Ann Inst Stat Math (2019) 71:657–677
https://doi.org/10.1007/s10463-018-0654-0

Spline estimator for ultra-high dimensional partially linear varying coefficient models

Zhaoliang Wang[1,2] · **Liugen Xue**[1] · **Gaorong Li**[3] · **Fei Lu**[1]

Received: 10 June 2017 / Revised: 31 December 2017 / Published online: 13 March 2018

Abstract In this paper, we simultaneously study variable selection and estimation problems for sparse ultra-high dimensional partially linear varying coefficient models, where the number of variables in linear part can grow much faster than the sample size while many coefficients are zeros and the dimension of nonparametric part is fixed. We apply the B-spline basis to approximate each coefficient function. First, we demonstrate the convergence rates as well as asymptotic normality of the linear coefficients for the oracle estimator when the nonzero components are known in advance. Then, we propose a nonconvex penalized estimator and derive its oracle property under mild conditions. Furthermore, we address issues of numerical implementation and of data adaptive choice of the tuning parameters. Some Monte Carlo simulations and an application to a breast cancer data set are provided to corroborate our theoretical findings in finite samples.

Keywords High dimensionality · Partially linear varying coefficient model · Variable selection · Nonconvex penalty · Oracle property

✉ Zhaoliang Wang
wangzhaoliang@hpu.edu.cn

1 College of Applied Sciences, Beijing University of Technology, Beijing 100124, China

2 School of Mathematics and Information Science, Henan Polytechnic University, Jiaozuo 454000, China

3 Beijing Institute for Scientific and Engineering Computing, Beijing University of Technology, Beijing 100124, China