Statistica Sinica **33** (2023), 1-22
doi:https://doi.org/10.5705/ss.202021.0073

ORACLE-EFFICIENT GLOBAL INFERENCE FOR VARIANCE FUNCTION IN NONPARAMETRIC REGRESSION WITH MISSING COVARIATES

Li Cai and Suojin Wang

Zhejiang Gongshang University and Texas A&M University

Abstract: We propose a new bias-corrected spline-kernel estimator and a smooth simultaneous confidence band (SCB) as a global inference tool for the conditional variance function in a nonparametric regression when the covariates are missing at random. To adapt to the possible missingness of the covariates, we employ a Horvitz–Thompson-type weighted spline smoothing to fit the nonparametric regression function. Based on the squared residuals, the weighted kernel method is then applied to estimate the variance function. Synthesizing the spline smoothing and kernel regression in one estimator takes advantage of the fast computing speed of the spline regression, and of the flexible local estimation and easy SCB constructions of the kernel smoothing. The proposed estimator is shown to be oracle-efficient in the sense that it is as efficient as the ideal one when the mean function and the selection probabilities are known by the "oracle", which we use to establish an asymptotically correct SCB for the variance function. The findings of our empirical finite-sample studies support our asymptotic theory. An application to a data set from the Canada 2010/2011 Youth Student Survey illustrates the usefulness of the proposed techniques.

Key words and phrases: B-spline regression, local linear regression, missing at random, oracle efficiency, simultaneous confidence band.

1. Introduction

Variance function estimation is an important procedure in many statistical analyses, such as stochastic control, risk analysis, the construction of confidence intervals for a regression function, and the estimation of smoothing parameters. Research on inferences for the variance function includes the works of Hall and Carroll (1989) and Wang et al. (2008), who studied the effect of the unknown mean on the estimation of the variance function, and Müller and Stadtmüller (1987, 1993), Brown and Levine (2007), and Cai and Wang (2008), who considered difference-based adaptive nonparametric estimators of the variance function.

Corresponding author: Suojin Wang, Department of Statistics, Texas A&M University, College Station, TX 77843, USA. E-mail: sjwang@stat.tamu.edu.