Statistica Sinica **35** (2025), 1-24
doi:https://doi.org/10.5705/ss.202022.0379

GENERALIZED FUNCTIONAL FEATURE REGRESSION MODELS

Qingzhi Zhong, Wei Liu, Li Liu, Hua Liang and Huazhen Lin*

Jinan University, Sichuan University, Wuhan University, George Washington University and Southwestern University of Finance and Economics

Abstract: The existing methods for functional regression can be roughly divided into two categories: direct functional regression (DFR) and functional regression based on functional principal component analysis (FR-FPCA). DFR may contain too much noise, while FR-FPCA may be inefficient because FPCA is independent of the response. In this paper, we investigate the effect of a vector of random curves on a response by extracting the latent features of the random curves that are associated with the response. Furthermore, to improve flexibility and predictive accuracy, we propose a generalized additive multiple index model that captures the relationship between the latent features and the response, without specifying component and link functions. We form an objective function based on a penalized quasi-likelihood function and FPCA to extract features, and to estimate the parameters and functions. We further develop an iterative algorithm, which is proven to be convergent and can expediently implement the proposed procedures. The convergence rates, oracle property, selection consistency and asymptotic normality for the proposed estimators are established. Numerical studies including extensive simulation experiments and two empirical applications show that the proposed procedures and methodology outperform the existing methods in interpretability, predictive accuracy and computation.

Key words and phrases: Functional principal component analysis (FPCA), generalized additive functional regression model (GAFRM), generalized linear functional regression (GLFR), group-SCAD penalty, penalized quasi-likelihood .

1. Introduction

New and advanced technologies enable us to collect greater quantities of functional data, in diverse areas including but not limited to financial exchange, medical data from wearable devices, MRI or CT scans, biological growth, climatology, traffic and online auction data. Consequently, the demands for analysis and prediction based on functional data have increased exponentially. A challenge analyzing functional data is that functional data may be irregularly and sparsely observed and typically contain too much noise. As a result, to build the relationship between a response and functional covariates, it is crucial to extract features from functional covariates that are associated with the response.

*Corresponding author. E-mail: linhz@swufe.edu.cn