Consistent variable selection in additive models

by

Lan Xue Oregon State University Department of Statistics, OSU Covallis, Oregon 97330 xuel@stat.orst.edu

Abstract

A penalized polynomial spline method will be introduced for simultaneous model estimation and variable selection in additive models. The proposed method approximates the nonparametric functions by polynomial splines, and minimizes the sum of squared errors subject to an additive penalty on norms of spline functions. This approach sets estimators of certain function components to exactly zero, thus performing variable selection. Under mild regularity conditions, I show that the proposed method estimates the non-zero function components in the model with the same optimal mean square convergence rate as the standard polynomial spline estimators, and correctly set the zero function components to zero with probability approaching one, as sample size goes to infinity. The theoretical results are well supported by simulation studies. The proposed method is also applied to two real data examples for illustration.