EFSTATHIA BURA
Department of Statistics, George Washington University, Washington DC

SUFFICIENT REDUCTIONS IN REGRESSION AND CLASSIFICATION

17.1.2017, 16.00 Uhr
MUW-Informatikbibliothek (88.03.806)

Gastgeber: Florian Frommlet

Abstract:
A main objective of statistical inference is the reduction of data: variables are replaced by relatively few quantities (reductions) which adequately represent the relevant information contained in the original data. Sufficient dimension reduction methodology (a) identifies and (b) provides estimation algorithms of sufficient reductions in regressions/classifications with many predictors. The reductions are sufficient or exhaustive in the sense that they are all that is needed to model and predict the response(s) and result in reducing big data to "analyzable" size.

An overview of the sufficient dimension reduction (SDR) methodology and its progress over the last two decades will be presented and juxtaposed with other data reducing methods. In particular, I will start from moment-based sufficient dimension reduction, its accomplishments and challenges, and move on to model-based sufficient reductions for regressions with predictors in the elliptically contoured family and multivariate exponential family of distributions. The latter includes regressions where predictors are all continuous, all categorical or mixtures of the two. The minimal sufficient reduction of the predictors and its maximum likelihood estimator are derived by modeling the conditional distribution of the predictors given the response.

Whereas nearly all extant estimators of sufficient reductions are linear and only partly capture the sufficient reduction, model-based SDR identifies both the linear and the nonlinear components of the sufficient reduction. It also provides the exact form of the sufficient reduction, which is exhaustive, its maximum likelihood estimate via an iterative re-weighted least squares (IRLS) estimation algorithm and asymptotic tests for the dimension of the regression.