

Smoothing Spline Modeling via Adaptive Bases Sampling

Ping Ma
Department of Statistics
University of Illinois at Urbana-Champaign
pingma@illinois.edu

Smoothing splines provide flexible nonparametric regression estimators. However, the computational cost of smoothing splines is high, which has hindered their wide applications. In this talk, I present a new method named adaptive basis sampling (ABS) for efficient computation of smoothing splines in super-large samples. The key idea of the proposed method is to estimate smoothing splines in a lower dimensional effective model space, which is constructed through an adaptive sampling method based on the scale of the response. The sampling strategy gives rise to more scalable computation. Our asymptotic analysis shows smoothing splines computed via adaptive basis sampling converge to the true function at the same rate as regular smoothing splines. The excellent empirical performance of the proposed method is demonstrated through simulation studies and real data analysis.

The talk is based on joint works with Jianhua Huang, Nan Zhang and Wenxuan Zhong.

Key Words: Nonparametric regression, big data, computation