

- Abraham, B., and Ledolter, J. (2006), *Introduction to Regression Modeling*, Thomson Brooks/Cole, Belmont, CA.
- Aggarwal, C.C. (2017), *Outlier Analysis*, 2nd ed., Springer, New York, NY.
- Agostinelli, C., Leung, A., Yohai, V., and Zamar, R. (2015), "Robust Estimation of Multivariate Location and Scatter in the Presence of Cellwise and Casewise Contamination," *Test*, 24, 441-461.
- Agresti, A. (2002), *Categorical Data Analysis*, 2nd ed., Wiley, Hoboken, NJ.
- Agresti, A. (2013), *Categorical Data Analysis*, 3rd ed., Wiley, Hoboken, NJ.
- Akaike, H. (1973), "Information Theory and an Extension of the Maximum Likelihood Principle," in *Proceedings, 2nd International Symposium on Information Theory*, eds. Petrov, B.N., and Csakim, F., Akademiai Kiado, Budapest, 267-281.
- Akaike, H. (1977), "On Entropy Maximization Principle," in *Applications of Statistics*, ed. Krishnaiah, P.R., North Holland, Amsterdam, 27-41.
- Akaike, H. (1978), "A New Look at the Bayes Procedure," *Biometrics*, 65, 53-59.
- Atkinson, A., and Riani, R. (2000), *Robust Diagnostic Regression Analysis*, Springer, New York, NY.
- Atkinson, A., Riani, R., and Cerioli, A. (2004), *Exploring Multivariate Data with the Forward Search*, Springer, New York, NY.
- Bartelsmeyer, C. (2017), "Prediction Intervals for Lasso and Relaxed Lasso Using d Variables," Master's Research Paper, Southern Illinois University.
- Becker, R.A., Chambers, J.M., and Wilks, A.R. (1988), *The New S Language: a Programming Environment for Data Analysis and Graphics*, Wadsworth and Brooks/Cole, Pacific Grove, CA.
- Belsley, D.A. (1984), "Demeaning Conditioning Diagnostics Through Centering," *The American Statistician*, 38, 73-77.
- Belsley, D.A., Kuh, E., and Welsch, R.E. (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, Wiley, New York, NY.
- Berk, R., Brown, L., Buja, A., Zhang, K., and Zhao, L. (2013), "Valid Post-Selection Inference," *The Annals of Statistics*, 41, 802-837.
- Berk, R.A. (2016), *Statistical Learning From a Regression Perspective*, 2nd ed., Springer, New York, NY.
- Bertsimas, D., King, A., and Mazmunder, R. (2016), "Best Subset Selection Via a Modern Optimization Lens," *The Annals of Statistics*, 44, 813-852.
- Bhatia, R., Elsner, L., and Krause, G. (1990), "Bounds for the Variation of the Roots of a Polynomial and the Eigenvalues of a Matrix," *Linear Algebra and Its Applications*, 142, 195-209.
- Bickel, P.J., and Freedman, D.A. (1981), "Some Asymptotic Theory for the Bootstrap," *The Annals of Statistics*, 1196-1217.

- Bickel, P.J., and Ren, J.-J. (2001), "The Bootstrap in Hypothesis Testing," in *State of the Art in Probability and Statistics: Festschrift for William R. van Zwet*, eds. de Gunst, M., Klaassen, C., and van der Vaart, A., The Institute of Mathematical Statistics, Hayward, CA, 91-112.
- Bickel, P.J., Götze, F., and Van Zwet, W.R. (1997), "Resampling Fewer Than n Observations: Gains, Losses and Remedies for Losses," *Statistica Sinica*, 7, 1-31.
- Bogdan, M., Ghosh, J., and Doerge, R. (2004), "Modifying the Schwarz Bayesian Information Criteria to Locate Multiple Interacting Quantitative Trait Loci," *Genetics*, 167, 989-999.
- Boudt, K., Rousseeuw, P.J., Vanduffel, S., and Verdonck, T. (2017), "The Minimum Regularized Covariance Determinant Estimator," unpublished document at (<https://arxiv.org/pdf/1701.07086.pdf>).
- Box, G.E.P., and Cox, D.R. (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society, B*, 26, 211-246.
- Breiman, L. (1996), "Bagging Predictors," *Machine Learning*, 24, 123-140.
- Brillinger, D.R. (1977), "The Identification of a Particular Nonlinear Time Series," *Biometrika*, 64, 509-515.
- Brillinger, D.R. (1983), "A Generalized Linear Model with "Gaussian" Regressor Variables," in *A Festschrift for Erich L. Lehmann*, eds. Bickel, P.J., Doksum, K.A., and Hodges, J.L., Wadsworth, Pacific Grove, CA, 97-114.
- Bühlmann, P., and Yu, B. (2002), "Analyzing Bagging," *The Annals of Statistics*, 30, 927-961.
- Buckland, S.T., Burnham, K.P., and Augustin, N.H. (1997), "Model Selection: an Integral Part of Inference," *Biometrics*, 53, 603-618.
- Budny, K. (2014), "A Generalization of Chebyshev's Inequality for Hilbert-Space-Valued Random Variables," *Statistics & Probability Letters*, 88, 62-65.
- Burnham, K.P., and Anderson, D.R. (2004), "Multimodel Inference Understanding AIC and BIC in Model Selection," *Sociological Methods & Research*, 33, 261-304.
- Burr, D. (1994), "A Comparison of Certain Bootstrap Confidence Intervals in the Cox Model," *Journal of the American Statistical Association*, 89, 1290-1302.
- Butler, R., and Rothman, E. (1980), "Predictive Intervals Based on Reuse of the Sample," *Journal of the American Statistical Association*, 75, 881-889.
- Buxton, L.H.D. (1920), "The Anthropology of Cyprus," *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 50, 183-235.
- Cai, T., and Liu, W. (2011), "A Direct Approach to Sparse Linear Discriminant Analysis," *Journal of the American Statistical Association*, 106, 1566-1577.
- Cai, T., Liu, W., and Luo, X. (2011), "A Constrained l_1 Minimization Approach to Sparse Precision Matrix Estimation," *Journal of the American Statistical Association*, 106, 594-607.

- Cai, T., Tian, L., Solomon, S.D., and Wei, L.J. (2008), "Predicting Future Responses Based on Possibly Misspecified Working Models," *Biometrika*, 95, 75-92.
- Cameron, A.C., and Trivedi, P.K. (1998), *Regression Analysis of Count Data*, 1st ed., Cambridge University Press, Cambridge, UK.
- Cameron, A.C., and Trivedi, P.K. (2013), *Regression Analysis of Count Data*, 2nd ed., Cambridge University Press, Cambridge, UK.
- Camponovo, L. (2015), "On the Validity of the Pairs Bootstrap for Lasso Estimators," *Biometrika*, 102, 981-987.
- Candes, E., and Tao, T. (2007), "The Dantzig Selector: Statistical Estimation When p Is Much Larger Than n ," *The Annals of Statistics*, 35, 2313-2351.
- Chang, J., and Hall, P. (2015), "Double Bootstrap Methods That Use a Single Double-Bootstrap Simulation," *Biometrika*, 102, 203-214.
- Chang, J., and Olive, D.J. (2010), "OLS for 1D Regression Models," *Communications in Statistics: Theory and Methods*, 39, 1869-1882.
- Chao, S.-K., Ning, Y., and Liu, H. (2014), "On High Dimensional Post-Regularization Prediction Intervals," unpublished manuscript at (http://www.stat.purdue.edu/~skchao74/HD_PCI.pdf).
- Charkhi, A., and Claeskens, G. (2018), "Asymptotic Post-Selection Inference for the Akaike Information Criterion," *Biometrika*, 105, 645-664.
- Chatterjee, A., and Lahiri, S.N. (2011), "Bootstrapping Lasso Estimators," *Journal of the American Statistical Association*, 106, 608-625.
- Chen, J., and Chen, Z. (2008), "Extended Bayesian Information Criterion for Model Selection With Large Model Spaces," *Biometrika*, 95, 759-771.
- Chen, S.X. (2016), "Peter Hall's Contributions to the Bootstrap," *The Annals of Statistics*, 44, 1821-1836.
- Chen, X. (2011), "A New Generalization of Chebyshev Inequality for Random Vectors," see arXiv:0707.0805v2.
- Chew, V. (1966), "Confidence, Prediction and Tolerance Regions for the Multivariate Normal Distribution," *Journal of the American Statistical Association*, 61, 605-617.
- Chihara, L., and Hesterberg, T. (2011), *Mathematical Statistics with Resampling and R*, Hoboken, NJ: Wiley.
- Cho, H., and Fryzlewicz, P. (2012), "High Dimensional Variable Selection Via Tilting," *Journal of the Royal Statistical Society, B*, 74, 593-622.
- Chun, H., and Keleş, S. (2010), "Sparse Partial Least Squares Regression for Simultaneous Dimension Reduction and Predictor Selection," *Journal of the Royal Statistical Society, B*, 72, 3-25.
- Claeskens, G., and Hjort, N.L. (2008), *Model Selection and Model Averaging*, Cambridge University Press, New York, NY.
- Clarke, B.R. (1986), "Nonsmooth Analysis and Fréchet Differentiability of M Functionals," *Probability Theory and Related Fields*, 73, 137-209.
- Clarke, B.R. (2000), "A Review of Differentiability in Relation to Robustness With an Application to Seismic Data Analysis," *Proceedings of the Indian National Science Academy, A*, 66, 467-482.

- Cleveland, W. (1979), "Robust Locally Weighted Regression and Smoothing Scatterplots," *Journal of the American Statistical Association*, 74, 829-836.
- Cleveland, W.S. (1981), "LOWESS: a Program for Smoothing Scatterplots by Robust Locally Weighted Regression," *The American Statistician*, 35, 54.
- Collett, D. (1999), *Modelling Binary Data*, 1st ed., Chapman & Hall/CRC, Boca Raton, FL.
- Collett, D. (2003), *Modelling Binary Data*, 2nd ed., Chapman & Hall/CRC, Boca Raton, FL.
- Cook, R.D. (2018), *An Introduction to Envelopes: Dimension Reduction for Efficient Estimation in Multivariate Statistics*, Wiley, Hoboken, NJ.
- Cook, R.D., and Forzani, L. (2008), "Principal Fitted Components for Dimension Reduction in Regression," *Statistical Science*, 23, 485-501.
- Cook, R.D., and Forzani, L. (2018), "Big Data and Partial Least Squares Prediction," *The Canadian Journal of Statistics*, 46, 62-78.
- Cook, R.D., and Forzani, L. (2019), "Partial Least Squares Prediction in High-Dimensional Regression," *The Annals of Statistics*, 47, 884-908.
- Cook, R.D., Forzani, L., and Rothman, A. (2013), "Prediction in Abundant High-Dimensional Linear Regression," *Electronic Journal of Statistics*, 7, 30593088.
- Cook, R.D., Helland, I.S., and Su, Z. (2013), "Envelopes and Partial Least Squares Regression," *Journal of the Royal Statistical Society, B*, 75, 851-877.
- Cook, R.D., and Olive, D.J. (2001), "A Note on Visualizing Response Transformations in Regression," *Technometrics*, 43, 443-449.
- Cook, R.D., and Su, Z. (2016), "Scaled Predictor Envelopes and Partial Least-Squares Regression," *Technometrics*, 58, 155-165.
- Cook, R.D., and Weisberg, S. (1997), "Graphics for Assessing the Adequacy of Regression Models," *Journal of the American Statistical Association*, 92, 490-499.
- Cook, R.D., and Weisberg, S. (1999), *Applied Regression Including Computing and Graphics*, Wiley, New York, NY.
- Cook, R.D., and Zhang, X. (2015), "Foundations of Envelope Models and Methods," *Journal of the American Statistical Association*, 110, 599-611.
- Cox, D.R. (1972), "Regression Models and Life-Tables," *Journal of the Royal Statistical Society, B*, 34, 187-220.
- Cramér, H. (1946), *Mathematical Methods of Statistics*, Princeton University Press, Princeton, NJ.
- Crawley, M.J. (2005), *Statistics an Introduction Using R*, Wiley, Hoboken, NJ.
- Crawley, M.J. (2013), *The R Book*, 2nd ed., Wiley, Hoboken, NJ.
- Croux, C., Filzmoser, P., and Fritz, H. (2013), "Robust Sparse Principal Component Analysis," *Technometrics*, 55, 202-214.
- Croux, C. and Öllerer, V. (2016), "Robust and Sparse Estimation of the Inverse Covariance Matrix Using Rank Correlation Measures," in *Recent Advances in Robust Statistics: Theory and Applications*, eds. Agostinelli, C.,

- Basu, A., Filzomer, P., and Mukherjee, D., Springer, New Delhi, India, 35-56.
- Daniel, C., and Wood, F.S. (1980), *Fitting Equations to Data*, 2nd ed., Wiley, New York, NY.
- Datta, B.N. (1995), *Numerical Linear Algebra and Applications*, Brooks/Cole Publishing Company, Pacific Grove, CA.
- Davison, A.C., and Hinkley, D.V. (1997), *Bootstrap Methods and their Application*, Cambridge University Press, Cambridge, UK.
- Denham, M.C. (1997), "Prediction Intervals in Partial Least Squares," *Journal of Chemometrics*, 11, 39-52.
- Devroye, L., and Wagner, T.J. (1982), "Nearest Neighbor Methods in Discrimination," in *Handbook of Statistics*, Vol. 2, eds. Krishnaiah, P.R., and Kanai, L.N., North Holland, Amsterdam, 193-197.
- Dezeure, R., Bühlmann, P., Meier, L., and Meinshausen, N. (2015), "High-Dimensional Inference: Confidence Intervals, p -Values and R-Software hdi," *Statistical Science*, 30, 533-558.
- Efron, B. (1979), "Bootstrap Methods, Another Look at the Jackknife," *The Annals of Statistics*, 7, 1-26.
- Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, SIAM, Philadelphia, PA.
- Efron, B. (2014), "Estimation and Accuracy After Model Selection," (with discussion), *Journal of the American Statistical Association*, 109, 991-1007.
- Efron, B., and Hastie, T. (2016), *Computer Age Statistical Inference*, Cambridge University Press, New York, NY.
- Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004), "Least Angle Regression," (with discussion), *The Annals of Statistics*, 32, 407-451.
- Efron, B., and Tibshirani, R.J. (1993), *An Introduction to the Bootstrap*, Chapman & Hall/CRC, New York, NY.
- Efroymson, M.A. (1960), "Multiple Regression Analysis," in *Mathematical Methods for Digital Computers*, eds. Ralston, A., and Wilf, H.S., Wiley, New York, New York, 191-203.
- Ein-Dor, P., and Feldmesser, J. (1987), "Attributes of the Performance of Central Processing Units: a Relative Performance Prediction Model," *Communications of the ACM*, 30, 308317.
- Ewald, K., and Schneider, U. (2018), "Uniformly Valid Confidence Sets Based on the Lasso," *Electronic Journal of Statistics*, 12, 1358-1387.
- Fahrmeir, L. and Tutz, G. (2001), *Multivariate Statistical Modelling Based on Generalized Linear Models*, 2nd ed., Springer, New York, NY.
- Fan, J., and Li, R. (2001), "Variable Selection Via Noncave Penalized Likelihood and Its Oracle Properties," *Journal of the American Statistical Association*, 96, 1348-1360.
- Fan, J., and Lv, J. (2010), "A Selective Overview of Variable Selection in High Dimensional Feature Space," *Statistica Sinica*, 20, 101-148.
- Fan, R. (2016), "A Squared Correlation Coefficient of the Correlation Matrix," unpublished manuscript at (<http://parker.ad.siu.edu/Olive/sfan.pdf>).

- Farcomeni, A., and Greco, L. (2015), *Robust Methods for Data Reduction*, Chapman & Hall/CRC, Boca Rotan, FL.
- Feng, L., and Sun, F. (2015), "A Note on High-Dimensional Two-Sample Test," *Statistics & Probability Letters*, 105, 29-36.
- Feng, X., and He, X. (2014), "Statistical Inference Based on Robust Low-Rank Data Matrix Approximation," *The Annals of Statistics*, 42, 190-210.
- Ferguson, T.S. (1996), *A Course in Large Sample Theory*, Chapman & Hall, New York, NY.
- Fernholtz, L.T. (1983), *von Mises Calculus for Statistical Functionals*, Springer, New York, NY.
- Ferrari, D., and Yang, Y. (2015), "Confidence Sets for Model Selection by F -Testing," *Statistica Sinica*, 25, 1637-1658.
- Filzmoser, P., Joossens, K., and Croux, C. (2006), "Multiple Group Linear Discriminant Analysis: Robustness and Error Rate," *Compstat 2006: Proceedings in Computational Statistics*, eds. Rizzi, A., and Vichi, M., Physica-Verlag, Heidelberg, 521-532.
- Fithian, W., Sun, D., and Taylor, J. (2014), "Optimal Inference After Model Selection," ArXiv e-prints.
- Flury, B., and Riedwyl, H. (1988), *Multivariate Statistics: a Practical Approach*, Chapman & Hall, New York.
- Fogel, P., Hawkins, D.M., Beecher, C., Luta, G., and Young, S. (2013), "A Tale of Two Matrix Factorizations," *The American Statistician*, 67, 207-218.
- Fox, J., and Weisberg, S. (2019), *An R Companion to Applied Regression*, 3rd ed., Sage Publications, Thousand Oaks, CA.
- Frank, I.E., and Friedman, J.H. (1993), "A Statistical View of Some Chemometrics Regression Tools," (with discussion), *Technometrics*, 35, 109-148.
- Freedman, D.A. (1981), "Bootstrapping Regression Models," *The Annals of Statistics*, 9, 1218-1228.
- Frey, J. (2013), "Data-Driven Nonparametric Prediction Intervals," *Journal of Statistical Planning and Inference*, 143, 1039-1048.
- Friedman, J., Hastie, T., Hoeffling, H., and Tibshirani, R. (2007), "Pathwise Coordinate Optimization," *Annals of Applied Statistics*, 1, 302-332.
- Friedman, J., Hastie, T., Simon, N., and Tibshirani, R. (2015), *glmnet: Lasso and Elastic-net Regularized Generalized Linear Models*, R Package version 2.0, (<http://cran.r-project.org/package=glmnet>).
- Friedman, J., Hastie, T., and Tibshirani, R. (2008), "Sparse Inverse Covariance Estimation with the Graphical Lasso," *Biostatistics*, 9, 432-441.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010), "Regularization Paths for Generalized Linear Models Via Coordinate Descent," *Journal of Statistical Software*, 33, 1-22.
- Friedman, J.H. (1989), "Regularized Discriminant Analysis," *Journal of the American Statistical Association*, 84, 165-175.
- Friedman, J.H., and Hall, P. (2007), "On Bagging and Nonlinear Estimation," *Journal of Statistical Planning and Inference*, 137, 669-683.

- Fujikoshi, Y., Sakurai, T., and Yanagihara, H. (2014), "Consistency of High-Dimensional AIC-Type and C_p -Type Criteria in Multivariate Linear Regression," *Journal of Multivariate Analysis*, 123, 184-200.
- Fujikoshi, Y., Ulyanov, V.V., and Shimizu, R. (2010), *Multivariate Statistics: High-Dimensional and Large-Sample Approximations*, Wiley, Hoboken, NJ.
- Gao, X., and Huang, J. (2010), "Asymptotic Analysis of High-Dimensional LAD Regression with Lasso," *Statistica Sinica*, 20, 1485-1506.
- García-Escudero, L.A., Gordaliza, A., Matrán, C., and Mayo-Isar, A. (2010), "A Review of Robust Clustering Methods," *Advances in Data Analysis and Clustering*, 4, 89-109.
- Ghosh, S., and Polansky, A.M. (2014), "Smoothed and Iterated Bootstrap Confidence Regions for Parameter Vectors," *Journal of Multivariate Analysis*, 132, 171-182.
- Gill, R.D. (1989), "Non- and Semi-Parametric Maximum Likelihood Estimators and the von Mises Method, Part 1," *Scandinavian Journal of Statistics*, 16, 97-128.
- Gladstone, R.J. (1905), "A Study of the Relations of the Brain to the Size of the Head," *Biometrika*, 4, 105-123.
- Graybill, F.A. (1983), *Matrices With Applications to Statistics*, 2nd ed., Wadsworth, Belmont, CA.
- Gregory, K.B., Carroll, R.J., Baladandayuthapani, V., and Lahari, S.N. (2015), "A Two-Sample Test for Equality of Means in High Dimension," *Journal of the American Statistical Association*, 110, 837-849.
- Grübel, R. (1988), "The Length of the Shorth," *The Annals of Statistics*, 16, 619-628.
- Gruber, M.H.J. (1998), *Improving Efficiency by Shrinkage: the James-Stein and Ridge Regression Estimators*, Marcel Dekker, New York, NY.
- Gunst, R.F., and Mason, R.L. (1980), *Regression Analysis and Its Application*, Marcel Dekker, New York, NY.
- Guo, C., Yang, H., and Lv, J. (2017), "Robust Variable Selection for Generalized Linear Models With a Diverging Number of Parameters," *Communications in Statistics: Theory and Methods*, 46, 2967-2981.
- Guttman, I. (1982), *Linear Models: an Introduction*, Wiley, New York, NY.
- Haggstrom, G.W. (1983), "Logistic Regression and Discriminant Analysis by Ordinary Least Squares," *Journal of Business & Economic Statistics*, 1, 229-238.
- Haile, M. (2017), "Prediction Intervals After Forward Selection Using EBIC," Master's Research Paper, Southern Illinois University.
- Haitovsky, Y. (1987), "On Multivariate Ridge Regression," *Biometrika*, 74, 563-570.
- Hall, P. (1986), "On the Bootstrap and Confidence Intervals," *The Annals of Statistics*, 14, 1431-1452.

- Hall, P. (1988), "Theoretical Comparisons of Bootstrap Confidence Intervals," (with discussion), *The Annals of Statistics*, 16, 927-985.
- Hall, P., Lee, E.R., and Park, B.U. (2009), "Bootstrap-Based Penalty Choice for the Lasso Achieving Oracle Performance," *Statistica Sinica*, 19, 449-471.
- Hall, P., Martin, M.A., and Schucany, W.R. (1989), "Better Nonparametric Bootstrap Confidence Intervals for the Correlation Coefficient," *Journal of Statistical Computation and Simulation*, 33, 161-172.
- Hand, D.J. (2006), "Classifier Technology and the Illusion of Progress," (with discussion), *Statistical Science*, 21, 1-34.
- Hastie, T.J., and Tibshirani, R.J. (1990), *Generalized Additive Models*, Chapman & Hall, London, UK.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, New York, NY.
- Hastie, T., Tibshirani, R., and Wainwright, M. (2015), *Statistical Learning with Sparsity: the Lasso and Generalizations*, CRC Press Taylor & Francis, Boca Raton, FL.
- Haughton, D.M.A. (1988), "On the Choice of a Model to Fit Data From an Exponential Family," *The Annals of Statistics*, 16, 342-355.
- Haughton, D. (1989), "Size of the Error in the Choice of a Model to Fit Data From an Exponential Family," *Sankhyā, A*, 51, 45-58.
- Hebbler, B. (1847), "Statistics of Prussia," *Journal of the Royal Statistical Society, A*, 10, 154-186.
- Hesterberg, T., (2014), "What Teachers Should Know About the Bootstrap: Resampling in the Undergraduate Statistics Curriculum," available from (<http://arxiv.org/pdf/1411.5279v1.pdf>). (An abbreviated version was published (2015), *The American Statistician*, 69, 371-386.)
- Hilbe, J.M. (2011), *Negative Binomial Regression*, Cambridge University Press, 2nd ed., Cambridge, UK.
- Hillis, S.L., and Davis, C.S. (1994), "A Simple Justification of the Iterative Fitting Procedure for Generalized Linear Models," *The American Statistician*, 48, 288-289.
- Hjort, N.L., and Claeskins, G. (2003), "Frequentist Model Average Estimators," *Journal of the American Statistical Association*, 98, 879-899.
- Hoerl, A.E., and Kennard, R. (1970), "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, 12, 55-67.
- Hoffman, I., Serneels, S., Filzmoser, P., and Croux, C. (2015), "Sparse Partial Robust M Regression," *Chemometrics and Intelligent Laboratory Systems*, 149, Part A, 50-59.
- Hogg, R.V., and Tanis, E.A. (2005), *Probability and Statistical Inference*, 7th ed., Prentice Hall, Englewood Cliffs, NJ.
- Hong, L., Kuffner, T.A., and Martin, R. (2018), "On Overfitting and Post-Selection Uncertainty Assessments," *Biometrika*, 105, 221-224.

- Hosmer, D.W., and Lemeshow, S. (2000), *Applied Logistic Regression*, 2nd ed., Wiley, New York, NY.
- Hsieh, C.J., Sustik, M.A., Dhillon, I.S., and Ravikumar, P. (2011), "Sparse Inverse Covariance Matrix Estimation Using Quadratic Approximation," *Advances in Neural Information Processing Systems*, 24, 2330-2338.
- Huang, J., Liu, N., Pourahmadi, M., and Liu, L. (2006), "Covariance Matrix Selection and Estimation via Penalised Normal Likelihood," *Biometrika*, 93, 85-98.
- Huberty, C.J., and Olejnik, S. (2006), *Applied MANOVA and Discriminant Analysis*, 2nd ed., Wiley, Hoboken, NJ.
- Hurvich, C., and Tsai, C.L. (1989), "Regression and Time Series Model Selection in Small Samples," *Biometrika*, 76, 297-307.
- Hurvich, C., and Tsai, C.L. (1990), "The Impact of Model Selection on Inference in Linear Regression," *The American Statistician*, 44, 214-217.
- Hurvich, C.M., and Tsai, C.-L. (1991), "Bias of the Corrected AIC Criterion for Underfitted Regression and Time Series Models," *Biometrika*, 78, 499-509.
- Hyodo, M., and Nishiyama, T. (2017), "A One-Sample Location Test Based on Weighted Averaging of Two Test Statistics When the Dimension and the Sample Size are Large," *Communications in Statistics: Theory and Methods*, 46, 3526-3541.
- Hyndman, R.J. (1996), "Computing and Graphing Highest Density Regions," *The American Statistician*, 50, 120-126.
- Imhoff, D.C. (2018), "Bootstrapping Forward Selection with C_p ," Master's Research Paper, Southern Illinois University.
- Izenman, A.J. (2008), *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*, Springer, New York, NY.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013), *An Introduction to Statistical Learning With Applications in R*, Springer, New York, NY.
- Jansen, L., Fithian, W., and Hastie, T. (2015), "Effective Degrees of Freedom: a Flawed Metaphor," *Biometrika*, 102, 479-485.
- Javanmard, A., and Montanari, A. (2014), "Confidence Intervals and Hypothesis Testing for High-Dimensional Regression," *Journal of Machine Learning Research*, 15, 2869-2909.
- Jia, J., and Yu, B. (2010), "On Model Selection Consistency of the Elastic Net When $p \gg n$," *Statistica Sinica*, 20, 595-611.
- Jin, J., and Wang, W. (2016), "Influential Features PCA for High Dimensional Clustering," *The Annals of Statistics*, 44, 2323-2359.
- Johnson, M.E. (1987), *Multivariate Statistical Simulation*, Wiley, New York, NY.
- Johnson, M.P., and Raven, P.H. (1973), "Species Number and Endemism, the Galápagos Archipelago Revisited," *Science*, 179, 893-895.
- Johnson, R.A., and Wichern, D.W. (1988), *Applied Multivariate Statistical Analysis*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ.

- Johnson, R.A., and Wichern, D.W. (2007), *Applied Multivariate Statistical Analysis*, 6th ed., Prentice Hall, Englewood Cliffs, NJ.
- Johnstone, I.M., and Lu, A.Y. (2009), "On Consistency and Sparsity for Principal Component Analysis in High Dimension," (with discussion), *Journal of the American Statistical Association*, 104, 682-703.
- Jolliffe, I.T. (1983), "A Note on the Use of Principal Components in Regression," *Applied Statistics*, 31, 300-303.
- Jones, H.L. (1946), "Linear Regression Functions with Neglected Variables," *Journal of the American Statistical Association*, 41, 356-369.
- Kaufman, L., and Rousseeuw, P.J. (1990), *Finding Groups in Data: an Introduction to Cluster Analysis*, Wiley, New York, NY.
- Kay, R., and Little, S. (1987), "Transformations of the Explanatory Variables in the Logistic Regression Model for Binary Data," *Biometrika*, 74, 495-501.
- Kim, S. (2017), "Prediction Intervals for Partial Least Squares and Principal Component Regression Using d Variables," Master's Research Paper, Southern Illinois University.
- Kim, Y., Kwon, S., and Choi, H. (2012), "Consistent Model Selection Criteria on High Dimensions," *Journal of Machine Learning Research*, 13, 1037-1057.
- Knight, K., and Fu, W.J. (2000), "Asymptotics for Lasso-Type Estimators," *The Annals of Statistics*, 28, 1356-1378.
- Koch, I. (2014), *Analysis of Multivariate and High-Dimensional Data*, Cambridge University Press, New York, NY.
- Kuhn, M., and Johnson, K. (2013), *Applied Predictive Modeling*, Springer, New York, NY.
- Lai, T.L., Robbins, H., Wei, C.Z. (1979), "Strong Consistency of Least Squares Estimates in Multiple Regression II," *Journal of Multivariate Analysis*, 9, 343-361.
- Larsen, R.J., and Marx, M.L. (2011), *Introduction to Mathematical Statistics and Its Applications*, 5th ed., Prentice Hall, Upper Saddle River, NJ.
- Ledoit, O., and Wolf, M. (2004), "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices," *Journal of Multivariate Analysis*, 88, 365-411.
- Lee, J., Sun, D., Sun, Y., and Taylor, J. (2016), "Exact Post-Selection Inference With Application to the Lasso," *The Annals of Statistics*, 44, 907-927.
- Lee, J.D., and Taylor, J.E. (2014), "Exact Post Model Selection Inference for Marginal Screening," in *Advances in Neural Information Processing Systems*, 136-144.
- Leeb, H., and Pötscher, B.M. (2003), "The Finite-sample Distribution of Post-Model Selection Estimators and Uniform Versus Nonuniform Approximations," *Econometric Theory*, 19, 100-142.
- Leeb, H., and Pötscher, B.M. (2005), "Model Selection and Inference: Facts and Fiction," *Econometric Theory*, 21, 21-59.

- Leeb, H., and Pötscher, B.M. (2006), “Can One Estimate the Conditional Distribution of Post-Model-Selection Estimators?” *The Annals of Statistics*, 34, 2554-2591.
- Leeb, H. and Pötscher, B.M. (2008), “Can One Estimate the Unconditional Distribution of Post-Model-Selection Estimators?” *Econometric Theory*, 24, 338-376.
- Leeb, H., Pötscher, B.M., and Ewald, K. (2015), “On Various Confidence Intervals Post-Model-Selection,” *Statistical Science*, 30, 216-227.
- Lehmann, E.L. (1999), *Elements of Large-Sample Theory*, Springer, New York, NY.
- Lei, J., G’Sell, M., Rinaldo, A., Tibshirani, R.J., and Wasserman, L. (2018), “Distribution-Free Predictive Inference for Regression,” *Journal of the American Statistical Association*, 113, 1094-1111.
- Lesnoff, M., and Lancelot, R. (2010), “aod: Analysis of Overdispersed Data,” R package version 1.2, (<http://cran.r-project.org/package=aod>).
- Li, K.-C. (1987), “Asymptotic Optimality for C_p , C_L , Cross-Validation and Generalized Cross-Validation: Discrete Index Set,” *The Annals of Statistics*, 15, 958-975.
- Lin, D., Foster, D.P., and Ungar, L.H. (2012), “VIF Regression, a Fast Regression Algorithm for Large Data,” *Journal of the American Statistical Association*, 106, 232-247.
- Lindenmayer, D.B., Cunningham, R., Tanton, M.T., Nix, H.A., and Smith, A.P. (1991), “The Conservation of Arboreal Marsupials in the Mountain Ash-forests of Central Highlands of Victoria, South-East Australia: III. The Habitat Requirement’s of Leadbeater’s Possum *Gymnobelideus Leadbeateri* and Models of the Diversity and Abundance of Arboreal Marsupials,” *Biological Conservation*, 56, 295-315.
- Liu, L., Hawkins, D.M., Ghosh, S., and Young, S.S. (2003), “Robust Singular Value Decomposition Analysis of Microarray Data,” *Proceedings of the National Academy of Sciences*, 100, 13167-13172.
- Lockhart, R., Taylor, J., Tibshirani, R.J., and Tibshirani, R. (2014), “A Significance Test for the Lasso,” (with discussion), *The Annals of Statistics*, 42, 413-468.
- Lu, S., Liu, Y., Yin, L., and Zhang, K. (2017), “Confidence Intervals and Regions for the Lasso by Using Stochastic Variational Inequality Techniques in Optimization,” *Journal of the Royal Statistical Society, B*, 79 589-611.
- Lumley, T. (using Fortran code by Alan Miller) (2009), *leaps: Regression Subset Selection*, R package version 2.9, (<https://CRAN.R-project.org/package=leaps>).
- Luo, S., and Chen, Z. (2013), “Extended BIC for Linear Regression Models With Diverging Number of Relevant Features and High or Ultra-High Feature Spaces,” *Journal of Statistical Planning and Inference*, 143, 494-504.
- Machado, J.A.F., and Parente, P. (2005), “Bootstrap Estimation of Covariance Matrices Via the Percentile Method,” *Econometrics Journal*, 8, 70-78.

- Mai, Q., and Zou, H. (2013), "A Note on the Connection and Equivalence of Three Sparse Linear Discriminant Analysis Methods," *Technometrics*, 55, 243-246.
- Mai, Q., Zou, H., and Yuan, M. (2012), "A Direct Approach to Sparse Discriminant Analysis in Ultra-High Dimensions," *Biometrika*, 99, 29-42.
- Mallows, C. (1973), "Some Comments on C_p ," *Technometrics*, 15, 661-676.
- Marden, J.I. (2006), *Notes on Statistical Learning*, unpublished notes online at (www.stat.istics.net).
- Mardia, K.V., Kent, J.T., and Bibby, J.M. (1979), *Multivariate Analysis*, Academic Press, London, UK.
- Marquardt, D.W., and Snee, R.D. (1975), "Ridge Regression in Practice," *The American Statistician*, 29, 3-20.
- MathSoft (1999a), *S-Plus 2000 User's Guide*, Data Analysis Products Division, MathSoft, Seattle, WA.
- MathSoft (1999b), *S-Plus 2000 Guide to Statistics*, Vol. 2, Data Analysis Products Division, MathSoft, Seattle, WA.
- McCullagh, P., and Nelder, J.A. (1989), *Generalized Linear Models*, 2nd ed., Chapman & Hall, London, UK.
- McLachlan, G.J. (2004), *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, Hoboken, NJ.
- Meinshausen, N. (2007), "Relaxed Lasso," *Computational Statistics & Data Analysis*, 52, 374-393.
- Mevik, B.-H., Wehrens, R., and Liland, K.H. (2015), *pls: Partial Least Squares and Principal Component Regression*, R package version 2.5-0, (<https://CRAN.R-project.org/package=pls>).
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F. (2017), *e1071: Misc Functions of the Department of Statistics, Probability Theory Group*, R package version 1.6-8, (<https://CRAN.R-project.org/package=e1071>).
- Minor, R. (2012), "Poverty, Productivity, and Public Health: the Effects of "Right to Work" Laws on Key Standards of Living," *Thought & Action: the NEA Higher Education Journal*, 16-28. See (<http://www.nea.org/home/52880.htm>).
- Montgomery, D.C., Peck, E.A., and Vining, G. (2001), *Introduction to Linear Regression Analysis*, 3rd ed., Wiley, Hoboken, NJ.
- Montgomery, D.C., Peck, E.A., and Vining, G. (2012), *Introduction to Linear Regression Analysis*, 5th ed., Wiley, Hoboken, NJ.
- Mosteller, F., and Tukey, J.W. (1977), *Data Analysis and Regression*, Addison-Wesley, Reading, MA.
- Murphy, C. (2018), "Bootstrapping Forward Selection with BIC," Master's Research Paper, Southern Illinois University.
- Murphy, K.P. (2012), *Machine Learning: a Probabilistic Perspective*, MIT Press, Cambridge, MA.

Myers, R.H., Montgomery, D.C., and Vining, G.G. (2002), *Generalized Linear Models with Applications in Engineering and the Sciences*, Wiley, New York, NY.

Naul, B., and Taylor, J. (2017), "Sparse Steinian Covariance Estimation," *Journal of Computational and Graphical Statistics*, 26, 355-366.

Navarro, J. (2014), "Can the Bounds in the Multivariate Chebyshev Inequality be Attained?" *Statistics & Probability Letters*, 91, 1-5.

Navarro, J. (2016), "A Very Simple Proof of the Multivariate Chebyshev's Inequality," *Communications in Statistics: Theory and Methods*, 45, 3458-3463.

Nelder, J.A., and Wedderburn, R.W.M. (1972), "Generalized Linear Models," *Journal of the Royal Statistical Society, A*, 135, 370-380.

Ning, Y., and Liu, H. (2017), "A General Theory of Hypothesis Tests and Confidence Regions for Sparse High Dimensional Models," *The Annals of Statistics*, 45, 158-195.

Nishii, R. (1984), "Asymptotic Properties of Criteria for Selection of Variables in Multiple Regression," *The Annals of Statistics*, 12, 758-765.

Norman, G.R., and Streiner, D.L. (1986), *PDQ Statistics*, B.C. Decker, Philadelphia, PA.

Obozinski, G., Wainwright, M.J., and Jordan, M.I. (2011), "Support Union Recovery in High-Dimensional Multivariate Regression," *The Annals of Statistics*, 39, 1-47.

Olive, D.J. (2004), "Visualizing 1D Regression," in *Theory and Applications of Recent Robust Methods*, eds. Hubert, M., Pison, G., Struyf, A., and Van Aelst S., Birkhäuser, Basel.

Olive, D.J. (2007), "Prediction Intervals for Regression Models," *Computational Statistics & Data Analysis*, 51, 3115-3122.

Olive, D.J. (2008), *Applied Robust Statistics*, online course notes, see (<http://parker.ad.siu.edu/Olive/ol-bookp.htm>).

Olive, D.J. (2010), *Multiple Linear and 1D Regression Models*, online course notes, see (<http://parker.ad.siu.edu/Olive/regbk.htm>).

Olive, D.J. (2013a), "Asymptotically Optimal Regression Prediction Intervals and Prediction Regions for Multivariate Data," *International Journal of Statistics and Probability*, 2, 90-100.

Olive, D.J. (2013b), "Plots for Generalized Additive Models," *Communications in Statistics: Theory and Methods*, 42, 2610-2628.

Olive, D.J. (2014), *Statistical Theory and Inference*, Springer, New York, NY.

Olive, D.J. (2017a), *Linear Regression*, Springer, New York, NY.

Olive, D.J. (2017b), *Robust Multivariate Analysis*, Springer, New York, NY.

Olive, D.J. (2018), "Applications of Hyperellipsoidal Prediction Regions," *Statistical Papers*, 59, 913-931.

Olive, D.J. (2020a), *Theory for Linear Models*, online course notes, (<http://parker.ad.siu.edu/Olive/linmodbk.htm>).

- Olive, D.J. (2020b), *Robust Statistics*, online course notes, (<http://parker.ad.siu.edu/Olive/robbook.htm>).
- Olive, D.J. (2020c), *Survival Analysis*, online course notes, see (<http://parker.ad.siu.edu/Olive/survbk.htm>).
- Olive, D.J., and Hawkins, D.M. (2003), "Robust Regression with High Coverage," *Statistics & Probability Letters*, 63, 259-266.
- Olive, D.J., and Hawkins, D.M. (2005), "Variable Selection for 1D Regression Models," *Technometrics*, 47, 43-50.
- Olive, D.J., Rathnayake, R.C., and Haile, M.G. (2020), "Prediction Intervals for GLMs, GAMs, and Some Survival Regression Models," preprint at (<http://parker.ad.siu.edu/Olive/pppigam.pdf>).
- Öllerer, V., and Croux, C. (2015), "Robust High-Dimensional Precision Matrix Estimation," in *Modern Nonparametric, Robust and Multivariate Methods*, eds. Nordhausen, K., and Taskinen, S., Springer, New York, NY, 325-350.
- Pati, Y.C., Rezaiifar, R., and Krishnaprasad, P.S. (1993), "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition," in *Conference Record of the Twenty-Seventh Asilomar Conference on Signals, Systems and Computers* IEEE, 40-44.
- Pelawa Watagoda, L.C.R. (2017a), "Inference After Variable Selection," Ph.D. Thesis, Southern Illinois University. See (<http://parker.ad.siu.edu/Olive/slasanthiphd.pdf>).
- Pelawa Watagoda, L.C.R. (2017b), "Simulation for Inference After Variable Selection," unpublished manuscript online at (<http://parker.ad.siu.edu/Olive/slasanthisim.pdf>).
- Pelawa Watagoda, L.C.R. (2019), "A Sub-Model Theorem for Ordinary Least Squares," *International Journal of Statistics and Probability*, 8, 40-43.
- Pelawa Watagoda, L. C. R., and Olive, D.J. (2019), "Bootstrapping Multiple Linear Regression After Variable Selection," *Statistical Papers*, to appear. See (<http://parker.ad.siu.edu/Olive/ppboottest.pdf>).
- Pelawa Watagoda, L.C.R., and Olive, D.J. (2020), "Comparing Six Shrinkage Estimators With Large Sample Theory and Asymptotically Optimal Prediction Intervals," *Statistical Papers*, to appear. See (<http://parker.ad.siu.edu/Olive/pppicomp.pdf>).
- Politis, D.N., and Romano, J.P. (1996), "A General Theory for Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions," *The Annals of Statistics*, 22, 2031-2050.
- Pötscher, B. (1991), "Effects of Model Selection on Inference," *Econometric Theory*, 7, 163-185.
- Pourahmadi, M.P. (2011), "Covariance Estimation: the GLM and Regularization Perspectives," *Statistical Science*, 26, 369-387.
- Pourahmadi, M. (2013), *High-Dimensional Covariance Estimation*, Wiley, Hoboken, NJ.
- Pratt, J.W. (1959), "On a General Concept of "in Probability"," *The Annals of Mathematical Statistics*, 30, 549-558.

- Press, S.J. (2005), *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*, 2nd ed., Dover, New York, NY.
- Qi, X., Luo, R., Carroll, R.J., and Zhao, H. (2015), "Sparse Regression by Projection and Sparse Discriminant Analysis," *Journal of Computational and Graphical Statistics*, 24, 416-438.
- Rajapaksha, K.W.G.D. (2017), "Prediction Intervals After Forward Selection Using d Variables," Master's Research Paper, Southern Illinois University.
- Rao, C.R. (1965), *Linear Statistical Inference and Its Applications*, 1st ed., Wiley, New York, NY.
- Rathnayake, R.C. (2019), *Inference For Some GLMs and Survival Regression Models After Variable Selection*, Ph.D. thesis, Southern Illinois University, at (<http://parker.ad.siu.edu/Olive/srasanjphd.pdf>).
- Rathnayake, R.C., and Olive, D.J. (2020), "Bootstrapping Some GLMs and Survival Regression Models After Variable Selection," preprint at (<http://parker.ad.siu.edu/Olive/ppbootglm.pdf>).
- R Core Team (2016), "R: a Language and Environment for Statistical Computing," R Foundation for Statistical Computing, Vienna, Austria, (www.R-project.org).
- Rejchel, W. (2016), "Lasso With Convex Loss: Model Selection Consistency and Estimation," *Communications in Statistics: Theory and Methods*, 45, 1989-2004.
- Ren, J.-J. (1991), "On Hadamard Differentiability of Extended Statistical Functional," *Journal of Multivariate Analysis*, 39, 30-43.
- Ren, J.-J., and Sen, P.K. (1995), "Hadamard Differentiability on $D[0,1]^p$," *Journal of Multivariate Analysis*, 55, 14-28.
- Rencher, A., and Pun, F. (1980), "Inflation of R^2 in Best Subset Regression," *Technometrics*, 22, 49-53.
- Rinaldo, A., Wasserman, L., and G'Sell, M. (2019), "Bootstrapping and Sample Splitting for High-Dimensional, Assumption-Lean Inference," *The Annals of Statistics*, 47, 3438-3469.
- Rish, I., and Grabarnik, G.N. (2015), *Sparse Modeling: Theory, Algorithms, and Applications*, CRC Press Taylor & Francis, Boca Raton, FL.
- Ritter, G. (2014), *Robust Cluster Analysis and Variable Selection*, Chapman & Hall/CRC Press, Boca Rotan, FL.
- Ro, K., Zou, C., Wang, W., and Yin, G. (2015), "Outlier Detection for High-Dimensional Data," *Biometrika*, 102, 589-599.
- Rohatgi, V.K. (1976), *An Introduction to Probability Theory and Mathematical Statistics*, Wiley, New York, NY.
- Rohatgi, V.K. (1984), *Statistical Inference*, Wiley, New York, NY.
- Rothman, A.J., Bickel, P.J., Levina, E., and Zhu, J. (2008), "Sparse Permutation Invariant Covariance Estimation," *Electronic Journal of Statistics*, 2, 494-515.

- Romera, R. (2010), "Prediction Intervals in Partial Least Squares Regression Via a New Local Linearization Approach," *Chemometrics and Intelligent Laboratory Systems*, 103, 122-128.
- Rousseeuw, P.J., and Leroy, A.M. (1987), *Robust Regression and Outlier Detection*, Wiley, New York, NY.
- Schaaffhausen, H. (1878), "Die Anthropologische Sammlung Des Anatomischen Der Universitat Bonn," *Archiv fur Anthropologie*, 10, 1-65, Appendix.
- Schäfer, J., and Strimmer, K. (2007), "A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics," *Statistical Applications in Genetics and Molecular Biology*, 4, Article 32.
- Schomaker, M., and Heumann, C. (2014), "Model Selection and Model Averaging After Multiple Imputation," *Computational Statistics & Data Analysis*, 71, 758-770.
- Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461-464.
- Seber, G.A.F., and Lee, A.J. (2003), *Linear Regression Analysis*, 2nd ed., Wiley, New York, NY.
- Sen, P.K., and Singer, J.M. (1993), *Large Sample Methods in Statistics: an Introduction with Applications*, Chapman & Hall, New York, NY.
- Serfling, R.J. (1980), *Approximation Theorems of Mathematical Statistics*, Wiley, New York, NY.
- Severini, T.A. (2005), *Elements of Distribution Theory*, Cambridge University Press, New York, NY.
- Shao, J. (1993), "Linear Model Selection by Cross-Validation," *Journal of the American Statistical Association*, 88, 486-494.
- Shao, J., and Tu, D.S. (1995), *The Jackknife and the Bootstrap*, Springer, New York, NY.
- Shibata, R. (1984), "Approximate Efficiency of a Selection Procedure for the Number of Regression Variables," *Biometrika*, 71, 43-49.
- Silver, N. (2015), *The Signal and the Noise: Why So Many Predictions Fail-but Some Don't*, Penguin Books, New York, NY.
- Silverman, B.A. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York, NY.
- Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011), "Regularization Paths for Cox's Proportional Hazards Model Via Coordinate Descent," *Journal of Statistical Software*, 39, 1-13.
- Simonoff, J.S. (2003), *Analyzing Categorical Data*, Springer, New York, NY.
- Slawski, M., zu Castell, W., and Tutz, G., (2010), "Feature Selection Guided by Structural Information," *The Annals of Applied Statistics*, 4, 1056-1080.
- Srivastava, M.S., and Khatri, C.G. (1979), *An Introduction to Multivariate Statistics*, North Holland, New York, NY.

Staudte, R.G., and Sheather, S.J. (1990), *Robust Estimation and Testing*, Wiley, New York, NY.

Steinberger, L., and Leeb, H. (2016), “Leave-One-Out Prediction Intervals in Linear Regression Models With Many Variables,” unpublished manuscript at (<https://arxiv.org/pdf/1602.05801.pdf>).

Stewart, G.M. (1969), “On the Continuity of the Generalized Inverse,” *SIAM Journal on Applied Mathematics*, 17, 33-45.

Stigler, S.M. (1994), “Citation Patterns in the Journals of Statistics and Probability,” *Statistical Science*, 9, 94-108.

Su, W., Bogdan, M. , and Candés, E. (2017), “False Discoveries Occur Early on the Lasso Path,” *The Annals of Statistics*, 45, 2133-2150.

Su, W.J. (2018), “When is the First Spurious Variable Selected by Sequential Regression Procedures?” *Biometrika*, 105, 517-527.

Su, Z., and Cook, R.D. (2012), “Inner Envelopes: Efficient Estimation in Multivariate Linear Regression,” *Biometrika*, 99, 687-702.

Su, Z., Zhu, G., and Yang, Y. (2016), “Sparse Envelope Model: Efficient Estimation and Response Variable Selection in Multivariate Linear Regression,” *Biometrika*, 103, 579-593.

Sun, T., and Zhang, C.-H. (2012), “Scaled Sparse Linear Regression,” *Biometrika*, 99, 879-898.

Tarr, G., Müller, S., and Weber, N.C. (2016), “Robust Estimation of Precision Matrices Under Cellwise Contamination,” *Computational Statistics & Data Analysis*, 93, 404-420.

Theoharakis, V., and Skordia, M. (2003), “How Do Statisticians Perceive Statistics Journals,” *The American Statistician*, 57, 115-123.

Therneau, T.M. and Atkinson, E.J. (2017), “An Introduction to Recursive Partitioning Using the RPART Routines,” at (<https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf>).

Tibshirani, R. (1996), “Regression Shrinkage and Selection Via the Lasso,” *Journal of the Royal Statistical Society, B*, 58, 267-288.

Tibshirani, R. (1997), “The Lasso Method for Variable Selection in the Cox Model,” *Statistics in Medicine*, 16, 385-395.

Tibshirani, R., Bien, J., Friedman, J., Hastie, T., Simon, N., Taylor, J., and Tibshirani, R.J. (2012), “Strong Rules for Discarding Predictors in Lasso-Type Problems,” *Journal of the Royal Statistical Society, B*, 74, 245-266.

Tibshirani, R.J. (2013) “The Lasso Problem and Uniqueness,” *Electronic Journal of Statistics*, 7, 1456-1490.

Tibshirani, R.J. (2015), “Degrees of Freedom and Model Search,” *Statistica Sinica*, 25, 1265-1296.

Tibshirani, R.J., Rinaldo, A., Tibshirani, R., and Wasserman, L. (2018), “Uniform Asymptotic Inference and the Bootstrap After Model Selection,” *The Annals of Statistics*, 46, 1255-1287.

Tibshirani, R.J., and Taylor, J. (2012), “Degrees of Freedom in Lasso Problems,” *The Annals of Statistics*, 40, 1198-1232.

- Tibshirani, R.J., Taylor, J., Lockhart, R., and Tibshirani, R. (2016), "Exact Post-Selection Inference for Sequential Regression Procedures," *Journal of the American Statistical Association*, 111, 600-620.
- Tibshirani, R., Tibshirani, R., Taylor, J., Loftus, J. and Reid, S. (2016), *SelectiveInference: Tools for Post-Selection Inference*, R Package version 1.1.3.
- Tremearne, A.J.N. (1911), "Notes on Some Nigerian Tribal Marks," *Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 41, 162-178.
- Tukey, J.W. (1957), "Comparative Anatomy of Transformations," *The Annals of Mathematical Statistics*, 28, 602-632.
- Uraibi, H.S., Midi, H., and Rana, S. (2017), "Robust Multivariate Least Angle Regression," *Science Asia*, 43, 56-60.
- Uraibi, H.S., Midi, H., and Rana, S. (2019), "Selective Overview of Forward Selection in Terms of Robust Correlations," *Communications in Statistics: Simulations and Computation*, to appear.
- van de Geer, S., Bühlmann, P., Ritov, Y., and Dezeure, R. (2014), "On Asymptotically Optimal Confidence Regions and Tests for High-Dimensional Models," *The Annals of Statistics*, 42, 1166-1202.
- Venables, W.N., and Ripley, B.D. (1997), *Modern Applied Statistics with S*, 2nd ed. Springer, New York, NY.
- Venables, W.N., and Ripley, B.D. (2010), *Modern Applied Statistics with S*, 4th ed., Springer, New York, NY.
- Wackerly, D.D., Mendenhall, W., and Scheaffer, R.L. (2008), *Mathematical Statistics with Applications*, 7th ed., Thomson Brooks/Cole, Belmont, CA.
- Walpole, R.E., Myers, R.H., Myers, S.L., and Ye, K. (2006), *Probability & Statistics for Engineers & Scientists*, 8th ed., Prentice Hall, Upper Saddle River, NJ.
- Wang, H. (2009), "Forward Regression for Ultra-High Dimensional Variable Screening," *Journal of the American Statistical Association*, 104, 1512-1524.
- Wang, H. (2018), "Outlier Detection for High Dimensional Data," Master's Research Paper, Southern Illinois University.
- Wang, H., and Zhou, S.Z.F. (2013), "Interval Estimation by Frequentist Model Averaging," *Communications in Statistics: Theory and Methods*, 42, 4342-4356.
- Wang, L., Liu, X., Liang, H., and Carroll, R.J. (2011), "Estimation and Variable Selection for Generalized Additive Partial Linear Models," *The Annals of Statistics*, 39, 1827-1851.
- Warton, D.I. (2008), "Penalized Normal Likelihood and Ridge Regularization of Correlation and Covariance Matrices," *Journal of the American Statistical Association*, 103, 340-349.
- Wasserman, L. (2014), "Discussion: A Significance Test for the Lasso," *The Annals of Statistics*, 42, 501-508.
- White, H. (1984), *Asymptotic Theory for Econometricians*, Academic Press, San Diego, CA.

- Wieczorek, J.A. (2018), *Model Selection and Stopping Rules for High-Dimensional Forward Selection*, Ph.D. thesis, Carnegie Mellon University.
- Winkelmann, R. (2000), *Econometric Analysis of Count Data*, 3rd ed., Springer, New York, NY.
- Winkelmann, R. (2008), *Econometric Analysis of Count Data*, 5th ed., Springer, New York, NY.
- Wiseman, S.U., Hopke, P.K., and Schindler-Kaudelka, E. (1987), "Multielemental and Multivariate Analysis of Italian Terra Sigillata in the World Heritage Museum, University of Illinois at Urbana-Champaign," *Archeomaterials*, 1, 101-107.
- Witten, D.M., and Tibshirani, R. (2011), "Penalized Classification Using Fisher's Linear Discriminant," *Journal of the Royal Statistical Society, B*, 73, 753-772.
- Wold, H. (1975), "Soft Modelling by Latent Variables: the Nonlinear Partial Least Squares (NIPALS) Approach," in *Perspectives in Probability and Statistics, Papers in Honor of M.S. Bartlett*, ed. Gani, J., Academic Press, San Diego, CA, 117-144.
- Wold, H. (1985), "Partial Least Squares," *International Journal of Cardiology*, 147, 581-591.
- Wold, H. (2006), "Partial Least Squares," *Encyclopedia of Statistical Sciences*, Wiley, New York, NY.
- Wood, S.N. (2017), *Generalized Additive Models: an Introduction with R*, 2nd ed., Chapman & Hall/CRC, Boca Roton, FL.
- Xia, D. (2017), "Variable Selection of Linear Programming Discriminant Analysis," *Communications in Statistics: Theory and Methods*, 46, 3321-3341.
- Xu, H., Caramanis, C., and Mannor, S. (2011), "Sparse Algorithms are Not Stable: a No-Free-Lunch Theorem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP(99), 1-9.
- Yang, Y. (2003), "Regression with Multiple Candidate Models: Selecting or Mixing?" *Statistica Sinica*, 13, 783-809.
- Yao, J., Zheng, S., and Bai, Z. (2015), *Large Sample Covariance Matrices and High-Dimensional Data Analysis*, Cambridge University Press, New York, NY.
- Ye, J. (1998), "On Measuring and Correcting the Effects of Data Mining and Model Selection," *Journal of the American Statistical Association*, 93, 120-131.
- Yu, P.L.H., Wang, X., and Zhu, Y. (2017), "High Dimensional Covariance Matrix Estimation by Penalizing the Matrix-Logarithm Transformed Likelihood," *Computational Statistics & Data Analysis*, to appear.
- Yuan, M., and Lin, Y. (2007), "Model Selection and Estimation in the Gaussian Graphical Model," *Biometrika*, 94, 19-35.
- Zhang, P. (1992), "On the Distributional Properties of Model Selection Criterion," *Journal of the American Statistical Association*, 87, 732-737.
- Zhang, T., and Yang, B. (2017), "Box-Cox Transformation in Big Data," *Technometrics*, 59, 189-201.

Zhang, X., and Cheng, G. (2017), “Simultaneous Inference for High-Dimensional Linear Models,” *Journal of the American Statistical Association*, 112, 757-768.

Zheng, Z., and Loh, W.-Y. (1995), “Consistent Variable Selection in Linear Models,” *Journal of the American Statistical Association*, 90, 151-156.

Zhou, M. (2001), “Understanding the Cox Regression Models with Time-Change Covariates,” *The American Statistician*, 55, 153-155.

Zou, H., and Hastie, T. (2005), “Regularization and Variable Selection Via the Elastic Net,” *Journal of the Royal Statistical Society Series, B*, 67, 301-320.

Zou, H., Hastie, T., and Tibshirani, R. (1993), “Sparse Principal Component Analysis,” *Journal of Computational and Graphical Statistics*, 15, 265-286.

Zou, H., Hastie, T., and Tibshirani, R. (2007), “On the “Degrees of Freedom” of the Lasso,” *The Annals of Statistics*, 35, 2173-2192.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., and Smith, G.M. (2009), *Mixed Effects Models and Extensions in Ecology with R*, Springer, New York, NY.

Index

- Öllerer, 347
- 1D regression, 5, 75
- 1D regression model, 209
- Abraham, 221
- active set, 168
- additive error regression, 2, 6, 93, 210
- additive error single index model, 214
- additive predictor, 5
- AER, 4
- Aggarwal, 347
- Agostinelli, 347
- Agresti, 210, 211, 225, 226, 287
- Akaike, 76, 84, 271
- Anderson, 84, 198, 239
- AP, 4
- apparent error rate, 314
- asymptotic distribution, 29, 32
- asymptotic theory, 29
- asymptotically optimal, 90, 97
- Atkinson, 216, 295, 339, 354
- Büchlmann, 110
- bagging estimator, 110
- Becker, 358
- Belsley, 194, 342
- Berk, v, 137, 195
- Bertsimas, 194, 197
- beta–binomial regression, 210
- Bhatia, 67, 162
- Bickel, 107, 110, 134, 136
- binary regression, 210, 215
- binomial regression, 210, 215
- bivariate normal, 15
- Bogdan, 198
- boosting, 332
- bootstrap, 29, 136
- Boudt, 67, 347
- Box, 12
- Box–Cox transformation, 12
- Breiman, 110
- Brillinger, 209, 287
- Buckland, 137
- Budny, 101
- Burnham, 84, 198, 239
- Butler, 198
- Buxton, 23, 69, 72, 350
- c, 320
- Cai, 332, 347
- Cameron, 258, 287
- Camponovo, 195
- Candes, 197
- case, 1, 48
- Cauchy Schwartz inequality, 121
- cdf, 4, 17
- centering matrix, 19
- cf, 4, 41
- Chang, 138, 214, 287
- Charkhi, 88, 272
- Chatterjee, 195
- Chebyshev’s Inequality, 34
- Chen, 84, 101, 103, 133, 136, 176, 187, 188, 198, 271
- Cheng, 195
- Chew, 99
- Cho, 198
- Chun, 157, 195
- CI, 4
- Claeskens, 87, 88, 137, 272, 287
- Claeskens, 194
- Clarke, 136
- classical prediction region, 99
- CLT, 4

- coefficient of multiple determination, 52
- Collett, 222, 258, 287, 333
- concentration matrix, 347
- conditional distribution, 15
- confidence region, 102, 134
- confusion matrix, 316
- consistent, 34
- consistent estimator, 34
- constant variance MLR model, 48
- Continuity Theorem, 41
- Continuous Mapping Theorem, 40
- converges almost everywhere, 36
- converges in distribution, 32
- converges in law, 32
- converges in probability, 33
- converges in quadratic mean, 34
- Cook, v, 9, 54, 67, 94, 120, 126, 156, 157, 195, 196, 224, 238, 241, 258, 265, 287, 290, 332
- covariance matrix, 13
- coverage, 99
- covmb2, 22, 66
- Cox, 12, 76, 188, 211
- Cramér, 53
- Crawley, 358, 360
- cross validation, 314
- Croux, 347
- CV, 4

- DA, 4
- Daniel, 82
- data frame, 358
- data splitting, 245
- Datta, 194, 344
- degrees of freedom, 53, 200
- Delta Method, 30
- Denham, 195
- Devroye, 315
- Dezeure, 195
- df, 53
- discriminant function, 216
- double bootstrap, 138
- Duan, 287

- EAP, 4
- EC, 4
- EE plot, 79, 239
- Efron, 84, 103, 109, 110, 120, 136, 161, 162, 166, 169, 194, 195
- Efroymsen, 194
- eigenvalue, 153
- eigenvector, 153
- Ein-Dor, 283
- elastic net, 172
- elliptically contoured, 28
- elliptically contoured distribution, 98
- elliptically symmetric, 28
- empirical cdf, 104
- empirical distribution, 104
- error sum of squares, 52, 63
- ESP, 4
- ESSP, 4
- estimated additive predictor, 5, 209
- estimated sufficient predictor, 5, 209, 302
- estimated sufficient summary plot, 5
- Euclidean norm, 42
- Ewald, 137, 185, 195
- exponential family, 212
- extrapolation, 94, 176

- Fahrmeir, 236
- Fan, 87, 194, 195, 197, 198
- Farcomeni, 354
- FDA, 4
- Feldmesser, 283
- Feng, 347
- Ferguson, 40, 67
- Fernholtz, 136
- Ferrari, 194
- FF plot, 58, 79
- Filzmoser, 309
- Fithian, 194
- fitted values, 49, 145, 189
- Flury, 266
- Fogel, 194
- Forzani, 156, 157, 195
- Fox, 358
- Frank, 196
- Freedman, 94, 120–122
- Frey, 91, 103, 140, 281
- Friedman, v, 110, 196, 271, 302, 332, 346, 347, 357
- Fryzlewicz, 198
- Fu, 137, 163, 167, 194–196
- Fujikoshi, 198
- Fujikoshi, 347
- full model, 76, 143, 189
- Furnival, 82

- GAM, 4
- Gamma regression model, 210
- Gao, 197
- García-Escudero, 354
- Gaussian MLR model, 49
- generalized additive model, 5, 209, 249
- generalized eigenvalue problem, 306

- generalized linear model, 5, 209, 212, 213
 Gill, 136
 Gini's index, 324
 Gladstone, 61, 72, 127, 235, 255, 337–339
 GLM, 4, 213, 239
 Grübel, 96
 Grabarnik, 347
 Gram matrix, 160
 graphical lasso, 347
 Graybill, 148
 Greco, 354
 Gregory, 347
 Gruber, 195
 Gunst, 162, 163, 194
 Guo, 347
 Guttman, 62

 Hadamard derivative, 136
 Haggstrom, 217
 Haitovsky, 197
 Hall, 103, 110, 138, 195
 Hand, 333, 347
 Hastie, v, 84, 87, 136, 154, 157, 159–162, 166, 167, 169, 172, 176, 186, 187, 194–196, 198, 200, 271, 272, 287, 327, 332, 347
 hat matrix, 49, 62
 Haughton, 272
 Hawkins, 78, 195, 198, 199, 214, 240, 246, 251, 271, 277, 333
 hazard function, 279
 He, 347
 Hebbler, 149
 Hesterberg, 29, 136
 Heumann, 137
 hierarchical clustering, 349
 highest density region, 93, 97
 Hilbe, 249, 287
 Hjort, 86, 87, 137, 194, 272, 287
 Hoerl, 195
 Hoffman, 198
 Hong, 94, 272
 Hosmer, 216, 218, 242, 287
 Hsieh, 346, 347
 Huang, 197, 347
 Huberty, 332
 Hurvich, 84, 85, 133, 187, 286
 Hyndman, 97
 Hyodo, 347

 i, 105
 identity line, 6, 50, 129
 iid, 4, 5, 17, 48

 Izenman, v, 347

 Jacobian matrix, 43
 James, v, 2, 147, 181, 194, 285, 291, 298, 323, 325, 332, 336, 351
 Javanmard, 195
 Jia, 174
 Jin, 347, 354
 Johnson, v, 14, 28, 99, 153, 258, 298, 306
 Johnstone, 347
 joint distribution, 14
 Jolliffe, 156
 Jones, 83, 198

 Karhunen Loeve direction, 154
 Karhunen Loeve directions, 195
 Kaufman, 351, 354
 Kay, 238, 255
 Keles, 157, 195
 Kennard, 195
 Khatri, 342
 Kim, 198
 Knight, 137, 163, 167, 194–196
 KNN, 4
 Koch, 309, 310, 332, 347
 Kuhn, v

 ladder of powers, 8
 ladder rule, 9, 64
 Lahiri, 195
 Lai, 121, 195
 Larsen, v
 lasso, 4, 10, 147, 197, 347
 Law of Total Probability, 88
 LDA, 4
 least squares, 49
 Ledoit, 347
 Ledolter, 221
 Lee, 16, 61, 77, 194, 197
 Leeb, 85, 137, 194, 195, 198
 Lehmann, 36, 37, 67
 Lei, 95, 198
 Lemeshow, 216, 218, 242, 287
 Leroy, 199, 293
 leverage, 94, 176
 Li, 86, 87, 197, 287
 limiting distribution, 29, 32
 Lin, 197, 347
 Lindenmayer, 265
 Little, 238, 255
 Liu, 195, 332, 347
 location model, 17
 Lockhart, 194, 195
 log rule, 9, 64, 238

- logistic regression, 189, 215, 302
- Loh, 198
- LR, 4, 215
- LS CLT, 119, 136
- Lu, 195, 347
- Lumley, v, 357
- Luo, 84, 176, 187, 198
- Lv, 194, 195, 198
- Machado, 107
- MAD, 4, 17
- Mahalanobis distance, 19, 21, 22, 65
- Mai, 332, 347
- Mallows, 82, 83, 86, 198, 271
- Marden, v, 2
- Mardia, 307
- Markov's Inequality, 34
- Marquardt, 161
- Marx, v
- Mason, 162, 163, 194
- MathSoft, 351
- Mathsoft, 358
- McCullagh, 287
- McLachlan, 332
- MCLT, 4
- mean, 17
- MED, 4
- median, 17, 65
- median absolute deviation, 18, 65
- Meinshausen, 169, 196
- Mendenhall, v
- Mevik, v, 157, 357
- mgf, 4, 41
- minimum chi-square estimator, 225
- Minor, 334, 335
- mixture distribution, 46, 67
- MLD, 4
- MLR, 2, 4, 48
- model averaging, 137
- model sum of squares, 62
- modified power transformation, 10
- Montanari, 195
- Montgomery, 237
- Mosteller, 11
- multicollinearity, 58
- multiple linear regression, 2, 5, 48
- Multivariate analysis, 347
- Multivariate Central Limit Theorem, 43
- Multivariate Chebyshev's inequality, 100
- Multivariate Delta Method, 43
- Multivariate normal, 13
- Murphy, v
- MVN, 4, 14
- Myers, v, 226, 228
- Naul, 347
- Navarro, 101
- Nelder, 287
- Ning, 195
- Nishii, 86
- Nishiyama, 347
- nonparametric bootstrap, 105, 122, 139
- nonparametric prediction region, 99
- norm, 172
- normal equations, 62
- normal MLR model, 49
- Norman, 21, 341
- null classifier, 322
- Obozinski, 197, 347
- observation, 1
- OD plot, 258
- Olejnik, 332
- Olive, v, 2, 8, 48, 67, 78, 85–87, 89, 94, 97, 99, 101, 106, 108, 112, 136, 169, 174–176, 186, 194, 195, 198, 199, 214, 219, 240, 246, 251, 260, 265, 271–273, 277, 283, 287, 332, 333, 354
- OLS, 4, 10, 49
- order statistics, 17, 65, 91
- outlier resistant regression, 22
- outliers, 7, 16
- overdispersion, 219
- overfit, 77
- p, 341
- Pötscher, 85, 87, 88, 137, 194
- Parente, 107
- partial correlation, 342
- partial least squares, 147
- Pati, 196
- PCA, 4
- PCR, 4
- pdf, 4
- Pelawa Watagoda, 2, 76, 87, 89, 94, 97, 108, 111, 112, 136, 169, 174, 176, 185, 194, 195, 265, 272, 273
- percentile method, 103
- PI, 4
- PLS, 4
- pmf, 4
- Poisson regression, 210, 223, 287
- population correlation, 15
- population mean, 13
- positive definite, 153
- positive semidefinite, 153
- Pourahmadi, 347
- power transformation, 10

- Pratt, 38, 87
- precision matrix, 347
- predicted values, 49, 189
- prediction region, 97
- Press, 68
- principal component direction, 154
- principal component regression, 153
- principal components regression, 147, 153
- pval, 54, 59
- pvalue, 54
- QDA, 4
- Qi, 194, 197
- qualitative variable, 48
- quantitative variable, 48
- R, 357
- r, 283
- R Core Team, v, 139, 178
- random forests, 332
- Rao, 13
- Rathnayake, 85, 86, 136, 169, 175, 195, 260, 268, 273, 287
- Raven, 258
- Rayleigh quotient, 306
- regression sum of squares, 52
- regression through the origin, 62
- Rejchel, 195
- relaxed elastic net, 175, 185
- relaxed lasso, 147, 169, 185, 196
- Ren, 107, 110, 134, 136
- residual plot, 5, 49
- residuals, 49, 145, 190
- response plot, 5, 49, 79, 209
- response transformation, 12
- response transformation model, 210
- response variable, 1, 4
- Riani, 216
- ridge regression, 147, 197, 347
- Riedwyl, 266
- Rinaldo, 133, 286
- Ripley, v, 283, 321, 357, 358
- Rish, 347
- Ritter, 354
- Ro, 23, 347
- Rohatgi, 16, 41
- Rothman, 198, 347
- Rousseeuw, 199, 293, 351, 354
- RR plot, 58, 79
- rule of thumb, 21
- sample covariance matrix, 19, 65
- sample mean, 19, 29, 51, 65
- Schäfer, 347
- Schaaffhausen, 221, 236
- Scheaffer, v
- Schneider, 137, 185, 195
- Schomaker, 137
- Schwarz, 76, 84, 271
- score equations, 162
- SE, 4, 29
- Seber, 16, 61, 77
- selection bias, 85
- Sen, 67, 118, 136, 261
- separating hyperplane, 326
- Serfling, 67, 105
- Severini, 14, 45, 67, 162
- Shao, 86, 264
- Sheather, 140
- Shibata, 83
- shrinkage estimator, 136
- Silverman, 306, 332
- Simon, 283
- Simonoff, 210, 224, 249, 287
- Singer, 67, 118, 261
- singular value decomposition, 159
- Slawski, 174
- Slutsky's Theorem, 39, 45
- smallest extreme value distribution, 216
- smoothed bootstrap estimator, 110
- Snee, 161
- SP, 4
- sparse model, 3
- spectral decomposition, 153
- split conformal prediction interval, 95
- square root matrix, 154
- Srivastava, 342
- SSP, 4
- standard deviation, 18
- standard error, 29
- STATLIB, 242
- Staudte, 140
- Steinberger, 198
- Stewart, 67, 162
- Streiner, 21, 341
- Strimmer, 347
- Su, 54, 94, 120, 195, 197
- submodel, 76
- sufficient predictor, 5, 76, 209
- Sun, 197, 347
- supervised classification, 297
- supervised learning, 2
- support vector machines, 332
- survival function, 279
- SVD, 159

- SVM, 4
- Tanis, v
- Tao, 197
- Tarr, 23
- Taylor, 173, 197, 347
- test data, 1
- Therneau, 295, 339
- Tibshirani, 85, 136, 137, 168, 173, 194, 195, 271, 283, 287, 333, 347
- Tikhonov regularization, 195
- time series, 188
- total sum of squares, 52
- trace, 160
- training data, 1
- training error rate, 314
- transformation plot, 11, 12
- trees, 332
- Tremearne, 6, 202
- Trivedi, 258, 287
- truth table, 316
- Tsai, 84, 85, 133, 187, 286
- Tu, 264
- Tukey, 10–12
- Tutz, 236
- uncorrected total sum of squares, 62
- underfit, 77, 82
- underfitting, 76
- unimodal MLR model, 49
- unsupervised learning, 2
- Uraibi, 198
- van de Geer, 195
- variable selection, 239
- variance, 17, 18
- Venables, v, 283, 321, 357, 358
- von Mises differentiable statistical functions, 105
- VS-elastic net, 175
- VS-lasso, 169
- W, 36
- Wackerly, v
- Wagner, 315
- Walpole, v
- Wang, 74, 137, 198, 347, 354
- Warton, 344
- Wasserman, 198
- Wedderburn, 287
- Weisberg, v, 9, 126, 224, 238, 241, 258, 265, 287, 290, 358
- White, 44, 67
- Wichern, v, 14, 99, 153, 298, 306
- Wieczorek, 87, 198
- Wilson, 82
- Winkelmann, 224, 258, 287
- Wisseman, 337
- Witten, 333, 347
- Wold, 195
- Wolf, 347
- Wood, 82, 255, 287, 293
- Xia, 347
- Xu, 194
- Yang, 67, 110, 194, 198
- Yao, 347
- Ye, v
- Yu, 86, 110, 174, 347
- Yuan, 347
- Zhang, 67, 195, 197, 198, 287, 332
- Zhao, 86
- Zheng, 198
- Zhou, 137, 281, 287
- Zou, 172, 200, 347
- Zuur, 249, 287