., 376, 398, 425
  - , Invarianzeigenschaft 306
  - , Modifikation 278f., 306
  - , Optimalität 204, 307
  - , exhaustive 22, 106
  - , nicht disjunkte 15, 16, 22, 317ff., 331
  - , hierarchische 330, 331, 423ff.
  - , nicht exhaustive 22, 235ff.
  - , disjunkte 107, 233
  - , nicht disjunkte 332ff.
  - , hierarchische 420ff.
  - , optimale 194
  - , hierarchische 17, 22, 356ff.
  - , optimale 370ff.
  - , stabile 333
  - , Test auf Existenz 147ff., 391f.
  - , Vergleich von -en, 161
- Gruppierung (als Prozeß, auch → Klassifikation)
  - , disjunkte 104–315, 337
  - , Problem der 106
  - , mit hierarchischen Verfahren 176, 409ff., 419
  - durch reziproke Paare 308ff.
  - durch Zusammenhangskomponenten 204, 298ff., 398ff.
  - , Optimalität 204, 307
  - , gleichzeitige, von Objekten und Merkmalen 194f.
  - , hierarchische 176, 356–429
  - , mit nichthierarchischen Verfahren 411
  - , nichtdisjunkte 423ff.
  - , Problem der 358, 370
  - , nicht exhaustive 233, 279ff., 332ff.
  - , stufenweise 77, 245f., 411
- Gruppierungsverfahren (auch → Verfahren)
  - , iterative 171ff., 200, 222, 332, 351
  - , optimale 113, 116ff.
  - , randomisiertes 116
  - , zulässiges 121
- Gütekriterium 159ff., 382f.
  - , asymptotische Verteilung 216ff.
  - , axiomatische Forderungen an, 206
  - bei Präordnung 212ff.
  - für qualitative Merkmale 206ff.
  - , informationstheoretisches 235, 408ff. 414
  - , invariante 187, 213
  - , monotones 171, 194, 202, 407
  - , nicht monotones 194, 221
  - von Benzecri und de la Vega 213f.
  - , Vergleich derselben 214ff.
- Hauptkomponenten 103, 196f., 237ff.
  - -methode 237ff.
  - , Optimalität 238
  - -vektoren 238
- Heterogenität| einer Objektmenge 91, 227
  - , mittlere, der Klassen 202, 221
- Heterogenitätsmaß 91ff.
  - als Index 362
  - , informationstheoretisches 93ff., 102, 230, 362, 408, 414, 415
  - , subadditives 97, 99
  - , supraadditives 97, 202, 415
- Heterogenitätszuwachs 227, 371, 407, 410
- Hierarchie 360
  - , binäre (dichotome) 361, 364, 386, 411
  - , indizierte 362
  - $k\mathfrak{H}$  von Ling 422
  - $\mathfrak{H}^*$  der  $K$ -Gruppen 379ff.
  - , nichtdisjunkte 330, 331, 423ff.
  - , optimale 370ff., 382
  - , totale 360
  - von Apresjan 379ff.
  - , Invarianz 312, 377
  - , Konstruktion 397ff.
  - ( $\mathfrak{H}^-, h^-$ ), Dendrogramm 373
  - , Invarianz 377
  - , Konstruktion 389ff.
  - , Optimalität 383
  - , statistische Eigenschaften 391

- , Beziehung zu Zusammenhangskomponenten 376, 419
- ( $\mathfrak{H}^+$ ,  $h^+$ ), Dendrogramm 379ff.
- , Invarianz 377
- , Konstruktion 392ff.
- , Optimalität 395
- Histogramm 268, 278
- Homogenität einer Objektmenge 91
- Homogenitätsmaß 92, 186, 410
  - , probabilistisches 93
- Hybriden 107, 229
- hypergeometrische Verteilung 60f., 64
  
- Identifizierbarkeit 252ff.
- Index 361
  - -bedingung, schwache 361, 386, 393
  - -bedingung, starke 362, 388, 395
- indizierte Hierarchie 362
- Information retrieval 17
- Informations|abstand 89
  - -analyse 408
  - -inhalt 93ff., 102, 409
  - -matrix 260
  - -radius 89
- , Normal- 89f.
- -theoretische Maße (→ Heterogenitätsmaß, Gütekriterium)
- -zuwachs 98, 409, 410, 414
- Inhomogenität einer Objektmenge (→ Heterogenität) 91
- integer-link-linkage 391
- Interdependenz 63, 99
- Invarianz
  - von Distanzmaßen 26
  - bzgl. linearer Transformationen 42, 84ff., 187
  - bzgl. Vertauschung 50
- Inversion
  - bei Ähnlichkeiten 34
  - bei Dendrogrammen 362, 363, 386, 401
- Inzidenzmatrix
  - einer Partition 108, 211
  - einer Relation 108, 210
- iterative Verfahren 171ff., 200, 222, 332, 351
  - , Endlichkeit 172f., 333ff., 352
  
- $k$ -Cluster 422
- ( $k$ ,  $d$ )|-Cluster 422
  - -dicht 422
  - -zusammenhängend 422
- $k$ -Dendrogramm 423, 427
  - und  $k$ -Ultrametrik 428
- Kern|funktion 265
  - -gruppe 330, 344
  - -punkte 101, 225
- , Wahl der 226, 229ff.
- Kettenbildung (→ Verkettung)
- $K$ -Gruppe 312ff., 331, 379
  - als Gruppe der Stufe  $d$  315
  - als maximale Clique 315
  - als 1-Nachbarschaftsgruppe 331
- , Hierarchie  $\mathfrak{H}^*$  der -n 379, 397
- , Konstruktion 315, 398
- , Separation 312, 381
- , Invarianz 312
- $k$ -Hierarchie 423ff.
- Klasse 105, 360
  - (→ Gruppe)
  - , monothetische 23
  - , polythetische 23
  - , runde, kompakte 101f., 219, 227
  - , verzweigte, gekrümmte 101, 231, 275, 300, 306, 330, 424
- Klassen
  - -anzahl 105
  - , bekannte 106
  - , unbekannte 107, 146–158, 221
  - , Bestimmung mit Signifikanztest 158, 262
  - , optimale Festlegung 194
  - , linear trennbare 168, 189
  - , kleine 121, 163, 276, 408, 413
  - -mittelpunkt 83, 100
  - -zentren, Wahl der 224
- Klassifikation (→ Gruppierung)
  - , automatische 3, 14
  - und Diskriminanzanalyse 23
  - , künstliche 163
- $k$ -link-method 391
- $k$ -Nachbarschaftsgruppe 331
- Koeffizient
  - , nichtmetrischer 47, 71
  - von Jaccard 54
  - von K. Pearson 87
- kompakte Klassen 101, 219, 227
- Komplement 431
- konservatives Verfahren 391
- Kontingenz
  - , mittlere quadratische 61
  - -tafel 49, 61
- kontrahierendes Verfahren 390

- Kontraktionsbedingung 372, 390  
 konvexe | Hülle 168  
 – Menge 168  
 – Partition 169, 177  
 Konzentrationsellipsoid 42  
 Kovarianz 434  
 Kovarianzmatrix  
 –, bekannte 113ff., 147  
 –, Eigenwerte (-vektoren) der 41, 103, 196, 238ff.  
 –, empirische, der Merkmale 41, 84, 103, 138, 238, 243  
 –, in den Klassen 85f., 138, 196, 201, 243  
 –, kombinierte 84  
 –, zwischen den Klassen 244  
 –, empirische, der Objekte 44  
 –, gleiche, in den Klassen 85ff., 114ff., 137ff., 186, 234  
 –, unbekannte 137ff., 153  
 –, verschiedene, in den Klassen 89f., 130f., 142ff., 193  
 Korrelation 434  
 – der Merkmale ( $\rightarrow$  Merkmal)  
 Korrelationskoeffizient, empirischer  
 – der Merkmale 48, 66, 417  
 – der Objekte (als Ähnlichkeitsmaß) 44ff., 59ff., 71  
 –, multipler 48, 418  
 –, partieller 48  
 $k$ -Partition 424  
 Kreuzproduktverhältnis 59  
 Kriterium  
 ( $\rightarrow$  Gütekriterium)  
 Kugel 228, 231, 327  
 $k$ -Ultrametrik 427  
 –, maximale dominierte 428  
  
 latent-class-analysis 263  
 Lernen | mit Lehrer 285  
 – ohne Lehrer 285  
 Likelihood-Quotienten-Test 88, 155, 262  
 $L_k$ -Gruppierungsmethode 331  
 $L_r$ -Distanzen 39  
 –, normierte 39, 75  
 –, modifizierte 40  
  
 Mahalanobis-Abstand  
 – zwischen Objekten 28, 40ff.  
 – zwischen Objektklassen (Typ I–III) 84, 89f., 190, 191f., 246, 260  
 –, Verteilung der 43, 84ff.  
  
 matching coefficient 51  
 Matrix 432  
 –, positiv definite 433  
 –, Diagonal- 433  
 MAXIMINDIST 224, 234  
 Maximum-Likelihood | -Methode 115, 257  
 – -Verfahren 114ff., 122, 128ff., 133, 137f., 142f., 192f.  
 – -Schätzung 257ff.  
 Medianmethode 405ff.  
 Merkmal  
 –, binäres 20, 48  
 –, symmetrisches 50  
 –, unsymmetrisches 50  
 –, mehrstufiges 20, 66ff.  
 –, qualitatives 19  
 –, quantitatives 18  
 –, trennendes 416  
 Merkmale 18  
 –, abhängige 28, 66, 76, 85, 186  
 –, gemischte 20, 64, 74f., 409, 414, 417, 418  
 –, Gewichtung der 65, 68, 69ff., 76f.  
 –, korrelierte 43, 48, 187  
 –, linear abhängige 45  
 –, Normierung der 36, 37f., 40  
 –, unabhängige 52, 72, 75, 162, 243  
 –, unkorrelierte 45  
 –, Transformation auf 41ff., 246  
 Merkmalsalternativen  
 – -anzahl, Berücksichtigung der, 69  
 –, exklusive 68  
 –, geordnete 67, 71ff., 74  
 –, seltene 65, 69, 418  
 –, ungeordnete 67, 68  
 Merkmalshäufigkeit 65ff., 69ff., 76, 90  
 Merkmalsraum 19  
 Metrik 26  
 Minimalbaum 303ff., 307, 389ff.  
 –, Approximation 305  
 –, Konstruktion 304  
 Minimal-Distanz-Partition 165f., 190, 291, 236  
 –, iterierte 171ff., 222ff., 236  
 –, modifizierte 222, 236, 332ff.  
 –, verallgemeinerte 199, 201, 283  
 Minimal-Distanz-Regel 172, 191  
 Mischung  
 – von Verteilungen 250, 271  
 – von Normalverteilungen 173, 252–263, 271  
 –, Identifizierbarkeit 252ff.

- , Parameterschätzung bei 257ff.
- Mischungs|anteile 250
  - -komponenten 250
  - -problem 251ff.
- und Gruppierungsproblem 252
- Mittel|punkt 41, 83, 100
  - -senkrechte 170, 288, 290
  - -wertvektor 41, 83
- M-Koeffizient 32, 51ff., 56ff., 76, 214
  - , verallgemeinerter 68, 207
- ML-Verfahren (→ Maximum-Likelihood-Verfahren)
- Modalwert 224, 270, 282
  - , Bestimmung 272
- Modelle
  - , deterministische 23
  - , stochastische 22, 27, 113–158, 250–263, 279–290, 391
- Modus 270
- Momentenmethode 262
- $\mu$ -Abstand 32
- Nachbar
  - ,  $k$ -nächste -n 306
  - , nächster 308
- Nachbarschaft einer Menge 331
- Nachfolgerklassen 361, 381
- nichtmetrischer Koeffizient 47, 71
- Niveau einer Klasse 362
- Norm eines Vektors 432
- Normalverteilung 434
- Normierung der Daten 36, 37f., 40
- 0-1-Daten 48
- Objekt
  - , einklassiges 322
  - , isoliertes 321, 339
  - , transferierbares 344
  - , typischstes 100, 226, 230, 235
  - , vitales 344
- Objekte
  - , ähnliche 210, 298, 320
  - , äquivalente 25, 208, 362, 379, 386
  - , unklassifizierbare 104, 107, 2a9, 229, 317
- Objektmenge 18
  - mit Dendrogramm-Struktur 370, 382
- Optimalitätskriterium (→ Gütekriterium)
- Optimierung
  - , dynamische 203
  - , lineare 181ff.
  - , nichtlineare 180
- Ordnung 30
  - orthogonal 432
  - orthonormiert 432
- $\mathfrak{P}$  (= Menge zulässiger Partitionen) 114, 160
- Parameter
  - , Variation von -n 229, 307, 329, 419f.
  - -wahl 28, 205, 228ff., 307, 331
- Partition 106ff.
  - , Anzahl der -en 109ff., 176
  - der Stufe  $h$  363
  - , feinere 109
  - , gröbere 109
  - , konvexe 169f., 177, 190
  - , lokal optimale 161, 172, 220
  - , optimale 159ff.
  - , Approximation 218f.
  - , asymptotisches Verhalten 169, 208
  - , des  $R^P$  164ff., 280ff.
  - , des  $R^1$  178f.
  - , stabile 191
  - , stationäre 168, 170, 172, 292
  - , suboptimale 161, 219
  - , Typ einer 112, 214f.
  - , Unterschied zweier -en 206f.
  - , Verfeinerung 109
  - , Vergrößerung 109
  - , zentrale 207ff.
- Partitionenhierarchie 364, 371
  - , Anzahl der -n 368ff.
  - , indizierte 364
- Potentialfunktion 78, 294
  - , Methode der 293–297
  - , Wahl der 295
- Präordnung 29ff., 56, 80, 212, 303, 377, 383
  - einer Partition 109
  - , empirische Festlegung 30
  - , totale 30
  - , Unterschied zweier -en 33
- Projektionsmethoden 219, 237ff.
  - , nichtlineare 79f., 241, 248
- proportional-link-linkage 391
- Punkt
  - -dichte 101, 264, 267, 272ff., 278, 430
  - , Analyse der 249ff., 264, 270ff.
  - , Kern- 101, 225
  - -konzentration 101
  - , zentraler 100
- Radius 367

- , lokaler 379
- Rang, Ranggröße 46
- rank-order-Methode 331
- Rasterbilder 50, 226
- Relation 21, 29
  - , Graph einer 31
  - , reflexive 30, 210
  - , symmetrische 210ff.
  - , Approximation durch Äquivalenzrelation 211ff.
- , totale 30
- , transitive 30
- Rechenprogramme, Hinweise auf 203, 221, 234, 235, 241, 258, 305, 327, 330, 400, 427
- reziprokes Paar 308
- R-Gruppe 338ff.
  - , elementare 344
  - , größte 342
  - , Konstruktion 341ff.
- Risikofunktion 117
- $r_p$  (Ähnlichkeitsmaß) 78
- Rohdaten 18
  
- Scatter-Matrix (→ Streuungsmatrix)
- Schwerpunkt 83
- selbstadaptierende Verfahren 279ff., 286
- Separation| von Klassen 160, 192, 203, 317, 351
  - der  $K$ -Gruppen 312, 381
- sequentielle Verfahren (→ Verfahren)
- $S$ -Gruppe 347f.
- $S^*$ -Gruppe 347f.
- Signifikanzschranke
  - , konservative 229, 410
  - zur Festlegung von Parametern 28, 205, 229, 230, 307, 331
- Single-Linkage-Methode
  - , nichthierarchische Fassung 300ff., 419
  - , Modifikationen 278, 307
  - , hierarchische Fassung 315, 373, 387ff., 400, 404, 422, 427
  - , axiomatische Charakterisierung 391
  - , Modifikationen 420, 422
- Skaleninvarianz 27
- Skalierung, mehrdimensionale 79f., 241
- $S$ -Koeffizient 32, 53, 65
  - , modifizierter (4.7) 54, 56, 64
- Spannweite 37, 40, 74
- Spur 432
- Spurkriterium 191ff.
  
- stationäre Lösung (Partition)
  - , bei disk. Varianzkriterium 170, 172
  - , bei Extremalproblem 283
  - , bei kontinuierlichem Varianzkriterium 168, 287, 292
  - , sequentielle Bestimmung 284ff.
- Stereogramm 241
- Stirlingsche Zahlen  $S(N, m)$  109, 224, 369
- stochastische| Approximation 284ff.
  - Modelle (→ Modelle)
- Stopregel 410, 419
- Streuungsmatrix 138, 139
  - in Klassen 138, 185ff., 244
  - , totale 185ff.
  - zwischen Klassen 185ff., 243f.
- Stufe einer Klasse 362
- subdominierende Ultrametrik  $\delta^-$  372
  
- Teilgraph 301
- Transformation
  - auf unkorrelierte Merkmale 42
  - von Ähnlichkeiten und Distanzen 77
- Trainingsfolge 285
- Translationsinvarianz 37
- Trennfunktion (-fläche)
  - , lineare 168, 170, 189f.
  - , nichtlineare 200, 293ff.
- typischstes Objekt (→ Objekt)
  
- Überschneidung
  - bis  $k - 1$  Objekte 331, 423ff.
  - , kontrollierbare 336
  - -zahl 341, 424
- Unabhängigkeit
  - , empirische 59–63
  - , Test auf 61
- Unähnlichkeit (→ Distanz) 25
- Ultrametrik 366ff.
  - , maximale dominierte  $\delta^-$  372, 389
  - , Berechnung 373f.
  - , graphentheoretische Interpretation 374, 389
  - , minimale dominierende  $\delta^+$  377ff., 390, 395
  - , Konstruktion 395
  - , subdominierende 372
  - und Dendrogramm 365ff.
- Ungleichung
  - , Dreiecks- 26
  - , ultrametrische 366
  - , der Ordnung  $k$  428
- Untergraph 301

- , maximaler vollständiger 319
- unweighted-mean-pair-method 406
- Varianz** 433
  - , empirische 92, 362
  - in den Klassen 115, 157
  - , mittlere 192
  - zwischen den Klassen 116, 156
  - , Maximierung der 116, 162
- Varianzkriterium**
  - als Index 362
  - , diskretes 115, 122, 151, 154, 162–184, 220, 222, 232, 407, 412f.
  - , als ML-Verfahren 115
  - , Bayes-Optimalität 120, 122, 125, 127
  - , Monotonie 171
  - , optimale Partition 162ff.
  - , stationäre Lösung 170
  - , kontinuierliches 164, 178, 282, 287
  - , sequentielle Verfahren 287ff.
  - , stationäre Lösung 168, 287
- Vektor** 431f.
  - , komplementärer 48
- Vereinigung** 430
- Verfahren**
  - , agglomerative 371, 383–411
  - , Bayes-optimale 113–158
  - , dilatierendes 391
  - , divisive 371, 411–419
  - , monothetische 412, 416ff.
  - , polythetische 412ff.
  - , graphische 263
  - , heuristische, kombinierte 219, 223, 224–237, 332–338
  - , hierarchische 176, 356ff.
  - , iterative 171ff., 200, 222, 332, 351
  - , konservatives 391
  - , kontrahierendes 390
  - , optimales, zum mittleren Niveau  $1 - \alpha$  150f.
  - , rekursive 218, 224ff.
  - , sequentielle 235ff., 263, 279ff., 284
  - , Konvergenz 287, 290, 292
  - , Vergleich von 400, 409, 418
  - , Verhalten bei verschiedenen Punkt-konstellationen 234, 306
  - , Verhalten bei zufälligen Daten 400
  - von Braverman 289ff.
  - von Dorofeyuk 297
  - von Edwards-Cavalli-Sforza 176, 412
  - von Gower 413
  - von Lerman 374f.
  - von MacQueen 291
  - von Roux 373f., 429
  - von Schnell 272ff., 279, 420
  - von Ward 407
  - von Wishart 278ff., 306, 330, 420f.
- Vergleich von Verfahren 400, 409, 418
- Verkettung 300, 305f., 311, 376, 390, 406, 420
- Verlaufskurven 247
- Verlustfunktion**
  - , einfache 116, 130, 138, 143
  - , quadratische 122
- Verteilung**
  - , a-priori- 116f., 124, 130, 139, 143, 148
  - , Beta- 43
  - , Binomial- 52
  - ,  $\chi^2$ - (zentrale) 36, 84, 98, 434
  - ,  $\chi^2$ - (nichtzentrale) 36, 84
  - ,  $F$ - 43, 85, 86
  - , Gamma- 130
  - , inverse 130
  - , hypergeometrische 60f., 64
  - , Normal- 36, 43, 83ff., 89, 113ff., 234, 252ff.
  - , Polynomial- 251
  - , Wishart- 139, 143
- Verteilungsdichte**
  - , Analyse der 249
  - , multimodale 271
  - , nichtparametrische Schätzung 263–269, 272
  - , Konsistenz 265
  - , unimodale 270
  - -mischung 250
  - , Identifizierbarkeit 252ff.
  - , Zerlegung 250ff.
- Verzweigungsstufen 363
- Wahrscheinlichkeitsdistanz 64
- Weg in Graphen 301, 374
  - , kürzester 374
- weighted-mean-pair-method 248, 406
- Zentralitätsindex** 101
- Zentroid** 401
  - -methode 400ff., 405, 407
- Zusammenhangskomponenten**
  - eines Graphen 278f., 301ff., 309ff., 314, 376, 389, 421
  - , Gruppierung durch 298ff.
  - im  $R^p$  277