Statistica Sinica **36** (2026), 1-25
doi:https://doi.org/10.5705/ss.202023.0067

MODEL AVERAGING ESTIMATION FOR PARTIALLY LINEAR FUNCTIONAL SCORE MODELS

Shishi Liu[1], Chunming Zhang[2], Hao Zhang[3], Rou Zhong[3]
and Jingxiao Zhang*[3]

[1]*Hangzhou Dianzi University,* [2]*University of Wisconsin*
and [3]*Renmin University of China*

Abstract: The scalar-on-function regression is quite useful for modelling mixed-data in the context of scalar and functional variables. Under this class of regression, the paper aims at proposing a compelling alternative to model selection methods to address model selection uncertainty. The considered models characterize a scalar response using parametric effect of the scalar predictors and nonparametric effect of a functional predictor, and a model averaging estimation is developed based on Mallows-type criterion to assign weights for averaging. Further, the asymptotic optimality of the resulting estimator, in terms of achieving the smallest possible squared error loss, is established. Besides, simulation studies demonstrate its superiority to or comparability with some information criterion score-based model selection and averaging estimators. The proposed procedure is also applied to a mid-infrared spectra dataset for illustration.

Key words and phrases: Functional data, mallows-type criterion, model average.

1. Introduction

Functional data analysis (FDA) has received growing attention in recent decades due to its remarkable flexibility and widespread applicability in handling complex data, including variables defined on a continuum, such as time or space. Ramsay and Silverman (2005) offers a comprehensive introduction to FDA across various fields. One of the most extensively studied topics in FDA pertains to functional regression. A large body of research has been dedicated to developing regression models that incorporate functional predictors, with a predominant focus on functional parametric regression models, see Cardot, Ferraty and Sarda (1999, 2003), Yao, Müller and Wang (2005) and Cai and Hall (2006). These studies, falling into the category of scalar-on-function regression, assume specific forms of the regression model, such as functional linear model (FLM). The classic FLM is formulated to depict a linear relationship between a scalar response and a functional predictor, which is conventionally expressed as

$$Y = \beta_0 + \int_{\mathcal{T}} X(t)\beta(t)dt + \varepsilon, \tag{1.1}$$

*Corresponding author. E-mail: zhjxiaoruc@163.com