

Very High-Dimensional Data: Greedy Boosting and Convex Lasso-Relaxation

Peter BÜHLMANN, *ETH Zürich, Switzerland*, E-mail: buhlmann@stat.math.ethz.ch

Abstract: We consider data problems whose dimension (e.g. of the predictor or response) is very large relative to sample size. The setting includes high-dimensional multivariate regression, multi-category classification and graphical models. When the dimension is in the ten-thousands, either greedy methods or convex optimization are computationally attractive. We focus on greedy boosting algorithms and Lasso estimation (which is convex). We discuss (i) applications; (ii) asymptotic results for very high-dimensional but sparse settings; (iii) efficient and feasible computations for solutions which are closer to the often intractable subset variable selection problem (i.e. ℓ_0 -penalty problem) than the corresponding Lasso or Boosting estimates.