Statistica Sinica **35** (2025), 1-22
doi:https://doi.org/10.5705/ss.202021.0405

PARAMETRIC MODAL REGRESSION WITH AUTOCORRELATED ERROR PROCESS

Tao Wang*

University of Victoria

Abstract: We propose an efficient two-step estimation procedure for a parametric modal regression with autoregressive errors. The procedure relies on estimating a parametric transformation of the dependent variable from data using a (penalized) kernel-based objective function. We establish asymptotic normality for the resulting estimator and demonstrate that it possesses oracle properties, as if the true order of autoregressive error structure were known in advance. To numerically estimate modal parameter and determine the order of error structure, two modified (penalized) modal expectation-maximization (MEM) algorithms are developed. Furthermore, we present a modal residual-based autocorrelation test and show that the statistic is asymptotically distributed as a $\mathcal{X}^2$ distribution. Monte Carlo simulations and an empirical analysis are conducted to illustrate the finite sample performance of the resultant estimator. We also discuss the extension of the results to a nonparametric modal regression model.

Key words and phrases: Autoregressive error, MEM algorithm, modal regression, oracle property, order selection, residual-based test.

1. Introduction

Modal regression has recently attracted much attention due to its robustness for skewed and heavy-tailed data, which can be treated as a complement to mean or median (quantile) regression; see Ullah, Wang and Yao (2021, 2022, 2023). The main objective of modal regression is to capture how covariates X affect the "most likely" (mode) value of a response variable Y, as denoted by

$$Mode(Y \mid X) = \underset{Y}{\operatorname{argmax}}\, f_{Y|X}(Y \mid X), \tag{1.1}$$

where $f_{Y|X}(Y \mid X)$ represents the conditional density of Y given X. The modal regression line can then be obtained by nonparametrically estimating the aforementioned conditional density function (Chen et al., 2016). However, because of the "curse of dimensionty", such a density-based estimation is difficult to implement. Similar to mean or median (quantile) regression, we can avoid nonparametrically estimating conditional density and achieve different types of modal regression models by directly imposing structural assumptions on

*Corresponding author. E-mail: taow@uvic.ca