

PAC-Bayesian Estimation for High-Dimensional Multi-Index Regression with Unknown Active Dimension

Abstract

Increasing computational power and storage capacity have made high-dimensional datasets accessible to many areas of research such as medicine, natural and social sciences. While classical statistical methods are not compatible with high-dimensional data, especially due to the curse of dimensionality, machine learning methods have been successfully applied in practice. On the theoretical level, a popular way to circumvent the curse of dimensionality is the multi-index model. Building on the single-index analysis by Alquier and Biau (2013), we develop a PAC-Bayesian estimation method for a possibly miss-specified multi-index model with unknown active dimension and an orthogonal dimension reduction matrix. Our main result is a non-asymptotic oracle inequality, which shows that the estimation method adapts to the active dimension of the model, the sparsity of the dimension reduction matrix and the regularity of the link function. Under a Sobolev regularity assumption on the link function the estimator achieves the usual minimax rate of convergence and no additional price is paid for the unknown active dimension.

This presentation is based on joint work with Mathias Trabs from Universität Hamburg.

References

P. Alquier, G. Biau (2013): Sparse single-index model. *Journal of Machine Learning Research* 14(Jan), 243–280.