



Aggregation of estimators and high-dimensional statistics

Alexandre Tsybakov
CREST-ENSAE - alexandre.tsybakov@ensae.fr

This talk gives an overview of recent results on aggregation of estimators with an emphasis on sparsity issues. Given a collection of estimators, the problem of linear, convex or model selection type aggregation consists in constructing a new estimator, called the aggregate, which is nearly as good as the best among them (or nearly as good as their best linear or convex combination), with respect to a given risk criterion. This problem can be also considered with a given dictionary of functions instead of estimators. When the underlying model is sparse, which means that it is well approximated by a linear combination of a small number of functions in the dictionary, the aggregation techniques turn out to be very useful in taking advantage of sparsity. Aggregates are usually constructed by mixing the initial estimators or functions of the dictionary with data-dependent weights. An important example that we will focus on is sparsity pattern aggregate with exponential weights. It has some remarkable properties. In particular, it satisfies the sharpest known sparsity oracle inequalities and allows one to treat in a united way the settings with ordinary, group and fused sparsity.