



Universiteit
Leiden
The Netherlands

Feature network models for proximity data : statistical inference, model selection, network representations and links with related models

Frank, L.E.

Citation

Frank, L. E. (2006, September 21). *Feature network models for proximity data : statistical inference, model selection, network representations and links with related models*. Retrieved from <https://hdl.handle.net/1887/4560>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4560>

Note: To cite this publication please use the final published version (if applicable).

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723.
- Arabie, P. (1991). Was Euclid an unnecessarily sophisticated psychologist? *Psychometrika*, *56*, 567-587.
- Arabie, P., & Carroll, J. D. (1980). MAPCLUS: A mathematical programming approach to fitting the ADCLUS model. *Psychometrika*, *45*, 211-235.
- Ashby, F., & Maddox, W. (1990). Integrating information from separable dimensions. *Journal of Experimental Psychology: Human Perception & Performance*, *16*, 598-612.
- Attneave, F. (1950). Dimensions of similarity. *American Journal of Psychology*, *63*, 516-556.
- Backhaus, W., Menzel, R., & Kreißl, S. (1987). Multidimensional scaling of color similarities in bees. *Biological Cybernetics*, *56*, 293-304.
- Barthélemy, J. P., & Guénoche, A. (1991). *Trees and Proximity Representations*. New York: Wiley.
- Björk, A. (1996). *Numerical methods for least squares problems*. Philadelphia, PA.: SIAM.
- Borg, I., & Leutner, D. (1983). Dimensional models for the perception of rectangles. *Perception & Psychophysics*, *34*, 257-267.
- Bortz, J. (1974). Kritische Bemerkungen ueber den Einsatz nicht-euklidischer Metriken im Rahmen der multidimensionalen Skalierung [Critical remarks on the use of non-Euclidean metrics in the context of multidimensional scaling]. *Archiv für Psychologie*, *126*, 196-212.
- Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, *36*, 317-346.
- Brusco, M. (2001). A simulated annealing heuristic for unidimensional and multidimensional (City-Block) scaling of symmetric proximity matrices. *Journal of Classification*, *18*, 3-33.
- Brusco, M. (2002). Integer programming methods for seriation and unidimensional scaling of proximity matrices: A review and some extensions. *Journal of Classification*, *19*, 45-67.
- Buja, A., & Swayne, D. (2002). Visualization methodology for multidimensional scaling. *Journal of Classification*, *19*, 7-43.
- Bulmer, M. (1991). Use of the method of generalized least squares in reconstructing phylogenies from sequence data. *Molecular Biology and Evolution*, *8*, 868-883.
- Buneman, P. (1971). The recovery of trees from measures of dissimilarity. In F. R.

- Hodson, D. G. Kendall, & P. Tautu (Eds.), *Mathematics in the archeological and historical sciences* (p. 387-395). Edinburgh: University Press.
- Buneman, P. (1974). A note on the metric properties of trees. *Journal of Combinatorial Theory*, 17, 48-50.
- Busemann, H. (1955). *The geometry of geodesics*. New York: Academic Press (unabridged republication in 2005 by Dover Press, Mineola, New York).
- Carroll, J. D. (1976). Spatial, non-spatial and hybrid models for scaling. *Psychometrika*, 41(4), 439-463.
- Carroll, J. D., & Arabie, P. (1983). INDCLUS: An individual differences generalization of the ADCLUS model and the MAPCLUS algorithm. *Psychometrika*, 48, 157-169.
- Carroll, J. D., & Chang, J. J. (1973). A method for fitting a class of hierarchical tree structure models to dissimilarities data and its application to some "body parts" data of Miller's. In *Proceedings of the 81st annual convention of the american psychological association* (Vol. 8, p. 1097-1098).
- Carroll, J. D., Clark, L. A., & DeSarbo, W. S. (1984). The representation of three-way proximity data by single and multiple tree structure models. *Journal of Classification*, 1, 25-74.
- Carroll, J. D., & Corter, J. E. (1995). A graph-theoretic method for organizing overlapping clusters into trees, multiple trees, or extended trees. *Journal of Classification*, 12, 283-313.
- Carroll, J. D., & Pruzansky, S. (1980). Discrete and hybrid scaling models. In E. D. Lanterman & H. Feger (Eds.), *Similarity and Choice* (p. 108-139). Bern: Hans Huber.
- Chasles, M. (1875). *Aperçu historique sur l'origine et le développement des méthodes en géométrie : particulièrement de celles qui se rapportent à la géométrie moderne suivi d'un mémoire de géométrie sur deux principes généraux de la science, la dualité et l'homographie*. Paris: Gauthier-Villars.
- Chaturvedi, A., & Carroll, J. D. (1994). An alternating combinatorial optimization approach to fitting the INDCLUS and generalized INDCLUS models. *Journal of Classification*, 11, 155-170.
- Christensen, R. (2002). Linear and log-linear models. In A. E. Raftery, M. A. Tanner, & M. T. Wells (Eds.), *Statistics in the 21st Century* (p. 319-325). Boca Raton: Chapman & Hall.
- Colonius, H., & Schulze, H. (1981). Tree structures for proximity data. *British Journal of Mathematical and Statistical Psychology*, 34, 167-180.
- Corter, J. E. (1982). ADDTREE/P: A PASCAL program for fitting additive trees based on Sattath & Tversky's ADDTREE algorithm. *Behavior Research Methods and Instrumentation*, 14, 353-354.
- Corter, J. E. (1996). *Tree models of similarity and association*. Thousand Oaks: SAGE Publications.
- Corter, J. E. (1998). An efficient metric combinatorial algorithm for fitting additive trees. *Multivariate Behavioral Research*, 33, 249-272.
- Corter, J. E., & Tversky, A. (1986). Extended similarity trees. *Psychometrika*, 51, 429-451.
- Cross, D. (1965). Metric properties of multidimensional stimulus generalization.

- In D. Mostofsky (Ed.), *Stimulus generalization* (p. 72-93). Stanford: Stanford University Press.
- Cunningham, J. P. (1974, August). *Finding an optimal tree realization of a proximity matrix*. Paper presented at the Mathematical Psychology Meeting, Ann Arbor.
- Cunningham, J. P. (1978). Free trees and bidirectional trees as representations of psychological distance 19. *Journal of Mathematical Psychology*, 17, 165-188.
- Davey, B., & Priestley, H. (2002). *Introduction to Lattices and Order*. Cambridge, UK: Cambridge University Press.
- De Soete, G. (1983). A least squares algorithm for fitting additive trees to proximity data. *Psychometrika*, 48, 621-626.
- De Soete, G., & Carroll, J. D. (1996). Tree and other network models for representing proximity data. In P. Arabie, L. J. Hubert, & G. De Soete (Eds.), *Clustering and classification*. River Edge, NJ: World Scientific Publishing.
- Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis*. New York: Wiley.
- Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, 37, 409-428.
- Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. J. (2004). Least Angle Regression. *The Annals of Statistics*, 32, 407-499.
- Efron, B., & Tibshirani, R. J. (1998). *An introduction to the Bootstrap*. Boca Raton: CRC Press LLC.
- Eisler, H. (1973). The algebraic and statistical tractability of the city-block metric. *British Journal of Mathematical and Statistical Psychology*, 26, 212-218.
- Eisler, H., & Roskam, E. (1977). Multidimensional similarity: An experimental and theoretical comparison of vector, distance, and set-theoretical models. *Acta Psychologica*, 41, 1-46 and 335-363.
- Emde, G. Von der, & Ronacher, B. (1994). Perception of electric properties of objects in electrolocating weakly electric fish: Two-dimensional similarity scaling reveals a City-Block metric. *Journal of Comparative Physiology A*, 175, 801-812.
- Felsenstein, J. (1983). Statistical inference of phylogenies. *Journal of the Royal Statistical Society: Series A (General)*, 146, 246-272.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39, 783-791.
- Felsenstein, J. (2004). *Inferring Phylogenies*. Sunderland, Massachusetts: Sinauer Associates.
- Flament, C. (1963). *Applications of graph theory to group structure*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Frank, L. E., & Heiser, W. J. (2004, June). *Statistical inference in Feature Network Models and additive trees*. Paper presented at the International Meeting of the Psychometric Society, Pacific Grove, CA.
- Frank, L. E., & Heiser, W. J. (2005). Standard errors, prediction error and model tests in additive trees. Manuscript submitted for publication.
- Frank, L. E., & Heiser, W. J. (in press a). Estimating standard errors in Feature Network Models. *British Journal of Mathematical and Statistical Psychology*.
- Frank, L. E., & Heiser, W. J. (in press b). Feature selection in Feature Network Models: finding predictive subsets of features with the Positive Lasso. *British Journal of Mathematical and Statistical Psychology*.

- Freedman, D. A. (1981). Bootstrapping regression models. *The Annals of Statistics*, 9, 1218-1228.
- Freedman, D. A., & Peters, S. C. (1984). Bootstrapping a regression equation: some empirical results. *Journal of the American Statistical Association*, 79, 97-106.
- Friedman, J. H., & Popescu, B. E. (2006). Gradient directed regularization for linear regression and classification. Submitted manuscript.
- Galanter, E. H. (1956). An axiomatic and experimental study of sensory order and measure. *Psychological Review*, 63(1), 16-28.
- Gardner, M. (1972). The curious properties of the Gray code and how it can be use to solve puzzles. *Scientific American*, 227, 106-109.
- Garner, W. (1974). *The processing of information and structure*. New York: Wiley.
- Gascuel, O. (1994). A note on Sattath and Tversky's, Saitou and Nei's, and Studier and Keppler's algorithms for inferring phylogenies from evolutionary distances. *Molecular Biology and Evolution*, 11, 961-963.
- Gascuel, O., & Levy, D. (1996). A reduction algorithm for approximating a (non-metric) dissimilarity by a tree distance. *Journal of Classification*, 13, 129-155.
- Gati, I., & Tversky, A. (1982). Representations of qualitative and quantitative dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 325-340.
- Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology*, 16, 341-370.
- Gilbert, E. N. (1958). Gray codes and paths on the n -cube. *Bell Systems Technical Journal*, 37, 815-826.
- Goldstone, R. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, 123, 178-200.
- Golub, G. H., & Loan, C. F. van. (1989). *Matrix Computations* (second ed.). Baltimore: The Johns Hopkins University Press.
- Gonçalves, S., & White, H. (2005). Bootstrap standard error estimates for linear regression. *Journal of the American Statistical Association*, 100, 970-979.
- Goodman, N. (1951). *The Structure of Appearance*. Indianapolis, Indiana: Bobbs-Merrill.
- Goodman, N. (1977). *The Structure of Appearance* (3rd ed.). Dordrecht, Holland: Reidel.
- Gray, F. (1953, March). Pulse code communications. U.S. Patent 2632058.
- Groenen, P. J. F., & Heiser, W. J. (1996). The tunneling method for global optimization in multidimensional scaling. *Psychometrika*, 61(3), 529-550.
- Groenen, P. J. F., Heiser, W. J., & Meulman, J. J. (1998). City-block scaling: Smoothing strategies for avoiding local minima. In Balderjahn, R. R. Mathar, & M. Schader (Eds.), *Classification, data analysis and data highways* (p. 46-53). Berlin: Springer.
- Groenen, P. J. F., Heiser, W. J., & Meulman, J. J. (1999). Global optimization in least-squares multidimensional scaling by distance smoothing. *Journal of Classification*, 16, 225-254.
- Gulliksson, M., & Wedin, P.-Å. (2000). Perturbation theory for generalized and constrained linear least squares. *Numerical Linear Algebra with Applications*, 7, 181-195.

-
- Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157-1182.
- Hastie, T., Tibshirani, R. J., & Friedman, J. H. (2001). *The Elements of Statistical Learning; Data mining, Inference, and Prediction*. New York: Springer Verlag.
- Hays, W. (1958). An approach to the study of trait implication and trait similarity. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (p. 289-299). Stanford: Stanford University Press.
- Heiser, W. J. (1989). The city-block model for three-way multidimensional scaling. In R. Coppi & S. Bolasco (Eds.), *Multivariate data analysis* (p. 395-404). Amsterdam: North-Holland.
- Heiser, W. J. (1991). A generalized majorization method for least squares multidimensional scaling of pseudodistances that may be negative. *Psychometrika*, 56(1), 7-27.
- Heiser, W. J. (1998). Fitting graphs and trees with multidimensional scaling methods. In C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H.-H. Bock, & Y. Baba (Eds.), *Data science, classification and related methods* (p. 52-62). Tokyo: Springer Verlag.
- Heiser, W. J., & Busing, F. M. T. A. (2004). Multidimensional scaling and unfolding of symmetric and asymmetric proximity relations. In D. Kaplan (Ed.), *The SAGE handbook of quantitative methodology for the social sciences* (p. 25-48). Thousand Oaks, CA: Sage.
- Heiser, W. J., & Frank, L. E. (2005). On the relationship between the distinctive features model and the common features model. unpublished manuscript.
- Heiser, W. J., & Meulman, J. J. (1997). Representation of binary multivariate data by graph models using the hamming metric. In E. Wegman & S. Azen (Eds.), *Computing science and statistics* (Vol. 29 (2), p. 517-525). Fairfax, VA: Interface Foundation of North America.
- Hubert, L. J., & Arabie, P. (1988). Relying on necessary conditions for optimization: Unidimensional scaling and some extensions. In H. H. Bock (Ed.), *Classification and related methods of data analysis* (p. 463-472). Amsterdam: North-Holland.
- Hubert, L. J., Arabie, P., & Hesson-Mcinnis, M. (1992). Multidimensional scaling in the city-block metric: a combinatorial approach. *Journal of Classification*, 9, 211-136.
- Hubert, L. J., Arabie, P., & Meulman, J. J. (2001). *Combinatorial Data Analysis: Optimization by Dynamic Programming*. Philadelphia, NJ: SIAM.
- Huelsenbeck, J. P., & Rannala, B. (1997). Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science*, 276, 227-232.
- Jain, A., & Dubes, R. (1988). *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice Hall.
- Jakobson, R., Fant, C. G. M., & Halle, M. (1965). *Preliminaries to speech analysis. The distinctive features and their correlates. (Reprint of the original edition of 1953)*. Cambridge, MA: MIT Press.
- Joly, S., & Le Calvé, G. (1994). Similarity functions. In B. van Cutsem (Ed.), *Classification and Dissimilarity Analysis* (p. 67-86). New York: Springer Verlag.
- Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31, 7-15.

- Keren, G., & Baggen, S. (1981). Recognition models for alphanumeric characters. *Perception & Psychophysics*, *29*, 234-246.
- Kishino, H., & Hasegawa, M. (1990). Converting distance to time: application to human evolution. *Methods in Enzymology*, *183*, 550-570.
- Klatzky, R., Beall, A., Loomis, J., Golledge, R., & Philbeck, J. (1999). Human navigation ability: Tests of the encoding error model of path integration. *Spatial Cognition and Computation*, *1*, 31-65.
- Klauer, K. C. (1989). Ordinal network representation: representing proximities by graphs. *Psychometrika*, *54*(4), 737-750.
- Klauer, K. C. (1994). Representing proximities by network models. In E. Diday (Ed.), *New approaches in classification and data analysis*. Heidelberg: Springer Verlag.
- Klauer, K. C., & Carroll, J. D. (1989). A mathematical programming approach to fitting general graphs. *Journal of Classification*, *6*, 247-270.
- Krackhardt, D. (1988). Predicting with networks: nonparametric multiple regression analysis of dyadic data. *Social Networks*, *10*, 359-381.
- Kruschke, J. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22-44.
- Kuennapas, T., & Janson, A.-J. (1969). Multidimensional similarity of letters. *Perceptual and Motor Skills*, *28*, 3-12.
- Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, *17*, 1217-1241.
- Landahl, H. (1945). Neural mechanisms for the concepts of difference and similarity. *Bulletin of Mathematical Biophysics*, *7*, 83-88.
- Lawson, C. L., & Hanson, R. J. (1995). *Solving least squares problems*. Philadelphia: Society for Industrial and Applied Mathematics.
- Le Calvé, G. (1985). Distances à centre [center distances]. *Statistiques et Analyse des Données*, *10*, 29-44.
- Lee, M. D. (2001). On the complexity of additive clustering models. *Journal of Mathematical Psychology*, *45*, 131-148.
- Lee, M. D., & Navarro, D. J. (2002). Extending the ALCOVE model of category learning to featural stimulus domains. *Psychonomic Bulletin & Review*, *9*, 43-58.
- Lee, W. C., & Rodgers, J. L. (1998). Bootstrapping correlation coefficients using univariate and bivariate sampling. *Psychological Methods*, *3*, 91-103.
- Li, W.-H. (1989). A statistical test of phylogenies estimated from sequence data. *Molecular Biology and Evolution*, *6*, 424-435.
- Liew, C. K. (1976). Inequality constrained least-squares estimation. *Journal of the American Statistical Association*, *71*, 746-751.
- Liu, R. Y., & Singh, K. (1992a). Moving blocks jackknife and bootstrap capture weak dependence. In R. LePage & L. Billiard (Eds.), *Exploring the Limits of the Bootstrap* (p. 225-248). New York: Wiley.
- Liu, R. Y., & Singh, K. (1992b). Efficiency and robustness in resampling. *The Annals of Statistics*, *20*, 370-384.
- MacKay, D. (2001). Probabilistic multidimensional scaling using a city-block metric. *Journal of Mathematical Psychology*, *45*(3), 249-264.
- Mallows, C. L. (1973). Some comments on C_p . *Technometrics*, *15*, 661-675.
- Mallows, C. L. (1995). More comments on C_p . *Technometrics*, *37*, 362-372.

-
- Martinez, W. L., & Martinez, A. R. (2002). *Computational statistics handbook with Matlab*. Boca Raton: Chapman and Hall.
- Melara, R., Marks, L., & Lesko, K. (1992). Optional processes in similarity judgments. *Perception & Psychophysics*, *51*, 123-133.
- Meulman, J. J., & Heiser, W. J. (1999). *Categories*. Chicago: SPSS.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*, 156-166.
- Micko, H. C., & Fischer, W. (1970). The metric of multidimensional psychological spaces as a function of differential attention to subjective attributes. *Journal of Mathematical Psychology*, *7*, 118-143.
- Miller, G. A. (1969). Psychological method to investigate verbal concepts. *Journal of Mathematical Psychology*, *6*, 169-191.
- Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some english consonants. *Journal of the Acoustical Society of America*, *27*, 338-352.
- Mirkin, B. (1987). Additive clustering and qualitative factor analysis methods for similarity matrices. *Journal of Classification*, *6*, 247-270.
- Mirkin, B. (1990). A sequential fitting procedure for linear data analysis models. *Journal of Classification*, *7*, 167-195.
- Mirkin, B. (1996). Clustering for contingency tables: boxes and partitions. *Statistics and Computing*, *6*, 217-229.
- Mirkin, B. (1998). Least squares structuring, clustering, and data processing issues. *The Computer Journal*, *41*, 518-536.
- Mooijaart, A., van der Heijden, P. G. M., & van der Ark, L. (1999). A least squares algorithm for a mixture model for compositional data. *Computational Statistics & Data Analysis*, *30*, 359-379.
- Navarro, D. J., & Lee, M. D. (2001). Clustering using the contrast model. In J. Moore & K. Stenning (Eds.), *Proceedings of the 23rd annual conference of the cognitive science society* (p. 686-691). Mahwah, NJ: Lawrence Erlbaum.
- Navarro, D. J., & Lee, M. D. (2004). Common and distinctive features in stimulus similarity: A modified version of the contrast model. *Psychonomic Bulletin & Review*, *11*, 961-974.
- Nei, M., Stephens, J. C., & Saitou, N. (1985). Methods for computing the standard errors of branching points in an evolutionary tree and their application to molecular data from human and apes. *Molecular Biology and Evolution*, *2*, 66-85.
- Nijenhuis, A., & Wilf, H. S. (1978). *Combinatorial algorithms for computers and calculators*. New York: Academic Press.
- Nosofsky, R. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *10*, 104-114.
- Nosofsky, R. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.
- Nosofsky, R. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*, 87-108.
- Nosofsky, R. (1992). Exemplars, prototypes, and similarity rules. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning theory to connectionist theory: Es-*

- says in honour of william estes, vol. 1* (p. 149-167). Hillsdale, NJ: Lawrence Erlbaum.
- Nosofsky, R., & Zaki, S. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 924-940.
- Okada, A., & Imaizumi, T. (1980). Nonmetric method for extended INDSCAL model. *Behaviormetrika*, 7, 13-22.
- Ota, R., Waddell, P. J., Hasegawa, M., Shimodaira, H., & Kishino, H. (2000). Appropriate likelihood ratio tests and marginal distributions for evolutionary tree models with constraints on parameters. *Molecular Biology and Evolution*, 17, 798-803.
- Parault, S., & Schwanenflugel, P. (2000). The development of conceptual categories of attention during the elementary school years. *Journal of Experimental Child Psychology*, 75, 245-262.
- Pruzansky, S., Tversky, A., & Carroll, J. D. (1982). Spatial versus tree representations of proximity data. *Psychometrika*, 47(1), 3-19.
- Ramsay, J. O. (1982). Some statistical approaches to multidimensional scaling data. *Journal of the Royal Statistical Society. Series A (General)*, 145(3), 285-312.
- Restle, F. (1959). A metric and an ordering on sets. *Psychometrika*, 24, 207-220.
- Restle, F. (1961). *Psychology of judgment and choice*. New York: Wiley.
- Ronacher, B. (1992). Pattern recognition in honeybees: Multidimensional scaling reveals city-block metric. *Vision Research*, 32, 1837-1843.
- Rosenberg, S., & Kim, M. P. (1975). The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10(4), 489-502.
- Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology*, 53, 94-101.
- Rzhetsky, A., & Nei, M. (1992). A simple method for estimating and testing minimum-evolution trees. *Molecular Biology and Evolution*, 9, 945-967.
- Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4, 406-425.
- Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika*, 42, 319-345.
- Sattath, S., & Tversky, A. (1987). On the relation between common and distinctive feature models. *Psychological Review*, 94(1), 16-22.
- Savage, C. (1997). A survey of combinatorial Gray codes. *SIAM Review*, 39, 605-629.
- Schott, B. (2003). *Schott's Original Miscellany*. New York: Bloomsbury.
- Self, S. G., & Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82, 605-610.
- Sen, P. K., & Silvapulle, M. J. (2002). An appraisal of some aspects of statistical inference under inequality constraints. *Journal of Statistical Planning and Inference*, 107, 3-43.
- Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1, 54-87.
- Shepard, R. N. (1972). Psychological representation of speech sounds. In E. E.

- David Jr. & P. Denes (Eds.), *Human communication: A unified view* (p. 67-113). New York: McGraw-Hill.
- Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, *39*, 373-421.
- Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, *210*, 390-398.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317-1323.
- Shepard, R. N. (1991). Integrality versus separability of stimulus dimensions: From an early convergence of evidence to a proposed theoretical basis. In G. Lockhead & J. Pomerantz (Eds.), *The perception of structure* (p. 53-71). Washington, DC: APA.
- Shepard, R. N., & Arabie, P. (1979). Additive clustering: representation of similarities as combinations of discrete overlapping properties. *Psychological Review*, *2*, 87-123.
- Shepard, R. N., & Chang, J. J. (1963). Stimulus generalization in the learning of classifications. *Journal of Experimental Psychology*, *65*, 94-102.
- Shepard, R. N., Hovland, H. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, *75*, 13, Whole number 517.
- Shepard, R. N., Kilpatrick, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology*, *7*, 82-138.
- Sitnikova, T., Rzhetsky, A., & Nei, M. (1995). Interior-branch and bootstrap tests of phylogenetic trees. *Molecular Biology and Evolution*, *12*, 319-333.
- Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized Latent Variable Modeling. Multilevel, Longitudinal, and Structural Equations Models*. Boca Raton: Chapman & Hall.
- Soli, S. D., Arabie, P., & Carroll, J. D. (1986). Discrete representation of perceptual structure underlying consonant confusions. *Journal of the Acoustical Society of America*, *79*, 826-837.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Stram, D. O., & Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics*, *50*, 1171-1177.
- Suppes, P., Krantz, D., Luce, R., & Tversky, A. (1989). *Foundations of Measurement, Vol. II, Geometrical, Threshold, and Probabilistic Representations*. New York: Academic Press.
- Tajima, F. (1992). Statistical method for estimating the standard errors of branch lengths in a phylogenetic tree reconstructed without assuming equal rates of nucleotide substitution among different lineages. *Molecular Biology and Evolution*, *9*, 168-181.
- Takane, Y. (1981). Multidimensional successive categories scaling: a maximum likelihood method. *Psychometrika*, *46*, 9-28.
- Takane, Y. (1983). Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika*, *48*, 393-423.
- Takane, Y., & Carroll, J. D. (1981). Nonmetric maximum likelihood multidimensional scaling from directional rankings of similarities. *Psychometrika*, *46*, 389-405.

- Takane, Y., & Sergent, J. (1983). Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika*, *48*, 393-423.
- Tatsuoka, K. K. (1995). Architecture of knowledge structures and cognitive diagnosis: a statistical pattern recognition and classification approach. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively diagnostic assessment* (p. 327-359). Hillsdale, NJ: Erlbaum.
- Ten Berge, J. M. F., & Kiers, H. A. L. (2005). A comparison of two methods for fitting the INDCLUS model. *Journal of Classification*, *22*, 273-285.
- Theil, H. (1971). *Principles of Econometrics*. New York: Wiley.
- Tibshirani, R. J. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*, 267-288.
- Torgerson, W. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, *17*, 401-419.
- Torgerson, W. (1958). *Theory and methods of scaling*. New York: Wiley.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327-352.
- Tversky, A., & Gati, I. (1982). Similarity, separability, and the triangle inequality. *Psychological Review*, *89*, 123-154.
- Tversky, A., & Hutchinson, J. (1986). Nearest neighbor analysis of psychological spaces. *Psychological Review*, *93*, 3-22.
- Verbeemen, T., Storms, G., & Verguts, T. (2004). Similarity and taxonomy in categorization. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the cognitive science society* (p. 1393-1398). Mahwah, NJ: Lawrence Erlbaum.
- Verbeemen, T., Vanoverberghe, V., Storms, G., & Ruts, W. (2001). The role of contrast categories in natural language concepts. *Journal of Memory and Language*, *44*, 618-643.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge UK: Cambridge University Press.
- Weinberg, S. L., Carroll, J. D., & Cohen, H. S. (1984). Confidence regions for INDSCAL using the jackknife and bootstrap techniques. *Psychometrika*, *49*, 475-491.
- Werner, H. J. (1990). On inequality constrained generalized least squares estimation. *Linear Algebra and Its Applications*, *127*, 379-392.
- Werner, H. J., & Yapar, C. (1996). On inequality constrained generalized least squares selections in the general possibly singular gauss-markov model: a projector theoretical approach. *Linear Algebra and Its Applications*, *237*, 359-393.
- Winsberg, S., & Ramsay, J. O. (1981). Analysis of pairwise preference data using integrated B-splines. *Psychometrika*, *46*, 171-186.
- Wolak, F. A. (1987). An exact test for multiple inequality and equality constraints in the linear regression model. *Journal of the American Statistical Association*, *82*, 782-793.
- Wollan, P. C., & Dykstra, R. L. (1987). Algorithm AS 225: Minimizing linear inequality constrained Mahalanobis distances. *Applied Statistics*, *36*, 234-240.
- Yang, Y. (2005). Can the strengths of AIC and BIC be shared? A conflict between model identification and regression estimation. *Biometrika*, *92*, 937-950.
- Zaki, S., Nosofsky, R., Stanton, R., & Cohen, A. (2003). Prototype and exemplar ac-

- counts of category learning and attentional allocation: A reassessment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 1160-1173.
- Zielman, B., & Heiser, W. J. (1996). Models for asymmetric proximities. *British Journal of Mathematical and Statistical Psychology*, 49, 127-146.
- Zou, H., Hastie, T., & Tibshirani, R. J. (2006). On the degrees of freedom of the Lasso. Submitted manuscript.

Author Index

- Akaike, H., 66, 159
Arabie, P., 5, 14, 84, 88, 92, 100, 115–117,
133, 134, 137, 140, 143, 159, 160,
163, 167
Ashby, F., 116, 159
Attneave, F., 115, 116, 143, 144, 159, 176

Backhaus, W., 115, 159
Baggen, S., 5, 84, 164
Barthélemy, J. P., 142, 159
Beall, A., 145, 164
Björk, A., 29, 64, 159
Borg, I., 115, 121, 144, 159
Bortz, J., 127, 159
Box, G. E. P., 152, 159
Brusco, M., 115, 159
Buja, A., 116, 120, 159
Bulmer, M., 53, 80, 159
Buneman, P., 57, 117, 142, 159, 160
Busemann, H., 118, 160
Busing, F. M. T. A., 132, 163

Carroll, J. D., 5, 6, 10, 12, 14, 15, 22, 25, 38,
49, 52, 60, 66, 79, 84, 88, 92, 100,
117, 120, 124, 133, 134, 137, 138,
140, 142, 143, 145, 147, 154,
159–161, 164, 166–168
Chang, J. J., 116, 137, 138, 140, 143, 160, 167
Chasles, M., 160
Chaturvedi, A., 133, 160
Christensen, R., 152, 160
Clark, L. A., 142, 160
Cohen, A., 115, 168
Cohen, H. S., 38, 168
Colonus, H., 142, 160
Corter, J. E., 5, 6, 12, 14, 15, 22, 25, 42, 49, 52,
56, 84, 88, 117, 130, 131, 133, 134,
142, 143, 145, 147, 154, 160

Cross, D., 143, 160
Cunningham, J. P., 131, 142, 161, 167

Davey, B., 128, 161
DeSarbo, W. S., 142, 160
De Soete, G., 52, 60, 142, 161
Draper, N. R., 99, 161
Dubes, R., 121, 163
Durbin, J., 152, 161
Dykstra, R. L., 29, 64, 149, 168

Efron, B., 17, 30–33, 41, 69, 85, 95–97, 151,
153, 161, 179
Eisler, H., 115, 161
Elisseeff, A., 93, 99, 163
Emde, G. Von der, 115, 117, 161

Fant, C. G. M., 6, 163
Faust, K., 10, 168
Felsenstein, J., 52, 57, 58, 66, 80, 161
Fischer, W., 115, 165
Flament, C., 8, 25, 59, 89, 130, 161
Frank, L. E., 49, 84, 88, 89, 104, 132, 147, 148,
161, 163
Freedman, D. A., 41, 151, 162
Friedman, J. H., 67, 112, 162, 163

Galanter, E. H., 128, 162
Gardner, M., 91, 162
Garner, W., 115, 116, 162
Gascuel, O., 15, 53, 54, 79, 162
Gati, I., 2, 4, 5, 19, 59, 84, 162, 168
Gilbert, E. N., 91, 162
Goldstone, R., 116, 162
Golledge, R., 145, 164
Golub, G. H., 149, 162
Gonçalves, S., 151, 153, 162
Goodman, N., 7, 24, 55, 59, 87, 162, 176

- Gray, F., 90, 162
Groenen, P. J. F., 115, 121, 162
Guénoche, A., 142, 159
Gulliksson, M., 153, 162
Guyon, I., 93, 99, 163
- Halle, M., 6, 163
Hanson, R. J., 149, 164
Hasegawa, M., 53, 66, 164, 166
Hastie, T., 17, 67, 94, 96, 161, 163, 169
Hays, W., 129, 163
Heiser, W. J., 5, 7, 8, 10, 11, 14, 22, 24, 25, 27, 49, 52, 54–57, 59, 61, 83–85, 87–89, 104, 115, 116, 130–132, 147, 148, 155, 161–163, 165, 169, 176, 177
Hesson-Mcinnis, M., 116, 163
Hovland, H. I., 116, 167
Hubert, L. J., 116, 140, 163
Huelsenbeck, J. P., 66, 163
Hutchinson, J., 70, 168
- Imaizumi, T., 116, 166
- Jain, A., 121, 163
Jakobson, R., 6, 163
Janson, A.-J., 42, 164
Jenkins, H. M., 116, 167
Johnstone, I., 17, 161
Joly, S., 119, 163
- Kamada, T., 120, 163
Kawai, S., 120, 163
Keren, G., 5, 84, 164
Kiers, H. A. L., 133, 168
Kilpatric, D. W., 131, 167
Kim, M. P., 54, 166
Kishino, H., 53, 66, 164, 166
Klatzky, R., 145, 164
Klauer, K. C., 10, 15, 22, 25, 120, 164
Krackhardt, D., 15, 22, 164
Krantz, D., 119, 167
Kreißl, S., 115, 159
Kruschke, J., 115, 116, 164
Kuennapas, T., 42, 164
Künsch, H. R., 153, 164
- Landahl, H., 143, 164
Lawson, C. L., 149, 164
Le Calvé, G., 119, 124, 163, 164
Lee, J. W., 53, 167
- Lee, M. D., 5, 18, 22, 84, 85, 129, 130, 147, 148, 164, 165
Lee, W. C., 164
Lesko, K., 116, 165
Leutner, D., 115, 121, 144, 159
Levy, D., 15, 53, 79, 162
Li, W.-H., 53, 80, 164
Liang, K. Y., 53, 150, 151, 166
Liew, C. K., 16, 23, 28, 29, 53, 64, 65, 67, 149, 150, 164, 177
Liu, R. Y., 151, 153, 164
Loan, C. F. van, 149, 162
Loomis, J., 145, 164
Luce, R., 119, 167
- MacKay, D., 115, 164
Maddox, W., 116, 159
Mallows, C. L., 96, 164
Marks, L., 116, 165
Martinez, A. R., 32, 165
Martinez, W. L., 32, 165
Melara, R., 116, 117, 165
Menzel, R., 115, 159
Meulman, J. J., 10, 11, 14, 57, 61, 89, 116, 131, 140, 155, 162, 163, 165
Micceri, T., 152, 165
Micko, H. C., 115, 165
Miller, G. A., 23, 32, 86, 98, 136, 165
Mirkin, B., 133, 135, 165
Mooijaart, A., 50, 165
- Navarro, D. J., 5, 18, 22, 84, 85, 129, 130, 147, 148, 164, 165
Nei, M., 52–54, 70, 80, 81, 142, 165–167
Nicely, P. E., 23, 32, 86, 98, 165
Nijenhuis, A., 72, 93, 112, 165
Nosofsky, R., 115, 116, 148, 165, 166, 168
- Okada, A., 116, 166
Ota, R., 53, 166
- Parault, S., 130, 166
Peters, S. C., 41, 151, 162
Philbeck, J., 145, 164
Popescu, B. E., 112, 162
Priestley, H., 128, 161
Pruzansky, S., 12, 134, 142, 160, 166
- Rabe-Heskett, S., 152, 167
Ramsay, J. O., 64, 66, 79, 80, 166, 168
Rannala, B., 66, 163

- Restle, F., 7, 9, 24, 55, 59, 79, 87, 129, 166, 176
Rodgers, J. L., 164
Ronacher, B., 115, 117, 161, 166
Rosenberg, S., 54, 166
Roskam, E., 115, 161
Rothkopf, E. Z., 42, 166
Ruts, W., 142, 168
Rzhetsky, A., 53, 81, 166, 167
- Saitou, N., 52, 54, 70, 142, 165, 166
Sattath, S., 5, 6, 54, 84, 117, 134, 142, 145,
147, 154, 166
Savage, C., 91, 112, 166
Schott, B., 2, 166
Schulze, H., 142, 160
Schwanenflugel, P., 130, 166
Self, S. G., 53, 150, 151, 166
Sen, P. K., 16, 22, 149, 166
Sergent, J., 5, 14, 168
Shepard, R. N., 5, 21–23, 42, 84, 86, 100,
115–117, 131, 133, 134, 137, 140,
143, 166, 167
Shimodaira, H., 53, 166
Silvapulle, M. J., 16, 22, 149, 166
Singh, K., 151, 153, 164
Sitnikova, T., 81, 167
Skrondal, A., 152, 167
Smith, H., 99, 161
Soli, S. D., 133, 167
Stanton, R., 115, 168
Stephens, J. C., 52, 165
Stevens, J., 152, 167
Storms, G., 142, 148, 168
Stram, D. O., 53, 167
Suppes, P., 119, 167
Swayne, D., 116, 120, 159
- Tajima, F., 53, 80, 167
Takane, Y., 5, 14, 66, 79, 167, 168
Tatsuoka, K. K., 50, 168
Ten Berge, J. M. F., 133, 168
Theil, H., 28, 168
Tibshirani, R. J., 17, 30–33, 41, 67, 69, 93, 94,
96, 112, 150, 151, 161, 163, 168, 169
Torgerson, W., 26, 118, 168
Tversky, A., 2, 4–6, 12, 15, 19, 21, 22, 42, 52,
54, 59, 70, 84, 117, 119, 129–131,
133, 134, 142, 143, 145, 147, 154,
160, 162, 166–168, 175
- van der Ark, L., 50, 165
van der Heijden, P. G. M., 50, 165
Vanoverberghe, V., 142, 168
Verbeemen, T., 142, 148, 168
Verguts, T., 148, 168
- Waddell, P. J., 53, 166
Wasserman, S., 10, 168
Watson, G. S., 152, 161
Wedin, P.-Å., 153, 162
Weinberg, S. L., 38, 168
Werner, H. J., 80, 153, 168
White, H., 151, 153, 162
Wilf, H. S., 72, 93, 112, 165
Winsberg, S., 66, 168
Wolak, F. A., 49, 65, 66, 168
Wollan, P. C., 29, 64, 149, 168
- Yang, Y., 113, 168
Yapar, C., 80, 153, 168
- Zaki, S., 115, 116, 148, 166, 168
Zielman, B., 147, 169
Zou, H., 96, 97, 112, 169

Subject Index

- additive clustering, 14, 20, 115, 117, 127, 133, 134, 138, 148
 - ADCLUS, 5, 84
 - CLUSTREES, 5
 - MAPCLUS, 5, 84
 - three-way, 133
- additivity, 13, 21, 22, 28, 84, 115, 116, 118, 123, 129, 143, 148, 154
 - distinctive-feature, 28, 63
 - inter-dimensional, 117, 119, 144
 - intra-dimensional, 118, 119
 - metric-segmental, 13, 117, 119, 120
 - of distance, 115, 118, 143
 - segmental, 123
- betweenness, 9, 10, 12, 13, 59, 117, 118, 128–130, 179
 - intra-dimensional, 118
 - lattice, 127–129, 135
 - metric-segmental, 118, 127–129
- bias, 28, 32–34, 41, 44, 45, 49, 53, 69, 71, 74–76, 79, 94, 96, 154
 - relative, 41
- binary code, 6, 72, 85, 90–93, 112
- city-block
 - configuration, 119–123, 125, 127, 128
 - coordinates, 127, 143
 - dimensions, 116, 118
 - metric, 7, 13, 19, 21, 22, 24, 91, 115–118
 - model, 13, 19, 115–118, 127, 130, 144–146, 148, 154
 - space, 13, 118–120, 127, 128, 132, 143
- clique, 117
- cluster differences scaling, 27
 - algorithm, 27, 54, 85, 112
- common features model (CF), 5, 6, 14, 20, 49, 88, 113, 115, 117, 133–138, 145–147, 154, 176, 179
- confidence interval, 14, 16, 19, 23, 26, 31, 33, 46, 48, 49, 51–53, 61, 65, 68, 69, 78, 80, 81, 148, 149
 - BC_a , 34, 41, 49
 - t -, 16, 56, 61, 63, 65, 72, 76, 79, 89, 150, 152
 - bootstrap, 32, 34, 38, 54, 68, 76, 150
 - bootstrap- t , 32, 41, 46
 - branch length, 53
 - empirical, 34, 69
 - nominal, 32–34, 38, 46, 48, 69, 76
- Contrast Model (CM), 2–5, 14, 22, 52, 84, 142, 147, 148, 154, 175, 176
- coordinate-free representation, 119, 144
- coverage, 38, 41, 45, 46, 48, 49, 54, 68, 69, 71, 72, 76, 78, 79, 82, 150
- cross-validation, 85
 - K-fold, 102
 - leave-one-out, 63, 66, 80
- dissimilarity, 2, 3, 14, 31, 32, 39, 49, 69, 84, 104, 121, 138, 139
 - coefficient, 22, 84
 - data, 1, 14, 115, 152
 - distribution, 142
 - judgement of, 4
 - matrix, 54, 138, 139, 147, 152
 - measure, 54
 - simulated, 39
 - value, 1, 2, 7, 8, 14, 54, 68, 70, 71, 80, 151, 152
- distance
 - center, 124

- city-block, 13, 52, 115, 116, 118, 120, 123, 124, 126, 127, 130, 132, 143, 154
- Euclidean, 11, 19, 21, 61, 116, 132, 139, 155
- evolutionary, 53, 79, 81
- feature, 7–10, 13, 14, 17, 24–26, 55, 56, 59, 60, 87, 89, 91, 129, 132, 135, 136, 139, 142, 155
- featurewise, 60, 64, 88, 92–94, 99, 111, 153
- Hamming, 7, 22, 24, 52, 55, 87
- internodal, 80
- metric, 3, 8, 51
- path-length, 8, 25, 49, 59, 89
- set-theoretic, 128
- shortest path, 8
- distinctive features model (DF), 5, 6, 13, 20, 49, 84, 113, 115, 127–132, 134–138, 140, 142, 143, 145–148, 154
- dominance metric, 116
- Dijkstra algorithm (AS 225), 29, 30, 64, 149
- edge, 2, 8–15, 17–19, 25, 26, 28, 29, 52, 53, 56–60, 63, 64, 84, 88, 89, 94, 117, 119, 120, 123–125, 132, 135, 144, 145, 155
 - deletion method, 132
 - length, 124, 132, 145, 150
 - unique, 135
- feature, 1, 2
 - cluster, 58–60, 64, 73, 76–78, 80, 81, 156
 - common, 3–6, 84, 92, 113, 137, 154, 156
 - distinctive, 3–10, 13, 17, 18, 22, 25, 49, 56, 84, 85, 88, 92, 99, 102, 104, 106, 108, 111, 113, 117, 129, 147, 153, 154
 - distinctive, complete set, 72
 - matrix, 2, 3, 11, 12, 17, 24, 30, 31, 40, 54, 58, 60, 68, 70, 71, 81, 87, 88, 93, 99, 101, 103, 111, 128, 134, 135
 - network, 1, 13, 14, 16, 19, 26, 56, 89, 131, 134, 138, 139, 155, 156
 - non-common, 55
 - structure, 12, 13, 16, 19, 41, 42, 52, 59–62, 70, 71, 101, 130, 138, 142, 145, 148, 150, 153–156
 - structure (a priori known), 53, 63
 - structure (nested), 12, 58
 - structure (overlapping), 12
 - unique, 58–60, 64, 73, 76, 78, 92, 131–140, 142, 143, 146, 147, 156
 - universal, 43, 92
- FNM (Feature Network Models), 1, 2, 5–8, 10, 12, 14–19, 21–23, 25–28, 30, 40, 42, 48–57, 59, 60, 69, 70, 72, 81, 83–89, 91–94, 96, 99, 111, 112, 149–156, 175–179
- PROXGRAPH, 14, 15, 29, 56, 60, 70, 71, 88, 92, 93
- generalized cross-validation error, 63, 67
- generalized cross-validation statistic, 66, 67, 80, 156
- Gray code, 6, 17, 72, 83, 85, 90–93, 99, 100, 102, 104, 105, 107–113, 153, 154
 - binary reflected, 91
 - generate, 93, 111
 - generating, 72, 91
 - rank number, 72, 101, 104, 154
- grid, 18, 117
- hypercube, 13, 91, 92, 119
- hypercuboid, 13, 119
- Kuhn-Tucker
 - multipliers, 30, 64
 - test, 49, 53, 54, 63, 65–68, 74
- Lasso, 17, 85, 93, 94, 96–98, 112, 150, 153
 - constraint, 95
 - Least Angle Regression (LARS), 95–97
 - loss function, 94
 - positive, 17–19, 83, 85, 86, 92, 93, 96–103, 105, 106, 108, 109, 111–113, 153, 154, 156
 - shrinkage, 94
- lattice, 104, 111, 128, 130
- least squares, 21, 22, 38, 81, 84, 153
 - alternating, 130
 - estimator, 28
 - generalized, 53, 80, 152, 153
 - inequality constrained (ICLS), 16, 22, 23, 28, 29, 48, 50, 53, 54, 63, 64, 68, 80, 81, 98, 149, 150, 153
 - inequality constrained generalized (ICGLS), 80
 - loss function, 7, 25, 56
 - minimization, 52

- nonnegative, 14, 15, 29, 84, 88, 132, 140, 143
- nonnegative, loss function, 88, 93
- ordinary (OLS), 14, 29, 53, 65, 93, 94, 98, 149, 151
- ultrametric, 15

- Minkowski metric, 116
- model selection, 14, 17, 18, 85, 113
 - algorithm, 17, 85
- Modified Contrast Model (MCM), 5, 15, 18, 22, 84, 85, 129
- monotonicity, 119, 128
- multidimensional scaling (MDS), 3, 8, 10, 21, 22, 26, 38, 51, 79, 80, 83, 121, 129, 142, 145, 155
 - PROXSCAL, 10, 11, 13, 14, 18, 56, 61, 89, 132, 155
 - city-block, 132
 - nonmetric, 116

- network, 2, 8, 14, 28, 29, 52, 57, 117, 122–124, 137
 - complete, 8, 111, 118–120
 - connected, 104, 111
 - construction, 117, 120, 144, 145
 - coordinate-free, 144
 - embedding, 2, 10, 11, 127, 132, 143, 155
 - graph, 2, 8, 16, 57, 59
 - plot, 11, 105
 - representation, 2, 9, 10, 12–14, 17, 19, 22, 54, 56, 83–85, 88–90, 100, 104, 115, 118–120, 122, 123, 125–128, 131, 134, 136–138, 141, 143–145, 147, 148, 154, 155
 - representation (universal), 13, 19, 117
 - structure, 155
- node, 51, 57, 60, 61, 117, 119, 120, 122, 123, 126, 132, 133, 136, 137, 140, 142, 144, 145, 147
 - cluster, 140
 - external (leaf), 12, 57–59, 125, 126, 132, 135, 138
 - internal, 1, 2, 9–13, 57–61, 70, 71, 80, 115, 117, 118, 123–127, 130–137, 140, 142, 144, 147, 156, 178, 179
 - origin, 140

- partial isometry, 118, 126
- path-length metric, 49

- prediction error, 51, 53, 63, 66, 68, 71, 74, 77, 80, 85, 96–98, 101, 102, 107–110, 156
- proximity, 83
 - data, 8, 21–23, 51, 52, 54, 83, 86
 - transformation, 10, 13

- ridge regression, 94, 95, 150
- root mean squared error (*rmse*), 41, 44, 45, 48, 69

- sampling
 - bivariate, 30, 31
 - from binomial distribution, 38–40, 68, 69, 71, 104
 - from Gray codes, 101
 - multivariate, 30–32, 40, 69
 - pairs, 151
 - residuals, 31, 151
 - univariate, 31
- set-theoretical approach, 2, 3, 5, 21, 22
- shortest path, 120
- shrinkage, 93, 94, 96, 112
- similarity, 2–4, 21, 22, 42, 70, 84, 87, 115, 117, 118, 128, 129, 131, 133, 134, 137, 138, 143, 147, 148, 154
 - data, 19, 115, 117, 142
 - discrete models of, 127
 - distribution, 142
 - featural representation of, 115
 - judgement of, 3, 4, 42, 116
 - matrix, 136, 138, 147
 - measure, 23
 - value, 4, 70
- simplex, 124, 125
 - tree, 147
- social network models, 10
- standard errors, 14–16, 21, 22, 26, 31, 48, 49, 52, 53, 65, 68, 69, 80, 150
 - bootstrap, 23, 31, 32, 38, 41, 54, 68, 69, 74
 - branch lengths, 53
 - empirical, 16, 19, 21, 38, 42, 68
 - nominal, 32–34, 38, 40, 41, 44, 45, 48, 67–72, 74–76
 - theoretical, 15, 16, 19, 21, 23, 28–30, 32, 48, 49, 51, 53, 54, 63–65, 79–81, 84, 88, 89, 148–153
- star graph, 124, 125, 135, 140, 147

- statistical inference, 14, 16, 19, 22, 27, 48–54,
63, 67, 79, 80, 84, 88, 149, 153, 156
- subset selection, 17, 18, 85, 99, 149
algorithm, 18, 85
- symmetric set difference, 3, 5, 7, 8, 13, 24,
52, 55, 87, 93, 117, 127–129
- tree
- CLUSTREES, 5, 15, 22, 84
 - additive, 5, 12, 14–16, 19, 20, 25, 49,
51–54, 56–64, 67, 70, 74, 79–81, 84,
115, 117, 128, 130, 141–144,
146–148, 150, 151, 153, 156
 - algorithm, ADDTREE/P, 142
 - algorithm, ADDTREE, 54, 81, 142, 143
 - algorithm, GTREE, 142
 - algorithm, neighbor-joining (NJ), 54,
70, 81
 - bifurcating, 57, 58
 - bootstrap, 53, 81
 - consensus, 66
 - double, 147
 - double star, 117, 140, 141, 143, 147
 - embedding, 61
 - evolutionary, 53, 79
 - extended (EXTREE), 5, 12, 14, 15, 20, 22,
43, 49, 84, 115, 117, 130, 131, 134,
145, 146
 - graph, 12, 58, 59, 117, 142
 - hierarchical, 14, 20, 84, 115, 140, 142,
156
 - linear (chain), 147
 - multifurcating, 58, 60
 - multiple, 12, 49, 134, 140, 147
 - neighbor-joining, 73, 76–78
 - phylogenetic, 52–54, 63, 79–81, 156
 - reduced star, 147
 - resolved, 62, 74
 - simplex, 147
 - singular, 147
 - starlike, 81
 - statistical inference, 63
 - topology, 16, 19, 51–54, 60, 61, 63,
65–68, 70–72, 74, 76–81, 150, 154,
156
 - ultrametric, 49, 147
 - unresolved, 61, 62, 74, 140, 147
 - unrooted, 58
- triangle equality, 9, 10, 19, 118
algorithm, 14
- triangle inequality, 4, 8, 19, 25, 59, 89, 118
- uniqueness, 49, 116
coordinate system, 116, 143
dimensional, 116