Optimal Kullback-Leibler Aggregation in Mixture Density Estimation by Maximum Likelihood

Mehdi Sebbar

We consider the problem of density estimation under the assumption that it is well approximated by a mixture with a large number of components. The statistical properties of the maximum likelihood estimator are explored with respect to the Kullback-Leibler loss. We establish risk bounds taking the form of sharp oracle inequalities both in deviation and in expectation. A simple consequence of these bounds is that the maximum likelihood estimator attains the optimal rate, up to a logarithmic factor, in the problem of convex aggregation. This rate is of order $((\log K)/n)^{1/2}$, where K is the number of mixture components and n is the sample size. More importantly, under the additional assumption that the Gram matrix of the components satisfies the compatibility condition, the obtained oracle inequalities yield the optimal rate in the sparsity scenario. That is, if the weight vector is (nearly) D-sparse, we get the rate $(D \log K)/n$. Joint work with Arnak Dalalyan.