References

ALLBUS. (2006). *Germany general social survey* 1980–2004.

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*, 589–609.

Andrews, D. (1972). Plots of high-dimensional data. *Biometrics, 28*, 125–136.

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (1996). *Multivariate analysemethoden*. Berlin: Springer.

Bartlett, M. S. (1939). A note on tests of significance in multivariate analysis. *Proceedings of the Cambridge Philosophical Society, 35*, 180–185.

Bartlett, M. S. (1954). A note on multiplying factors for various chi-squared approximations. *Journal of the Royal Statistical Society: Series B, 16*, 296–298.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics*. New York: Wiley.

Blæsild, P., & Jensen, J. (1981). Multivariate distributions of hyperbolic type. In *Statistical distributions in scientific work—Proceedings of the NATO Advanced Study Institute held at the Università degli studi di Trieste* (Vol. 4, pp. 45–66).

Bock, R. D. (1975). *Multivariate statistical methods in behavioral research*. New York: Mc Graw-Hill.

Bouroche, J.-M., & Saporta, G. (1980). *L'analyse des données*. Paris: Presses Universitaires de France.

Breiman, L. (1973). *Statistics: With a view towards application*. Boston: Houghton Mifflin Company.

Breiman, L., Friedman, J. H., Olshen, R., & Stone, C. J. (1984). *Classification and regression trees*. Belmont: Wadsworth.

Chambers, J. M., Cleveland, W. S., Kleiner, B., & Tukey, P. A. (1983). *Graphical methods for data analysis*. Boston: Duxbury Press.

Chen, Y., Härdle, W., & Jeong, S.-O. (2008). Nonparametric risk management with generalized hyperbolic distributions. *Journal of the American Statistical Association, 103*, 910–923.

Chernoff, H. (1973). Using faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association, 68*, 361–368.

Cook, R. D., & Weisberg, S. (1991). Comment on "sliced inverse regression for dimension reduction". *Journal of the American Statistical Association, 86*(414), 328–332.

Dillon, W. R., & Goldstein, M. (1984). *Multivariate analysis*. New York: Wiley.

Duan, N., & Li, K.-C. (1991). Slicing regression: A link-free regression method. *Annals of Statistics, 19*(2), 505–530.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression (with discussion). *Annals of Statistics, 32*(2), 407–499.

© Springer-Verlag Berlin Heidelberg 2015

W.K. Härdle, L. Simar, *Applied Multivariate Statistical Analysis*,
DOI 10.1007/978-3-662-45171-7

Embrechts, P., McNeil, A. J., & Straumann, D. (1999, May). Correlation and dependence in risk management: Properties and pitfalls. *RISK Magazine*, 69–71.

EUNITE. (2001). *Electricity load forecast competition of the European network on intelligent technologies for smart adaptive systems*. http://neuron.tuke.sk/competition/

Everitt, B., & Dunn, G. (1998). *Applied multivariate data analysis*. London: Edward Arnold.

Fahrmeir, L., & Hamerle, A. (1984). *Multivariate statistische verfahren*. Berlin: De Gruyter.

Fang, K. T., Kotz, S., & Ng, K. W. (1990). *Symmetric multivariate and related distributions*. London: Chapman and Hall.

Fengler, M. R., Härdle, W., & Villa, C. (2003). The dynamics of implied volatilities: A common principal components approach. *Review of Derivative Research, 6*, 179–202.

Flury, B. (1988). *Common principle components analysis and related multivariate models*. New York: Wiley.

Flury, B., & Gautschi, W. (1986). An algorithm for simultaneous orthogonal transformation of several positive definite symmetric matrices to nearly diagonal form. *SIAM Journal on Scientific and Statistical Computing, 7*, 169–184.

Flury, B., & Riedwyl, H. (1988). *Multivariate statistics: A practical approach*. Cambridge: Cambridge University Press.

Franke, J., Härdle, W., & Hafner, C. (2011). *Introduction to statistics of financial markets* (3rd ed.). Heidelberg: Springer.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1–22.

Friedman, J. H., & Stuetzle, W. (1981). Projection pursuit classification. Unpublished manuscript.

Friedman, J. H., & Tukey, J. W. (1974). A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers, C 23*, 881–890.

Gale, D., Kuhn, H. W., & Tucker, A. W. (1951). Linear programming and the theory of games. In T. C. Koopmans (Ed.), *Activity analysis of production and allocation* (pp. 317–329). New York: John Wiley and Sons.

Gibbins, R. (1985). *Canonical analysis. A review with application in ecology*. Berlin: Springer.

Giri, N. C. (1996). *Multivariate statistical analysis*. New York: Marcel Dekker.

Gosset, W. S. (1908). The probable error of a mean. *Biometrika, 6*, 1–25.

Graham, M., & Kennedy, J. (2003). Using curves to enhance parallel coordinate visualisations. In *Proceedings of seventh international conference on information visualization, 2003 (IV 2003)* (pp. 10–16).

Hall, P. (1992). *The bootstrap and edgeworth expansion*. *Statistical Series*. New York: Springer.

Hall, P., & Li, K.-C. (1993). On almost linearity of low dimensional projections from high dimensional data. *Annals of Statistics, 21*(2), 867–889.

Härdle, W. (1991). *Smoothing techniques, with implementations in S*. New York: Springer.

Härdle, W., Hautsch, N., & Overbeck, L. (2009). *Applied quantitative finance* (2nd ed.). Heidelberg: Springer.

Härdle, W., Müller, M., Sperlich, S., & Werwatz, A. (2004). *Non- and semiparametric models*. Heidelberg: Springer.

Härdle, W., & Scott, D. (1992). Smoothing by weighted averaging of rounded points. *Computational Statistics, 7*, 97–128.

Harrison, D., & Rubinfeld, D. L. (1978). Hedonic prices and the demand for clean air. *Journal of Environment Economics & Management, 5*, 81–102.

Hoaglin, W., Mosteller, F., & Tukey, J. (1983). *Understanding robust and exploratory data analysis*. New York: Wiley.

Hodges, J. L., & Lehman, E. L. (1956). The efficiency of some non-parametric competitors of the t-test. *Annals of Mathematical Statistics, 27*, 324–335.

Hotelling, H. (1935). The most predictable criterion. *Journal of Educational Psychology, 26*, 139–142.

Hotelling, H. (1953). New light on the correlation coefficient and its transform. *Journal of the Royal Statistical Society, Series B, 15*, 193–232.

Huber, P. (1985). Projection pursuit. *Annals of Statistics, 13*(2), 435–475.

Inselberg, A. (1985). A goodness of fit test for binary regression models based on smoothing methods. *The Visual Computer, 1*, 69–91.

Johnson, R. A., & Wichern, D. W. (1998). *Applied multivariate analysis* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Jones, M. C., & Sibson, R. (1987). What is projection pursuit? (with discussion). *Journal of the Royal Statistical Society, Series A, 150*(1), 1–36.

Kaiser, H. F. (1985). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23*, 187–200.

Kendall, K., & Stuart, S. (1977). Distribution theory. In *The advanced theory of statistics* (Vol. 1). London: Griffin.

Klinke, S., & Polzehl, J. (1995). *Implementation of kernel based indices in XGobi*. Discussion paper 47, SFB 373. Humboldt-University of Berlin.

Kruskal, J. B. (1965). Analysis of factorial experiments by estimating a monotone transformation of data. *Journal of the Royal Statistical Society, Series B, 27*, 251–263.

Kruskal, J. B. (1969). Toward a practical method which helps uncover the structure of a set of observations by finding the line tranformation which optimizes a new "index of condensation". In R. C. Milton & J. A. Nelder (Eds.), *Statistical computation* (pp. 427–440). New York: Academic Press.

Kruskal, J. B. (1972). Linear transformation of multivariate data to reveal clustering. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioural sciences* (Vol. 1, pp. 179–191). London: Seminar Press.

Lachenbruch, P. A., & Mickey, M. R. (1968). Estimation of error rates in discriminant analysis. *Technometrics, 10*, 1–11.

Laplace, P.-S. (1774). Mémoire sur la probabilité des causes par les événements. *Savants étranges, 6*, 621–656.

Lawson, C., & Hansen, R. (1974). *Solving least square problems*. Englewood Cliffs: Prentice Hall.

Lebart, L., Morineau, A. and Fénelon, J. P. (1982). Traitement des données statistiques, Dunod, Paris.

Lewin-Koh, N. (2006). *Hexagon binnning*. Technical Report.

Li, K.-C. (1991). Sliced inverse regression for dimension reduction (with discussion). *Journal of the American Statistical Association, 86*(414), 316–342.

Li, K.-C. (1992). On principal Hessian directions for data visualization and dimension reduction: Another application of Stein's lemma. *Journal of the American Statistical Association, 87*, 1025–1039.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. Duluth/London: Academic Press.

Meier, L., van de Geer, S., & Bühlmann, P. (2008). The group lasso for logistic regression. *Journal of the Royal Statistical Society, Series B, 70*, 53–71.

Mercer, J. (1909). Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London, 209*, 415–446.

Morrison, D. (1990). *Multivariate statistical methods* (3rd ed., 495 pp.). New York: McGraw-Hill.

Muirhead, R. J. (1982). *Aspects of multivariate statistics*. New York: Wiley.

Nelsen, R. B. (1999). *An introduction to copulas*. New York: Springer.

Olkin, I., & Veath, M. (1980). Maximum likelihood estimation in a two-way analysis with correlated errors in one classification. *Biometrika, 68*, 653–660.

Osborne, M., Presnell, B., & Turlach, B. (2000). On the lasso and its dual. *Journal of Computational and Graphical Statistics, 9*(2), 319–337.

Parzen, E. (1962). On estimating of a probability density and mode. *Annals of Mathematical Statistics, 35*, 1065–1076.

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics, 27*, 832–837.

Schott, J. R. (1994). Determining the dimensionality in sliced inverse regression. *Journal of the American Statistical Association, 89*(425), 141–148.

Scott, D. (1985). Averaged shifted histograms: Effective nonparametric density estimation in several dimensions. *Annals of Statistics, 13*, 1024–1040.

SEC (2004). *Securities and exchange commission: Archive of historical documents*. http://www.sec.gov/cgi-bin/srch-edgar

Shevade, S. K., & Keerthi, S. S. (2003). A simple and efficient algorithm for gene selection using sparse logistic regression. *Bioinformatics, 19*, 2246–2253.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. *Monographs on statistics and applied probability* (Vol. 26). London: Chapman and Hall.

Simon, N., Friedman, J. H., Hastie, T., & Tibshirani, R. (2013). A sparse-group lasso. *Journal of Computational and Graphical Statistics, 22*(2), 231–245.

Sklar, A. (1959). Fonctions de répartition à *n* dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris, 8*, 229–231.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B, 58*, 267–288.

Tufte, E. (1983). *The visual display of quantitative information*. Cheshire: Graphics Press.

Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer.

Volle, V. (1985). *Analyse des Données*. Paris: Economica.

Whittle, P. (1958). On the smoothing of probability density functions. *Journal of the Royal Statistical Society, Series B, 55*, 549–557.

Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society, Series B, 68*, 49–67.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B, 67*, 301–320.

Index

Actual error rate (AER), 417
Admissible, 413
Agglomerative techniques, 393
Allocation rules, 407
Analysis of variance (ANOVA), 72
Andrews' Curves, 29
Angle between two vectors, 72
Apparent error rate (APER), 417

Bayes discriminant rule, 413
Bernoulli distributions, 144
Best line, 308
Binary structure, 387
Binomial sampling, 273
Biplots, 439
Bootstrap, 176
  sample, 179
Bootstrap sample, 179
Boston Housing, 40, 110, 242, 262, 346, 378, 400, 552
Boxplot, 4, 5
  construction, 7

Canonical correlation, 443
  analysis, 443
  coefficient, 445
  variable, 445
  vector, 445
Capital asset pricing model (CAPM), 497
Cauchy distribution, 154
Centering matrix, 90
Central limit theorem (CLT), 143–145
Centroid, 396
Characteristic functions, 123, 130
Classic blue pullovers, 82
Cluster algorithms, 392
Cluster analysis, 385
Cobb–Douglas production function, 254
Cochran theorem, 192
Coefficient of determination, 96, 107
  adjusted, 107
Column space, 74, 306
Common factors, 361
Common principal components analysis (CPCA), 342
Communality, 362
Complete linkage, 396
Computationally intensive techniques, 501
Concentration ellipsoid, 139
Conditional approximations, 188
Conditional covariance, 513
Conditional density, 119
Conditional distribution, 186
Conditional expectation, 126, 512, 513
Conditional pdf, 119
Confidence interval, 146
Confusion matrix, 417
Contingency tables, 264, 273, 425
Contrast, 227
Convex hull, 504
Copula, 120, 168
Copulae, 166
Correlation, 84
  multiple, 188
Correspondence analysis, 425
Covariance, 80
Covariance matrix
  decomposition, 320
  properties, 124

© Springer-Verlag Berlin Heidelberg 2015

W.K. Härdle, L. Simar, *Applied Multivariate Statistical Analysis*,
DOI 10.1007/978-3-662-45171-7

Cramer–Rao, 208
Cramer–Rao lower bound, 206
Cramer–Wold, 131
Cumulant, 132
Cumulative distribution function (cdf), 118
Curse of dimensionality, 511

Data depth, 504
Degrees of freedom, 102
Dendrogram, 394
Density estimates, 11
Density functions, 118
Determinant, 56
Deviance, 269
Diagonal matrix, 54
Dice, 388
Discriminant analysis, 407
Discriminant rule, 408
  in practice, 415
Dissimilarity of cars, 458
Distance
  d, 68
  Euclidean, 68
  iso-distance curves, 69
  matrix, 460
  measures, 389
Distribution, 118
Draftman's plot, 21
Duality relations, 313
Duality theorem, 464

Effective dimension reduction directions, 511, 513
Effective dimension reduction space, 511
Efficient portfolio, 488
Eigenvalues, 57
Eigenvectors, 57
Elastic net, 297
Elliptical distribution, 196
Elliptically symmetric distribution, 511
Existence of a riskless asset, 492
Expected cost of misclassification, 409
Explained variation, 96
Exploratory projection pursuit, 506
Extremes, 7

F-spread, 7
F-test, 104
Faces, 23
Factorial
  axis, 308, 309
  method, 336
  representation, 314, 317
  variable, 308, 316
Factors, 306
  analysis model, 359, 360
  model, 367
  scores, 376
Farthest Neighbor, 396
Fisher information, 209
Fisher information matrix, 207, 208
Fisher's linear discrimination function, 418
Five-number summary, 5
Flury faces, 24
Fourths, 6
French food expenditure, 340
Full model, 103

G-inverse, 57
  non-uniqueness, 61
General multinormal distribution, 193
Gradient, 65
Group-building algorithm, 386

Heavy-tailed distributions, 149
Hessian, 65
Hexagon, 37
  binning algorithm, 37, 50
  plot, 38
Hierarchical algorithm, 393
Histograms, 11
Hotelling T^2-distribution, 193
Hyperbolic, 151

Idempotent matrix, 54
Identity matrix, 54
Independence copula, 122
Independent, 85, 119
Inertia, 315, 317
Information matrix, 208
Interpretation of the factors, 363
Interpretation of the principal components, 327
Invariance of scale, 363
Inverse, 56
Inverse regression, 511, 513

Jaccard, 388
Jacobian, 135
Jordan decomposition, 60, 61

Kernel
  densities, 15
  estimator, 15
Kulczynski, 388

Laplace distribution, 154
Lasso, 282
Likelihood function, 202
Likelihood ratio test, 214
Limit theorems, 142
Linear discriminant analysis (LDA), 412
Linear regression, 93
Linear transformation, 91
Link function, 511
Loadings, 361, 362
  non-uniqueness, 364
Log-likelihood function, 202
Log-linear, 264
Logit models, 272

Mahalanobis distance, 412
Mahalanobis transformation, 93, 137, 138
Marginal densities, 119
Marketing strategies, 101
Maximum likelihood discriminant rule, 408
Maximum likelihood estimator, 202
MDS direction, 459
Mean-variance, 487, 488
Median, 5, 502
Metric methods, 459
Mixture model, 156
Model with interactions, 258
Moments, 123
Multidimentional scaling (MDS), 455
Multinormal, 139, 183
  distribution, 137
Multivariate deneralised hyperbolic distribution, 160
Multivariate laplace distribution, 163
Multivariate median, 504
Multivariate t-distribution, 163, 196

Nearest neighbor, 395
Non-metric solution, 479
Nonexistence of a riskless asset, 491
Nonhomogeneous, 92
Nonmetric methods of MDS, 459
Norm of a vector, 72
Normal distribution, 203
Normal-inverse Gaussian, 151
Normalized principal components (NPCs), 335
Null space, 74

Odds, 273
Order statistics, 5
Orthogonal complement, 75
Orthogonal matrix, 54
Orthonormed, 309
Outliers, 3
Outside bars, 7

Parallel profiles, 238
Partitioned covariance matrix, 184
Partitioned matrixes, 66
PAV algorithm, 466, 484
Pearson chi-square, 269
Pearson chi-square test for independence, 269
Pool-adjacent violators algorithm, 466, 484
Portfolio
  analysis, 487
  choice, 487
Positive
  definite, 62
  definiteness, 65
  or negative dependence, 22
  semidefinite, 62, 90
Principal
  axes, 70
  components, 324
    method, 372
    in practice, 324
    technique, 324
    transformation, 321, 323
  factors, 370
Principal components analysis (PCA), 512, 515
Profile
  analysis, 238
  method, 476
Projection
  matrix, 75
  pursuit, 505
  pursuit regression, 508
  vector, 511
Proximity between objects, 387
Proximity measure, 386
p-value, 269

Quadratic discriminant analysis, 414
Quadratic forms, 62
Quadratic response model, 253
Quality of the representations, 339

Randomized discriminant rule, 413
Rank, 55
Reduced model, 103
Rotations, 73, 374
Row space, 306
Russel and Rao (RR), 388

Sampling distributions, 142
Scatterplot matrix, 21
Separation line, 21
Similarity of objects, 387
Simple analysis of variance (ANOVA), 100
Simple Matching, 388, 389
Single linkage, 395
Singular normal distribution, 140
Singular value decomposition (SVD), 61, 313
Sliced inverse regression, 511, 516
  algorithm, 512
Sliced inverse regression II, 513, 514, 516, 517
  algorithm, 514
Solution
  nonmetric, 483
Specific factors, 361
Specific variance, 362
Spectral decompositions, 60
Spherical distribution, 195
Standardized linear combinations (SLC), 320
Statistics, 142
Stimulus, 475
Student's t with n, 152
Student's t-distribution, 94
Sum of squares, 102
Summary statistics, 89
Support vector machines (SVMs), 519
Swiss bank data, 4
Symmetric matrix, 54

T-test, 94
Tanimoto, 388
Three-way tables, 267
Total variation, 96
Trace, 56
Trade-off analysis, 476
Transformations, 135
Transpose, 56
Two factor method, 476

Unbiased estimator, 208
Uncorrelated factors, 361
Unexplained variation, 96
Unit vector, 72
Upper triangular matrix, 54

Variance explained by PCs, 333
Varimax
  criterion, 374
  method, 374
  rotation method, 374

Ward clustering, 396
Wishart
  density, 192
  distribution, 191, 192