Annals of the Institute of Statistical Mathematics (2022) 74:869–893
https://doi.org/10.1007/s10463-021-00816-0

Multi-round smoothed composite quantile regression for distributed data

Fengrui Di[1] · **Lei Wang**[1]

Received: 12 May 2021 / Revised: 25 September 2021 / Accepted: 26 November 2021 /
Published online: 10 January 2022

Abstract

Statistical analysis of large-scale dataset is challenging due to the limited memory constraint and computation source and calls for the efficient distributed methods. In this paper, we mainly study the distributed estimation and inference for composite quantile regression (CQR). For computational and statistical efficiency, we propose to apply a smoothing idea to the CQR loss function for the distributed data and then successively refine the estimator via multiple rounds of aggregations. Based on the Bahadur representation, we derive the asymptotic normality of the proposed multi-round smoothed CQR estimator and show that it also achieves the same efficiency of the ideal CQR estimator by analyzing the entire dataset simultaneously. Moreover, to improve the efficiency of the CQR, we propose a multi-round smoothed weighted CQR estimator. Extensive numerical experiments on both simulated and real data validate the superior performance of the proposed estimators.

Keywords Bahadur representation · Composite quantile regression · Divide-and-conquer · Multiple rounds · Kernel smoothing · Weighted composite quantile regression

✉ Lei Wang
lwangstat@nankai.edu.cn

[1] School of Statistics and Data Science & LPMC, Nankai University, No. 94 Weijin Road, Tianjin 300071, China