Regularization in statistical models with multidimensional prediction functions

Dr. Matthias Schmid

Institut für Medizininformatik, Biometrie und Epidemiologie
Friedrich-Alexander-Universität Erlangen-Nürnberg

Boosting is one of the most important methods for fitting regression models and building prediction rules from high-dimensional data. A notable feature of boosting is that the technique has a built-in mechanism for shrinking coefficient estimates and variable selection. This regularization mechanism makes boosting a suitable method for analyzing data characterized by small sample sizes and large numbers of predictors. We extend the existing methodology by developing a boosting method for prediction functions with multiple components. Such multidimensional functions occur in many types of statistical models, for example in count data models and in models involving outcome variables with a mixture distribution. As will be demonstrated, the new algorithm is suitable for both the estimation of the prediction function and regularization of the estimates. In addition, nuisance parameters can be estimated simultaneously with the prediction function.