

Automatic Model Structure Selection for Multivariate Nonparametric Regression

Hao Helen Zhang *

University of Arizona, Tucson, AZ, USA hzhang@math.arizona.edu

Guang Cheng

Purdue University, West Lafayette, chengg@purdue.edu

Yufeng Liu

University of North Carolina, Chapel Hill, yfliu@email.unc.edu

Health Multivariate nonparametric regression models provide useful statistical analysis tools in practice due to their high flexibility and powerful ability to capture complex patterns hidden in data. One fundamental issue and open problem associated with multivariate nonparametric regression models is how to overcome the curse of dimensionality. It turns out, if the refined model structure, such as sparsity, linearity and low dimensional structures, can be utilized, then the model estimation and prediction efficiency can be greatly improved. Most existing methods for structure selection and estimation are largely ad hoc and lack of theoretical support. In this paper, we tackle the structure detection problem from the new perspective of model selection. A unified regularization framework in reproducing kernel Hilbert space (RKHS) is developed to automatically identify sparse, linear and low-dimensional structure of the model. Asymptotic results suggest that the proposed procedure can discover the underlying true model structure correctly as the sample size goes to infinity. Numerical examples are given to illustrate performance of the proposed procedure.

Key Words: Low dimensional structure, model selection, sparsity, reproducing kernel Hilbert space, regularization, nonparametric regression