Journal of the Korean
Statistical Society
Vol. 27, No. 1, 1998

Asymptotic Properties of LAD Estimators in Censored Nonlinear Regression Model †

Seung Hoe Choi [1] and Hae Kyung Kim [2]

ABSTRACT

This paper is concerned with the asymptotic properties of the least absolute deviation estimators for the nonlinear regression model when dependent variables are subject to censoring time, and proposed the simple and practical sufficient conditions for the strong consistency and asymptotic normality of the least absolute deviation estimators in censored regression model. Some desirable asymptotic properties including the asymptotic relative efficiency of proposed model with respect to standard model are given.

Key Words : Censored Nonlinear Model; Consistency; Normality; Efficiency.

---

†This work was supported by Yonsei University Research Fund of 1997.
[1]Department of General Studies, Hankuk Aviation University, Koyang, 411-791, Korea.
[2]Department of Mathematics, Yonsei University, Seoul, 120-749, Korea.

1. INTRODUCTION

The standard nonlinear regression model is

$$y_t = f(x_t, \theta_o) + \epsilon_t, \quad t = 1, 2, \cdots, n \tag{1.1}$$

where f is a real valued function defined on $R^{(p+q)}$, x_t is a $(1 \times q)$ observed vector, the error term ϵ_t are independent and identically distributed (i.i.d.) with finite variance. The parameter vectors θ_o which is interior point in Θ is unknown and to be estimated. Various statistical properties derived from standard regression form are not available if the dependent variables are subject to censoring time. Censored data, which contains only partial information about random variables of interest, are important in applications to medical statistics , biology, engineering, etc.

We consider in this paper the following censored nonlinear regression model

$$y_t = \min\{c_t, f(x_t, \theta_o) + \epsilon_t\}, \quad t = 1, 2, \cdots, n. \tag{1.2}$$

In censored regression model we observe only the censored data (Z_t, δ_t, x_t) with $Z_t = \min\{y_t, c_t\}, \delta_t = I_{[y_t \le c_t]}$ where c_t is a censoring time and I is indicator function.

The censored least absolute deviation(LAD) estimator of the true parameter θ_o based on (Z_t, x_t, δ_t), denoted by $\hat{\theta}_n$, is a parameter which minimizes the objective function

$$D_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} |y_t - \min\{c_t, f(x_t, \theta)\}|. \tag{1.3}$$

Modifying (1.3), we have another objective function of the censored LAD estimator

$$Q_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \{|y_t - \min\{c_t, f(x_t, \theta)\}| - |y_t - \min\{c_t, f(x_t, \theta_o)\}|\}. \tag{1.4}$$

Asymptotic results for censored linear regression models are given by various authors; Amemiya(1973), Powell(1984, 1986) and Chen and Wu(1994).

Amemiya(1973) proved the strong consistency and asymptotic normality of the maximum likelihood estimator(MLE) for regression model where the dependent variable is normal but censored to the left of zero. Powell(1984, 1986) proposed a censored regression form like as model (1.2) and censored LAD and quantile estimator, and investigated asymptotic properties of the proposed censored estimator under some regularity conditions. Chen and Wu(1994) gave somewhat different sufficient conditions from those which employed in Powell(1984) of the censored LAD estimator.

The main object of this paper is to provide simple and practically sufficient conditions for the asymptotic properties of the nonlinear censored LAD estimator $\hat{\theta}_n$. The order of presentation in the paper is as follows; Section 2 states basic notations and assumptions and investigates the strong consistency of proposed estimator in model (1.2). The section which follows presents sufficient conditions for the asymptotic normality of the censored nonlinear LAD estimator $\hat{\theta}_n$. The final section gives asymptotic relative efficiency of proposed censored model with respect to standard model.

2. STRONG CONSISTENCY

We start this section by introducing simple notations and conditions of regression function $f(x,\theta)$ and error term ϵ_t.

Let $(\Omega, \mathcal{A}, \mathcal{P})$ be probability space on R^p, and G_t and H_t denote the distribution function of error term ϵ_t and input variable X_t, respectively. Suppose that X_t is independent random variable and H_t is not degenerate. To simplify the notations, we denote

$$\nabla f_t(\theta) = \left(\frac{\partial}{\partial\theta_1} f(x_t,\theta), \cdots, \frac{\partial}{\partial\theta_p} f(x_t,\theta)\right)_{(p\times 1)},$$

$$\nabla^2 f_t(\theta) = \left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} f(x_t,\theta)\right]_{(p\times p)},$$

$$V_t(\epsilon_t,\theta,\theta_o) = |y_t - \min\{c_t, f(x_t,\theta)\}| - |y_t - \min\{\epsilon_t, f(x_t,\theta_o)\}|.$$

Throughout this paper, we will use the following assumptions;

Assumption A

The parameter space Θ is a compact subspace of R^p.

Assumption B

B_1 : $f(x_t,\theta), \nabla f_t(\theta)$, and $\nabla^2 f_t(\theta)$ are continuous on $\Omega \times \Theta$ for each t.

B_2 : x_t and ϵ_t are independent and ϵ_t has a unique median at zero.

B_3 : x_t is bounded in probability, i.e.,for every $\eta > 0$ there exists M_η such that $P\{|x_t| > M_\epsilon\} < \epsilon$.

For the strong consistency of the LAD estimator, we need the additional assumptions.

Assumption C

$C_1 : P\{x \in \Omega : f(x, \theta_1) \neq f(x, \theta_2)\} > 0$ for each $\theta_1, \theta_2 \in \Theta, \theta_1 \neq \theta_2$.

$C_2 : \frac{1}{n}\sum_{t=1}^{n} I_{\{c_t > f_t(\theta_o)\}} \nabla^T f_t(\theta_o) \nabla f_t(\theta_o)$ converges to a positive definite matrix $V(\theta_o)$ as $n \to \infty$.

Remark 2.1 For the asymptotic properties of any estimators of true parameter θ_o, in censored regression model the number of $\{t : c_t > f(x_t, \theta_o)\}$ must be sufficiently large. So the factor $I_{\{c_t > f_t(\theta_o)\}}$ appears in Assumption C_2. See Chen and Wu(1994). If the investigators prolong the observation time infinitely, the number of $\{t : c_t > f(x_t, \theta_o)\}$ is same with the number of the observation. Assumption C_2 thus converts that $\frac{1}{n}\sum_{t=1}^{n} \nabla^T f_t(\theta_o) \nabla f_t(\theta_o)$ converges to a positive definite matrix as $n \to \infty$, which is regular condition that is often referred in nonlinear regression model. See Kim and Choi(1995).

Theorem 2.1 Suppose that Assumption A, B, and C are satisfied on the model (1.2). Then the LAD estimator $\hat{\theta}_n$ is strongly consistent for θ_o.

Proof. Note that

(i) $\sup_{\theta \in \Theta} |Q_n(\theta) - EQ_n(\theta)| = o_p(1)$,

(ii) $Q(\theta) = \lim_{n \to \infty} EQ_n(\theta)$ has a unique minimizer θ_o in Θ,

where $o_p(1)$ stands for convergence in probability. By these is results we have

$$\lim_{n \to \infty} \inf_{\|\theta - \theta_o\| \geq \delta} \{D_n(\theta) - D_n(\theta_o)\} > 0 \quad a.e,$$

for any $\delta > 0$. For a detailed proof, see Choi and Kim(1997).

3. ASYMPTOTIC NORMALITY

In present section we consider the asymptotic normality of the proposed estimator $\hat{\theta}_n$ which is one of the most important statistical properties in asymptotic theory.

Theorem 3.1

Suppose that Assumption A, B and C hold on the model(1.2). Then $\sqrt{n}(\hat{\theta}_n - \theta_o)$ converges in distribution to a p-variate normal random vector with mean zero and covariance matrix $(2g(0))^{-2}V^{-1}(\theta_o)$. That is,

$$\sqrt{n}(\hat{\theta}_n - \theta_o) \xrightarrow{\mathcal{L}} N_p(0, (2g(0))^{-2}V^{-1}(\theta_o)),$$

where $g(0)$ is the height of the density of the error ϵ_t at zero.

Proof. The main idea is to approximate to the function $|x|$ by a smooth function $\rho_n(x)$ such that $\lim_{n\to\infty} \rho_n(x) = |x|$. As such function we use

$$\rho_n(x) = \left[-\frac{\beta_n^2}{3}x^3 + \beta_n x^2 + \frac{1}{\beta_n}\right] I_{\{0 \le x < \frac{1}{\beta_n}\}} + xI_{\{x > \frac{1}{\beta_n}\}}$$

and

$$\rho_n(x) = \rho_n(-x),$$

where $n = o(\beta_n^2), \beta_n = o(n)$ and $n^{\frac{1+\delta}{2}} = o(\beta_n)$ for some $\delta > 0$. The sequence $\beta_n = \frac{1}{n^p}$ satisfies above conditions for $p \in (\frac{1}{2}, 1)$. Let $D_n^*(\theta) = \frac{1}{n}\sum_{t=1}^n \rho_n(y_t - \min\{c_t, f(x_t, \theta)\})$ and $\tilde{\theta}_n$ denote a minimizer of $D_n^*(\theta)$. The rest of the proof can be briefly described as follows.

(i) $\sup_{\theta\in\Theta} n\{D_n(\theta) - D_n^*(\theta)\} = o_p(1)$,

(ii) $\sqrt{n}(\hat{\theta}_n - \tilde{\theta}_n) = o_p(1)$,

(iii) $\sqrt{n}(\tilde{\theta}_n - \theta_o) \xrightarrow{\mathcal{L}} N_p(0, (2g(0))^{-2}V^{-1}(\theta_o))$.

The detailed proof of this theorem is given in appendix.

4. ASYMPTOTIC RELATIVE EFFICIENCY

This section considers the asymptotic relative efficiency of censored regression model to standard regression form based on confidence region of the true parameter θ_o which derived from the asymptotic normality of the LAD estimator $\hat{\theta}_n$.

Let C^i denote the vector $\{c_{1i}, \cdots, c_{ni}\}$ and u_n^i indicate number of elements of the set $U_n^i = \{t : f(x_t, \theta_o) \le c_{ti}\}$.

For the purpose of this section we also need the following Assumption:

Assumption D

The ratio of the number of the uncensored data to the number of the sample in model (1.1), denoted by $\frac{u_n^i}{n}$, converges to $p_i, 0 < p_i < 1$.

Note that Assumption C suggests

$$\frac{1}{u_n^i} \sum_{t \in U_n^i} \nabla^T f_t(\theta_o) \nabla f_t(\theta_o) \to [p_i]^{-1} V(\theta_o) \tag{4.1}$$

Furthermore, Theorem 3.1 in Kim and Choi(1995) and (4.1) imply that

$$\sqrt{u_n^i}(\theta_{u_n^i} - \theta_o) \xrightarrow{\mathcal{L}} N_p(0, (2g(0))^{-2} p_i V(\theta_o)).$$

From the above results, we have the following theorem.

Theorem 4.1 Suppose that Assumption D holds in model (1.2). Under the same conditions of the theorem 3.1, we have get

$$\sqrt{n}(\hat{\theta}_{u_n^i} - \theta_o) \xrightarrow{\mathcal{L}} N_p(0, (2g(0))^{-2} V^{-1}(\theta_o)) \tag{4.2}$$

The asymptotic normality of $\sqrt{n}(\hat{\theta}_n - \theta_o)$, derived in theorem 3.1, implies that $Q(\sqrt{n}(\hat{\theta}_n - \theta_o), V_n(\hat{\theta}_n))$ has asymptotically a chi-square distribution with p degrees of freedom, denoted by χ_p^2, where

$$Q(\sqrt{n}(\hat{\theta}_n - \theta_o), V_n(\hat{\theta}_n)) = (2g(0))^2 n(\hat{\theta}_n - \theta_o) V_n(\hat{\theta}_n)(\hat{\theta}_n - \theta_o)$$

and

$$V_n(\hat{\theta}_n) = \frac{1}{n} \sum_{t=1}^{n} \nabla^T f_t(\hat{\theta}_n) \nabla f_t(\hat{\theta}_n).$$

Therefore, the ellipsoidal confidence region for θ

$$E(\hat{\theta}) = \{\theta : Q(\sqrt{n}(\hat{\theta}_n - \theta), V_n(\hat{\theta}_n)) \le C_\alpha\}$$

has asymptotic confidence coefficient $1 - \alpha$, where $P(\chi_p^2 > C_\alpha) = \alpha$. As mentioned in Serpling(1980), if we define the asymptotic relative efficiency of two proposed model is the ratio of the volumes of the corresponding confidence ellipsoids, for a specified value of the limiting confidence coefficient, by proceeding approach we have the next consequence.

Theorem 4.2

Under the same conditions of the theorem 4.1, the asymptotic relative efficiency of censored model relative to standard model is p_i which coincides with limit of the ratio of the number of the censored data and the total sample.

On the other hand, in according to the fixed censoring time's altering there is another censored regression model. To seek the ARE of the new censored model with respect to old censored model, we need to the following notation. $C^i \leq C^j$ means that $c_{ti} \leq c_{tj}$ for all t. By similar method, we conclude that the ARE of the new model with the censoring time C^j relative to old model with C^i is $\frac{p_j}{p_i}$, which is identical with the limit of ratio of the uncensored data in each model.

The above result and theorem imply that the model including censored data is relatively more efficient than the standard model and the more the censored model(1.2) has enlarged censoring time, the more efficiency the proposed censored model has than the other censored model.

APPENDIX

Proof of theorem 3.1

For the first aim, note that

$$\{D_n(\theta) - D_n^*(\theta)\} \leq \sum_{t=1}^{n} \frac{1}{3\beta_n} I_{\{|y_t - \min\{c_t, f(x_t,\theta)\}| \leq \frac{1}{\beta_n}\}}$$

By continuity of $\{D_n(\theta) - D_n^*(\theta)\}$ and compactness of parameter space we choose θ^* such that

$$n\{D_n(\theta^*) - D_n^*(\theta^*)\} = \sup_{\theta\in\Theta} n\{D_n(\theta) - D_n^*(\theta)\}$$

On the other hand, Chebyshev's inequality gives

$$P[|n\{D_n(\theta) - D_n^*(\theta)\}| > \epsilon] \leq \frac{n \max_{1\leq t\leq n} Var I_{\{|r_t(\theta)| \leq \frac{1}{\beta_n}\}}}{9\beta_n^2 \epsilon^2}, \qquad (A.1)$$

where $r_t(\theta) = y_t - \min\{f_t(\theta), c_t\}$.

Meanwhile, let

$$\begin{aligned} \Omega_1 &= \{x \in \Omega : f_t(\theta_o) + \epsilon_t < c_t < f_t(\theta)\}, \\ \Omega_2 &= \{x \in \Omega : f_t(\theta_o) + \epsilon_t < c_t, f_t(\theta) < c_t\}, \\ \Omega_3 &= \{x \in \Omega : f_t(\theta_o) + \epsilon_t > c_t, c_t < f_t(\theta)\}, \\ \Omega_4 &= \{x \in \Omega : f_t(\theta_o) + \epsilon_t < c_t, c_t > f_t(\theta_o)\}, \end{aligned}$$

where $f_t(\theta) = f(x_t, \theta)$. Then in the case Ω_2 we obtain

$$EI_{\{|r_t(\theta)| \le \frac{1}{\beta_n}\}} = \int_{d_t(\theta)-\frac{1}{\beta_n}}^{d_t(\theta)-\frac{1}{\beta_n}} g_t(\lambda) d\lambda.$$

By similar method we get

$$EI_{\{|r_t(\theta)| \le \frac{1}{\beta_n}\}} = \begin{cases} (\frac{2}{\beta_n})[g_t(a_t(\theta_o)) + o(1)], & \text{on} \quad \Omega_1, \\ (\frac{2}{\beta_n})[g_t(d_t(\theta)) + o(1)], & \text{on} \quad \Omega_2, \\ 1, & \text{on} \quad \Omega_3, \\ 1 \text{ or } 0, & \text{on} \quad \Omega_4. \end{cases} \tag{A.2}$$

where $d_t(\theta) = f_t(\theta) - f_t(\theta_o)$ and $a_t(\theta) = c_t - f_t(\theta)$.

Combining with (A.1) and (A.2), we conclude that

$$\sup_{\theta \in \Theta} n\{D_n(\theta) - D_n^*(\theta)\} = 0_p(1). \tag{A.3}$$

For the second purpose, note that $D_n^*(\hat{\theta}_n) - D_n^*(\tilde{\theta}_n)$ is less than

$$[D_n^*(\hat{\theta}_n) - D_n(\hat{\theta}_n)] + [D_n(\tilde{\theta}_n) - D_n^*(\tilde{\theta}_n)]$$

because of $[D_n(\hat{\theta}_n) - D_n(\tilde{\theta}_n)] \le 0$. This result and first conclusion mean that

$$n|D_n^*(\hat{\theta}_n) - D_n^*(\tilde{\theta}_n)| = 0_p(1).$$

Further, we can rewrite $D_n^*(\theta)$ as following;

$$D_n^*(\theta) = \frac{1}{n} \sum_{t=1}^{n} \{\rho_n(y_t - f_t(\theta)) I_{\{f_t(\theta) \le c_t\}} + \rho_n(y_t - c_t) I_{\{f_t(\theta) \ge c_t\}}\}$$

Thus, we obtain

$$\nabla D_n^*(\theta) = -\frac{1}{n}\sum_{t=1}^{n}\{\rho_n'(y_t - f_t(\theta))I_{\{f_t(\theta)\leq c_t\}}\}\nabla f_t(\theta)$$

and

$$\begin{aligned}\nabla^2 D_n^*(\theta) &= \frac{1}{n}\sum_{t=1}^{n}\{\rho_n''(y_t - f_t(\theta))I_{\{f_t(\theta)\leq c_t\}}\}\nabla^T f_t(\theta)\nabla f_t(\theta) \\ &\quad - \frac{1}{n}\sum_{t=1}^{n}\{\rho_n'(y_t - f_t(\theta))I_{\{f_t(\theta)\leq c_t\}}\}\nabla^2 f_t(\theta).\end{aligned}$$

Let λ_k be denote k-th largest eigenvalue of $\nabla^2 D_n(\bar{\theta}_n)$. Then Assumption B follows that $\nabla^2 D_n(\theta)$ is symmetric matrix and Courant-Fisher minimax characterization gives

$$(\hat{\theta}_n - \tilde{\theta}_n)^T(\hat{\theta}_n - \tilde{\theta}_n) \leq \frac{1}{\lambda_1}(\hat{\theta}_n - \tilde{\theta}_n)^T\nabla^2 D_n^*(\bar{\theta}_n)(\hat{\theta}_n - \tilde{\theta}_n),$$

where $\bar{\theta}_n$ is the point of segment of the line joining $\hat{\theta}_n$ and $\tilde{\theta}_n$. Moreover, Second-order Taylor series presents

$$D_n^*(\hat{\theta}_n) - D_n^*(\tilde{\theta}_n) = \frac{1}{2}(\hat{\theta}_n - \tilde{\theta}_n)^T\nabla^2 D_n^*(\bar{\theta}_n)(\hat{\theta}_n - \tilde{\theta}_n).$$

Therefore, we obtain

$$n(\hat{\theta}_n - \tilde{\theta}_n)^T(\hat{\theta}_n - \tilde{\theta}_n) \leq \frac{2n}{\lambda_1}|D_n^*(\hat{\theta}_n) - D_n^*(\tilde{\theta}_n)|.$$

To complete the proof of purpose (ii), we have to show that $\lambda_1 \neq 0$. It is sufficient to show that $\nabla^2 D_n^*(\bar{\theta}_n)$ is positive definite. For this, we first will evaluate

$$[\rho_n''(r_t(\bar{\theta}_n))I_{\{f_t(\bar{\theta}_n)\leq c_t\}}] - 2g_t(0) = o_p(1).$$

Since $\rho_n''(x) = [-2\beta_n^2 x + 2\beta_n]I_{\{0\leq x\leq \frac{1}{\beta_n}\}}$, $\rho_n''(-x) = \rho_n''(x)$ and $a_t(\theta)$ is finite, we get

$$E_{\epsilon_t}[\rho_n''(r_t(\bar{\theta}_n))I_{\{f_t(\bar{\theta}_n)\leq c_t\}}I_{\{\epsilon_t > a_t(\theta_o)\}}] = 0.$$

While, since

$$G_t(d_t(\bar{\theta}_n) + \frac{1}{\beta_n}) - G_t(d_t(\bar{\theta}_n) - \frac{1}{\beta_n}) - 2g_t(0) = o(1),$$

$$|\beta_n d_t(\bar{\theta}_n)[g_t(\theta_n^1) - g_t(\theta_n^2)]| \le \|\beta_n(\bar{\theta}_n - \theta_o)\|_2^2 \|\nabla f_t(\bar{\theta}_n)\|_2^2 |g_t(\theta_n^1) - g_t(\theta_n^2)|,$$

and

$$\beta_n^2\Big[\int_{d_t(\bar{\theta}_n)-\frac{1}{\beta_n}}^{d_t(\bar{\theta}_n)} \lambda g_t(\lambda)d\lambda - \int_{d_t(\bar{\theta}_n)}^{d_t(\bar{\theta}_n)+\frac{1}{\beta_n}} \lambda g_t(\lambda)d\lambda\Big] - g_t(0) = o(1).$$

Thus, we have

$$E_{\epsilon_t}[\rho_n''(r_t(\bar{\theta}_n))I_{\{f_t(\bar{\theta}_n)\le c_t\}}] - 2g_t(0) = o_p(1).$$

Hence

$$\frac{1}{n}\sum_{t=1}^{n}\{\rho_n''(r_t(\bar{\theta}_n))I_{\{f_t(\bar{\theta}_n)\le c_t\}}\} = 2g(0) + o_p(1)$$

follows Chebyshev's inequality and the fact that

$$Var[\rho_n''(r_t(\bar{\theta}_n)] \le E[\rho_n''(r_t(\bar{\theta}_n))]^2 \le E[2\beta_n I_{\{|r_t(\bar{\theta}_n)|\le\frac{1}{\beta_n}\}}]^2.$$

Using the mean value theorem for final object, we acquire

$$\nabla D_n^*(\theta_o) = \nabla D_n^*(\tilde{\theta}_n) + \nabla^2 D_n^*(\bar{\theta}_n)(\tilde{\theta}_n - \theta_o),$$

where $\bar{\theta}_n = \lambda\theta_o + (1-\lambda)\tilde{\theta}_n$. Thus,

$$\sqrt{n}(\tilde{\theta}_n - \theta_o) = [\nabla^2 D_n^*(\bar{\theta}_n)]^{-1}\sqrt{n}\nabla D_n^*(\theta_o). \qquad (A.4)$$

By theorem 2.1 and (A.3), we easily prove that $\tilde{\theta}_n$ converges almost surely to θ_o. To check normality of $\sqrt{n}\nabla D_n^*(\theta_o)$ we observe that

$$\begin{aligned}\sqrt{n}D_n^*(\theta_o) &= -\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{n}\{\rho_n'(r_t(\theta_o))I_{\{f_t(\theta_o)<c_t\}}\}\nabla f_t(\theta_o)\\ &= \tfrac{1}{n}\textstyle\sum_{t=1}^{n}\{\sqrt{n}h_n(r_t(\theta_o))I_{\{|r_t(\theta_o)|\le\frac{1}{\beta_n}\}}I_{\{f_t(\theta_o)<c_t\}}\}\nabla f_t(\theta_o)\\ &+ \tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{n}\{Sign(r_t(\theta_o))I_{\{|r_t(\theta_o)|\ge\frac{1}{\beta_n}\}}I_{\{f_t(\theta_o)<c_t\}}\}\nabla f_t(\theta_o),\end{aligned} \qquad (A.5)$$

where $h_n(x) = -\beta_n^2x^2 + 2\beta_n x$. Markov's theorem gives

$$\frac{1}{n}\sum_{t=1}^{n}\{\sqrt{n}h_n(r_t(\theta_o))I_{\{|r_t(\theta_o)|\le\frac{1}{\beta_n}\}}I_{\{f_t(\theta_o)<c_t\}}\} = o_p(1)$$

due to

$$E|\sqrt{n}\{h_n(r_t(\theta_o))I_{\{|r_t(\theta_o)|\leq\frac{1}{\beta_n}\}}|^{1+\delta} \leq \frac{n^{\frac{1+\delta}{2}}}{\beta_n}g_t(\eta_n),$$

where $\eta_n = o(1)$.

The second term of (A.5) equals to as following;

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t I_{\{f_t(\theta_o)<c_t\}}\nabla f_t(\theta_o),$$

where

$$U_t = \{Sign(\epsilon_t)I_{\{|\epsilon_t|\geq\frac{1}{\beta_n}\}}I_{\{\epsilon_t<a_t(\theta_o)\}} + I_{\{\epsilon_t>a_t(\theta_o)\}}\}.$$

Let

$$Z_t = \frac{1}{\sqrt{n}}\sum_{t=1}^{k} U_t I_{\{f_t(\theta_o)<c_t\}}\lambda_k \frac{\partial}{\partial\theta_k}f_t(\theta_o)$$

and

$$B_n^2 = \frac{1}{n}\sum_{t=1}^{n}[\sum_{k=1}^{p}\lambda_k\frac{\partial}{\partial\theta_k}f_t(\theta_o)]^2.$$

Since $\frac{1}{B_n^2}\sum_{t=1}^{n} EZ_t^2 I_{\{|Z_t|>\epsilon B_n\}}$ converges to zero, from Cramér-Wold device and the application of Linderberg Central Limit Theorem we conclude that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n} U_t I_{\{f_t(\theta_o)<c_t\}}\nabla f_t(\theta_o) \xrightarrow{\mathcal{L}} N_p(0, V(\theta_o)). \qquad (A.6)$$

The asymptotic normality of $\tilde{\theta}_n$, which is the final aim holds by (A.4) and (A.6).

REFERENCES

(1) Amemiya, T., (1973), Regression Analysis when the Dependent Variable is Truncated Normal, *Econometrica* , Vol.41, 997-1016.

(2) Chen, X. R. and Wu, Y., (1994), Consistency of L_1 Estimates in Censored Linear Regression Models, *Communication in Statistics* , Vol.23, 1849-1858.

(3) Choi, S. H. and Kim, H. K., (1997), The Strong Consistency of the L_1-norm Estimators in Censored Nonlinear Regression Models, *Bulletin of the Korean Mathematical Society* , Vol.34, 573-581.

(4) Kim, H. K. and Choi, S. H. (1995), Asymptotic Properties of Nonlinear Least Absolute Deviation Estimators, *Journal of Korean Statistical Society* , Vol.24, 127-139.

(5) Powell, J. L., (1984), Least Absolute Deviation Estimation for the Censored Regression Model, *Journal of Econometrics* , Vol.25, 303-325.

(6) Powell, J. L., (1986), Censored Regression Quantiles, *Journal of Econometrics* , Vol.27, 143-155.

(7) Serpling, R. J., (1980), *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons.