

Generalized Additive Models with Correlated Data

Xiangming Fang
Department of Statistics and Actuarial Science
University of Iowa

Abstract

Generalized additive models (GAMs) assume that the mean of the response variable depends on an additive predictor through a link function. Like generalized linear models (GLMs), generalized additive models permit the response probability distribution to be any member of the exponential family of distributions. The only difference between GAMs and GLMs is that the GAMs allow for unknown smooth functions in the linear predictor. Due to the high flexibility in model specification, GAMs have been widely used. While the procedure for fitting a generalized additive model to independent data has been well established, not as much work has been done when the data are correlated. Generally speaking, there are currently two frameworks to fit a GAM with correlated data: the mixed model framework, which transforms the generalized additive model into a generalized mixed model, and the Bayesian framework, in which the inference is based on MCMC simulations. The mixed model framework turns out to be numerically unstable, while the Bayesian framework is often computationally expensive, especially when the sample size is large. Therefore, the currently available methods are not completely satisfactory in practice. We propose a new approach to fit generalized additive models with correlated data via the penalized likelihood approach which estimates the smooth functions and covariance parameters by iteratively maximizing the penalized log likelihood. Both ML and REML estimation schemes are developed. Simulation study shows the new approach works fine. Finally, the new approach is used for modeling the abundance distribution of jellyfish in the Bering Sea.