

# Statistical Diagnostics for Semiparametric Mixed Models

Wing-Kam Fung<sup>1</sup>, Zhong-Yi Zhu<sup>2</sup>, Bo-Cheng Wei<sup>3</sup> and Xuming He<sup>4</sup>

<sup>1</sup> The University of Hong Kong, Pokfulam Road, Hong Kong, China. E-mail: wingfung@hku.hk

<sup>2</sup> East China Normal University, Shanghai, China

<sup>3</sup> Southeast University, Nanjing, China

<sup>4</sup> University of Illinois at Urbana-Champaign, USA

**Keywords:** Cook's distance, Longitudinal data, Penalized likelihood, Repeated measure, Semiparametric regression, Smoothing spline.

## Abstract

Semiparametric mixed models are useful in biometric and econometric applications, especially for longitudinal data. Maximum penalized likelihood estimators (MPLE) have been shown to work well by Zhang, Lin, Raz and Sowers (1998) for both linear coefficients and nonparametric functions. The present paper considers the role of influence diagnostics in the MPLE by extending the case-deletion and subject-deletion analysis of linear models to accommodate the inclusion of a nonparametric component. We focus on influence measures for the fixed effects and provide formulae that are analogous to those for simpler models and readily computable with the MPLE algorithm. We also establish an equivalence between the case- or subject-deletion model and a mean-shift outlier model from which we derive tests for outliers. The proposed influence diagnostics are illustrated through a longitudinal hormone study on progesterone and a simulated example.