Annals of the Institute of Statistical Mathematics (2020) 72:45–63
https://doi.org/10.1007/s10463-018-0682-9

Discovering model structure for partially linear models

Xin He[1] · **Junhui Wang**[2]

Received: 6 December 2017 / Revised: 4 June 2018 / Published online: 30 July 2018

Abstract
Partially linear models (PLMs) have been widely used in statistical modeling, where prior knowledge is often required on which variables have linear or nonlinear effects in the PLMs. In this paper, we propose a model-free structure selection method for the PLMs, which aims to discover the model structure in the PLMs through automatically identifying variables that have linear or nonlinear effects on the response. The proposed method is formulated in a framework of gradient learning, equipped with a flexible reproducing kernel Hilbert space. The resultant optimization task is solved by an efficient proximal gradient descent algorithm. More importantly, the asymptotic estimation and selection consistencies of the proposed method are established without specifying any explicit model assumption, which assure that the true model structure in the PLMs can be correctly identified with high probability. The effectiveness of the proposed method is also supported by a variety of simulated and real-life examples.

Keywords Lasso · Gradient learning · Partially linear models · Proximal gradient descent · Reproducing kernel Hilbert space (RKHS)

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10463-018-0682-9) contains supplementary material, which is available to authorized users.

✉ Xin He
xinhe6-c@my.cityu.edu.hk

Junhui Wang
j.h.wang@cityu.edu.hk

[1] School of Statistics and Management, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai 200433, China

[2] Department of Mathematics, City University of Hong Kong, 83 Tat Chee Ave, Kowloon Tong 999077, Hong Kong, China