

Performance limits of nonparametric estimators

Ph.D. thesis summary

Andras Antos

The thesis focuses on the theoretical limits of nonparametric statistical estimation. Namely, we are concerned with regression function estimation, pattern recognition, and estimation of some functionals of distributions.

In all cases our aim is to estimate some characteristics of the unknown distribution of a random variable based on independent samples. Several results are known which guarantee consistency of the estimator and bound the rate of convergence of the expected loss as the sample size increases.

The sharpness of these results can be shown by proving corresponding lower bounds or sometimes the nonexistence of universal rates of convergence. Thus we examine the worst-case loss of the best estimate. In typical such *minimax* lower bounds, the “bad” distribution may change with the sample size. The rate of convergence for every single distribution can be faster than that of the worst-case loss. The *individual* lower bounds, which we are interested in, show the existence of a fixed distribution that may be “bad” for an arbitrary large sample size.

For a parametric class of distributions, one can construct usually a consistent estimate such that the optimal minimax and individual rates of convergence are inversely proportional to the square root of the sample size.

In the nonparametric case, however, there is no guaranteed rate of convergence for rich classes of distributions. On the other hand, the optimal minimax and individual rates of convergence can be different. While the minimax rate of convergence gives the worst-case loss for every sample size, the individual rate of convergence gives the slowest one among the effective rates of convergence for the particular distributions.

In the Introduction, we summarize related results concerning distribution and density estimation. For regression function estimation and pattern recognition, we extend known minimax lower bounds regarding certain classes of distributions to individual ones. In the case of the expectation, entropy, mutual information and Bayes-error, we prove that, while they have consistent estimates, there is no universal rate of convergence, which can also be interpreted as an individual lower bound.