Berrendero, José R.; **Cárcamo, Javier**
Linear components of quadratic classifiers. (English) Zbl 1459.62104
Adv. Data Anal. Classif., ADAC 13, No. 2, 347-377 (2019).

Summary: We obtain a decomposition of any quadratic classifier in terms of products of hyperplanes. These hyperplanes can be viewed as relevant linear components of the quadratic rule (with respect to the underlying classification problem). As an application, we introduce the associated multidirectional classifier; a piecewise linear classification rule induced by the approximating products. Such a classifier is useful to determine linear combinations of the predictor variables with ability to discriminate. We also show that this classifier can be used as a tool to reduce the dimension of the data and helps identify the most important variables to classify new elements. Finally, we illustrate with a real data set the use of these linear components to construct oblique classification trees.

MSC:

62H30 Classification and discrimination; cluster analysis (statistical aspects)

Keywords:

supervised classification; Fisher linear discriminant analysis; quadratic discriminant analysis; reduction of the dimension; feature extraction; oblique classification trees

Software:

R; caret; tree; UCI-ml

Full Text: DOI Link

References:

[1] Bache K, Lichman M (2013) UCI machine learning repository. http://archive.ics.uci.edu/ml

[2] Devroye L, Györfi L, Lugosi G (1996) A probabilistic theory of pattern recognition. Applications of mathematics (New York), vol 31. Springer, New York · Zbl 0853.68150 · doi:10.1007/978-1-4612-0711-5

[3] Fan, J.; Ke, ZT; Liu, H.; Xia, L., QUADRO: a supervised dimension reduction method via Rayleigh quotient optimization, Ann Stat, 43, 1498-1534, (2015) · Zbl 1317.62054 · doi:10.1214/14-AOS1307

[4] Friedman, JH, Regularized discriminant analysis, J Am Stat Assoc, 84, 165-175, (1989)· doi:10.1080/01621459.1989.10478752

[5] Golub GH, Van Loan CF (2013) Matrix computations, 4th edn. Johns Hopkins studies in the mathematical sciences. Johns Hopkins University Press, Baltimore

[6] Hand, DJ, Classifier technology and the illusion of progress, Stat Sci, 21, 1-34, (2006) · Zbl 1426.62188 · doi:10.1214/088342306000000060

[7] Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. Springer series in statistics, 2nd edn. Springer, New York · Zbl 1273.62005

[8] Huang, H.; Liu, Y.; Marron, JS, Bidirectional discrimination with application to data visualization, Biometrika, 99, 851-864, (2012) · Zbl 1452.62449 · doi:10.1093/biomet/ass029

[9] Kuhn, M., Building predictive models in R using the caret package, J Stat Softw, 28, 1-26, (2008)· doi:10.18637/jss.v028.i05

[10] Park SH, Fürnkranz J (2007) Efficient pairwise classification. In: European conference on machine learning. Springer, pp 658-665

[11] R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/

[12] Rifkin, R.; Klautau, A., In defense of one-vs-all classification, J Mach Learn Res, 5, 101-141, (2004) · Zbl 1222.68287

[13] Ripley B (2014) Tree: classification and regression trees. R package version 1.0-35. http://CRAN.R-project.org/package=tree

[14] Truong A (2009) Fast growing and interpretable oblique trees via logistic regression models. Doctoral dissertation, University of Oxford

[15] Wald, PW; Kronmal, R., Discriminant functions when covariances are unequal and sample sizes are moderate, Biometrics, 33, 479-484, (1977) · Zbl 0371.62091 · doi:10.2307/2529362

paper as accurately as possible without claiming the completeness or perfect precision of the matching.