Accepted Manuscript

Estimation of partially linear regression models under the partial consistency property

Xia Cui, Ying Lu, Heng Peng

PII: S0167-9473(17)30099-3
DOI: http://dx.doi.org/10.1016/j.csda.2017.05.004
Reference: COMSTA 6464

To appear in: *Computational Statistics and Data Analysis*

Received date : 11 July 2014
Revised date : 12 May 2017
Accepted date : 15 May 2017

Please cite this article as: Cui, X., et al., Estimation of partially linear regression models under the partial consistency property. *Computational Statistics and Data Analysis* (2017), http://dx.doi.org/10.1016/j.csda.2017.05.004

Estimation of Partially Linear Regression Models under the Partial Consistency Property

Xia Cui[a], Ying Lu[b,*], Heng Peng[c]

[a]*School of Mathematics and Information Science, Guangzhou University, Guangzhou, China. Email: cuixia@gzhu.edu.cn*
[b]*Department of Applied Statistics, Social Science and Humanities, Steinhardt School of Culture, Education and Human Development,New York, USA. Email: ying.lu@nyu.edu New York University, New York, USA*
[c]*Department of Mathematics, The Hong Kong Baptist University, Kowloon Tong, Hong Kong. Email: hpeng@math.hkbu.edu.hk*

Abstract

Utilizing recent theoretical results in high dimensional statistical modeling, a flexible yet computationally simple approach is proposed to estimate the partially linear models. Motivated by the partial consistency phenomena, the nonparametric component in the partially linear model is modeled via incidental parameters and estimated by a simple local average over small partitions of the support of the nonparametric variables. The proposed least-squares based method seeks to strike a balance between computation burden and efficiency of the estimators while minimizing model bias. It is shown that given inconsistent estimators of the nonparametric component, square root-n consistent estimators of the parameters of the parametric component can be obtained with little loss in efficiency. Moreover, conditional on the parametric estimates, an optimal estimator of the nonparametric component can be obtained using classic nonparametric methods. The statistical inference problems regarding the parametric parameters and a two-population nonparametric testing problem regarding the nonparametric component are considered. The results show that the behavior of the test statistics is satisfactory. To assess the performance of the new method in comparison with other methods, three simulation studies are conducted and a real data set about risk factors of birth weights is analyzed.

*Correspondence to: 246 Greene Street, New York, NY 10003. Tel: 1-212-998-5560.