Dyadic Thinking about Density Estimation and Tree Classifiers

Robert Nowak
Rice University

Our recent work in tree-based, complexity-penalized likelihood estimation extends the well known near minimax optimality results for (wavelet-based) Gaussian denoising to a broader class of observation models, including multinomial models arising in density estimation. The key to these estimators is the exploitation of recursive dyadic partitions of the data space.

Even more recently, we have begun to investigate the application of similar tree-based techniques for constructing classifiers from training data. A crucial difference between tree-based, complexity-penalized estimation and classification is that in the estimation problem the proper penalty is proportional to the number of leaves, while in the classification problem the proper penalty is proportional to the square-root of the number of leaves. This distinction makes classifier design a more challenging problem than estimator design. I will describe a quadratic-time algorithm that computes the optimal tree classifier. I will also discuss some results concerning the rate of convergence of optimal tree classifiers to the Bayes decision boundary. Under mild regularity conditions, our results demonstrate that the adaptive resolution of tree classifiers enables them to focus on the lower dimensional decision boundary, rather than the higher dimensional density, providing much faster rates of convergence than classifiers based on density estimation in certain cases.