



Error density estimation in high-dimensional sparse linear model

Feng Zou¹ · Hengjian Cui¹

Received: 19 March 2018 / Revised: 18 September 2018 / Published online: 16 November 2018
© The Institute of Statistical Mathematics, Tokyo 2018

Abstract

This paper is concerned with the error density estimation in high-dimensional sparse linear model, where the number of variables may be larger than the sample size. An improved two-stage refitted cross-validation procedure by random splitting technique is used to obtain the residuals of the model, and then traditional kernel density method is applied to estimate the error density. Under suitable sparse conditions, the large sample properties of the estimator including the consistency and asymptotic normality, as well as the law of the iterated logarithm are obtained. Especially, we gave the relationship between the sparsity and the convergence rate of the kernel density estimator. The simulation results show that our error density estimator has a good performance. A real data example is presented to illustrate our methods.

Keywords High-dimensional sparse linear model · Kernel density estimation · Refitted cross-validation method · Asymptotic properties · Law of the iterated logarithm

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10463-018-0699-0>) contains supplementary material, which is available to authorized users.

✉ Hengjian Cui
hjcui@bnu.edu.cn

Feng Zou
zoufengjiayou@163.com

¹ Department of Statistics, School of Mathematical Sciences, Capital Normal University, No.105, West Third Ring Road, Haidian District, Beijing 100048, China