Journal of Multivariate Analysis 102 (2011) 349–362

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On linear models with long memory and heavy-tailed errors

Zhou Zhou *, Wei Biao Wu

University of Toronto, Canada
University of Chicago, United States

ARTICLE INFO

Article history:
Received 18 October 2007
Available online 22 September 2010

AMS subject classifications:
62J05
60G18
60G51

Keywords:
Bahadur representation
Heavy tails
Long memory
M-estimation

ABSTRACT

We consider the robust estimation of regression parameters in linear models with long memory and heavy-tailed errors. Asymptotic Bahadur-type representations of robust estimates are developed and their limiting distributions are obtained. It is shown that the limiting distributions are very different from those obtained under short memory. A simulation study is carried out to compare the performance of various asymptotic representations.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The estimation of unknown regression parameters in linear models has been extensively studied. The least squares estimate (LSE) is widely used in practice and its finite and asymptotic distribution theory has been well developed; see for example the texts by Davidson and MacKinnon [11] and Rao and Toutenburg [26]. For linear models with heavy-tailed errors, the LSE may perform poorly and robust estimates are attractive alternatives. The last three decades have witnessed a rapid growth in quantile estimation and other robust procedures. See [17,14,19] for excellent treatments.

Consider the p-variate linear model:

$$y_i = \mathbf{x}_i^T \beta + e_i, \quad i = 1, 2, \ldots, n, \tag{1}$$

where $\mathbf{A}^T$ denotes the matrix transpose, and $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T$, $1 \le i \le n$, are $p \times 1$ known deterministic design vectors. As a typical robust estimation procedure, let ρ be a convex function and we estimate the unknown parameter vector β by the minimizer

$$\hat{\beta}_n = \operatorname*{argmin}_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} \rho(y_i - \mathbf{x}_i^T \beta). \tag{2}$$

Note that, $\rho(u) = u^2$ leads to LSE. Other popular choices of ρ include quantile regression with $\rho(x) = \alpha x_+ + (1-\alpha)(-x)_+$, $0 < \alpha < 1$, where $x_+ = \max(0, x)$, Huber's procedure [17] with $\rho(x) = (x^2 \mathbf{1}\{|x| > c\})/2 + (cx - c^2/2)\mathbf{1}\{|x| \le c\}$, $c > 0$, and $\mathcal{L}^q$ regression with $\rho(x) = |x|^q$, $1 \le q \le 2$. In the literature, asymptotic properties of $\hat{\beta}_n - \beta$ have been studied mainly under the assumption that the errors are independent (Bassett and Koenker [5], Babu [3], Bai et al. [4] and He and Shao [15] among

* Corresponding address: Department of Statistics, 100 St. George St. Toronto, M5S 3G3, Canada.
E-mail addresses: zhou@utstat.toronto.edu, zhouzhou@uchicago.edu (Z. Zhou), wbwu@galton.uchicago.edu (W.B. Wu).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.09.009

others) or strong mixing [13,23,10] or short-range dependent [33]. See the latter paper for additional references for robust estimation under independence and weak dependence. Hampel et al. [14] argued that many science and engineering data exhibit significant temporal dependence and the assumption of independence is violated; see Chapter 8 therein. However, there seem to be few results on robust estimators of linear models with long memory (or long-range dependent) errors.

Recently processes with both heavy tails and long memory have received considerable attention. Willinger et al. [32] showed that self-similarity and heavy tails exist in network traffic data, while Rachev and Mittnik [25] did an extensive empirical study and showed that high frequency asset return data exhibited both long memory and heavy tails. To the best of our knowledge, most of the existing results focused on estimation and inference of long memory and heavy-tail parameters, while little attention was paid to regression analysis. The latter problem is clearly of great interest if one wants to include covariates or predictors into the model for explanatory purpose.

The paper aims to study properties of $\hat{\beta}_n$ under the assumption that the errors e_i in model (1) are long memory as well as heavy tailed; see Section 2.1 for assumptions on the error structure. It is shown that asymptotic behavior of $\hat{\beta}_n$ is very different from that obtained under independence and weak dependence. We will also provide Bahadur representations of the robust estimates of model (1). Those representations are useful for further analysis of the asymptotics of robust estimates.

The rest of the paper is structured as follows. Regularity conditions are given in Section 2. Section 3 presents main results including consistency and asymptotic distributions of robust estimates. Section 5 provides proofs and Section 4 presents a simulation study.

2. Preliminaries

We now introduce some notation. For a vector $\mathbf{v} = (v_1, \ldots, v_p) \in \mathbb{R}^p$, let $|\mathbf{v}| = (\sum_{i=1}^p v_i^2)^{1/2}$. For a $p \times p$ matrix A, define $|A| = \sup\{|A\mathbf{v}| : |\mathbf{v}| = 1\}$. For a random vector $\mathbf{V}$, write $\mathbf{V} \in \mathcal{L}^q$ $(q > 0)$ if $\|\mathbf{V}\|_q := [\mathbb{E}(|\mathbf{V}|^q)]^{1/q} < \infty$. Let (η_n) be a sequence of random variables and (d_n) a positive sequence. We write $\eta_n = o_p(d_n)$ if $\eta_n/d_n \to 0$ in probability and $\eta_n = O_p(d_n)$ if η_n/d_n is bounded in probability. Denote by $\Rightarrow$ the weak convergence. Let $\mathcal{C}^i$, $i \in \mathbb{N}$, be the collection of functions that have ith order continuous derivatives. Let C denote a generic constant independent of n and its value may vary from place to place.

2.1. The error structure

We assume that (e_i) is a moving average process

$$e_i = \sum_{j=0}^{\infty} a_j \varepsilon_{i-j}, \tag{3}$$

where $\varepsilon_j, j \in \mathbb{Z}$, are independent and identically distributed (iid) random variables with mean 0 and $\varepsilon_j \in \mathcal{D}(\alpha)$, $\alpha \in (1, 2)$. Here $\mathcal{D}(\alpha)$ denotes the α-stable domain of attraction (see [9]); namely, there exists real sequences (A_n) and (B_n) such that $A_n^{-1}(\varepsilon_1 + \cdots + \varepsilon_n) - B_n$ converges to an α-stable law whose characteristic function is

$$\varphi(t) = \exp(-\sigma^\alpha |t|^\alpha (1 - \sqrt{-1}\varrho w_\alpha(t)) + \sqrt{-1}\mu t), \quad \text{where } w_\alpha(t) = \tan \frac{\pi \alpha \operatorname{sgn}(t)}{2}. \tag{4}$$

Here σ, μ and ϱ $(-1 \le \varrho \le 1)$ are the scale, shift and skewness parameters, respectively, and $\sqrt{-1}$ is the imaginary unit. Let F_ε be the distribution function of ε_j and $f_\varepsilon = F_\varepsilon'$ be its density. Then $\varepsilon_j \in \mathcal{D}(\alpha)$ can be characterized by

$$1 - F_\varepsilon(u) = \frac{c_1 + o(1)}{u^\alpha} L(u) \quad \text{and} \quad F_\varepsilon(-u) = \frac{c_2 + o(1)}{u^\alpha} L(u) \quad \text{as } u \to \infty, \tag{5}$$

where $c_1, c_2 \ge 0$, $c_1 + c_2 > 0$ and L is a slowly varying function, i.e., $\lim_{x\to\infty} L(tx)/L(x) = 1$, for all $t > 0$ (cf. [7]). It is easy to see that $\inf\{x : \mathbb{P}(|\varepsilon_i| > x) \le 1/n\} = n^{1/\alpha} L_1(n)$, where L_1 is also a slowly varying function. Observe that $\varepsilon_i \in \mathcal{L}^{\alpha'}$, for all $\alpha' \in (0, \alpha)$, and α is called the heavy-tail index, and $\mathbb{E}(\varepsilon_i^2) = \infty$. If $\varrho = \mu = 0$, then (4) becomes the symmetric-α-stable ($S\alpha S$) law. In this case (5) holds with $L(t) = 1$, $c_1 = c_2 = \sigma^\alpha/(2C_\alpha)$, where $C_\alpha = \cos(\alpha\pi/2)\Gamma(2-\alpha)/(1-\alpha)$.

Let $\varepsilon_\alpha(u), u \in \mathbb{R}$, be a two-sided Levy α-stable process [28] with independent increments, $\varepsilon_\alpha(0) = 0$, and $\varepsilon_\alpha(u+t) - \varepsilon_\alpha(u)$ having characteristic function $\varphi(t)$ (cf. (4)) with $\mu = 0$. By Theorem 2.7 in [29], in the space $\mathcal{D}[0, 1]$ of functions that are right continuous and have left limit, we have the weak convergence

$$\lim_{n\to\infty} \left\{ \frac{1}{n^{1/\alpha} L_1(n)} \sum_{i=1}^{\lfloor nu \rfloor} \varepsilon_i, 0 \le u \le 1 \right\} = \{\varepsilon_\alpha(u), 0 \le u \le 1\}, \tag{6}$$

where $\lfloor v \rfloor = \max\{j \in \mathbb{Z} : j \le v\}$. See also [2].

For the coefficients $(a_j)_0^\infty$, we assume $a_0 = 1$, and, for $j \ge 1$,

$$a_j = j^{-\gamma} l(j), \quad 1/\alpha < \gamma < 1, \text{ where } l(\cdot) \text{ is a slowly varying function.} \tag{7}$$

Download English Version:

https://daneshyari.com/en/article/1146898

Download Persian Version:

https://daneshyari.com/article/1146898

Daneshyari.com