Broken adaptive ridge regression for right-censored survival data

Zhihua Sun1 · Yi Liu1 · Kani Chen2 · Gang Li3

Received: 1 February 2020 / Revised: 29 December 2020 / Accepted: 18 January 2021 / Published online: 5 April 2021
© The Institute of Statistical Mathematics, Tokyo 2021

Abstract
Broken adaptive ridge (BAR) is a computationally scalable surrogate to $L_0$-penalized regression, which involves iteratively performing reweighted $L_2$ penalized regressions and enjoys some appealing properties of both $L_0$ and $L_2$ penalized regressions while avoiding some of their limitations. In this paper, we extend the BAR method to the semi-parametric accelerated failure time (AFT) model for right-censored survival data. Specifically, we propose a censored BAR (CBAR) estimator by applying the BAR algorithm to the Leurgan’s synthetic data and show that the resulting CBAR estimator is consistent for variable selection, possesses an oracle property for parameter estimation and enjoys a grouping property for highly correlation covariates. Both low- and high-dimensional covariates are considered. The effectiveness of our method is demonstrated and compared with some popular penalization methods using simulations. Real data illustrations are provided on a diffuse large-B-cell lymphoma data and a glioblastoma multiforme data.

Keywords Accelerated failure time model · Grouping effect · $L_0$ penalization · Right censoring · Variable selection
Authors and Affiliations

Zhihua Sun¹ · Yi Liu¹ · Kani Chen² · Gang Li³

Zhihua Sun
zhihuasun@ouc.edu.cn

Yi Liu
liuyi@amss.ac.cn

Kani Chen
makchen@ust.hk

¹ Department of Mathematics, Ocean University of China, Qingdao 266000, China
² Department of Mathematics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong
³ Biostatistics and Computational Medicine, University of California, Los Angeles, CA 90095-1772, USA