Statistica Sinica **32** (2022), 2359-2380
doi:https://doi.org/10.5705/ss.202020.0404

EFFICIENT ESTIMATION FOR DIMENSION REDUCTION WITH CENSORED SURVIVAL DATA

Ge Zhao, Yanyuan Ma and Wenbin Lu

Portland State University, Penn State University, and North Carolina State University

Abstract: We propose a general index model for survival data, that generalizes many commonly used semiparametric survival models and belongs to the framework of dimension reduction. Using a combination of a geometric approach in semiparametrics and a martingale treatment in survival data analysis, we devise estimation procedures that are feasible and do not require covariate-independent censoring, as assumed in many dimension-reduction methods for censored survival data. We establish the root-n consistency and asymptotic normality of the proposed estimators and derive the most efficient estimator in this class for the general index model. Numerical experiments demonstrate the empirical performance of the proposed estimators, and an application to an AIDS data set further illustrates the usefulness of the work.

Key words and phrases: Dimension reduction, general index model, kernel estimation, semiparametric theory, survival analysis.

1. Introduction

The Cox proportional hazards model (Cox (1972)) is probably the most widely used semiparametric model for analyzing survival data. In the Cox model, the covariate effect is described by a single linear combination of the covariates in an exponential function, and is multiplicative in modeling the hazard function. Although this special way of modeling the hazard function permits a convenient estimation procedure, such as the maximum partial likelihood estimation (Cox (1975)), it has limitations. As widely studied in the literature, there are many situations where the Cox model may not be proper. Owing to the limitations of the Cox model, many other semiparametric survival models have been proposed, such as the accelerated failure time model (Buckley and James (1979)), proportional odds model (McCullagh (1980)), and linear transformation model (Dabrowska and Doksum (1988)), among others. Despite of all these efforts, the link between the summarized covariate effect, typically in the form of a linear

Corresponding author: Ge Zhao, Department of Mathematics and Statistics, Portland State University, Portland, OR 97201, USA. E-mail: gzhao@pdx.edu.