응용통계연구 제 7권 2호
1994년 9월 pp. 173-182

Asymptotically Efficient L-Estimation for Regression Slope When Trimming is Given

Sang Moon Han[1)]

Abstract

By applying slope estimator under the arbitrary error distributions proposed by Han(1993), if we define regression quantiles to give upper and lower trimming part and blocks of data, we show the proposed slope estimator has asymptotically efficient slope estimator when the number of regression quantiles to from blocks of data goes to sufficiently large.

1. Introduction

We are concerned with estimating asymptotically efficient regression slope estimator when trimming is given under the regression model

$$\underline{y} = X\underline{\beta} + \underline{z}, \tag{1.1}$$

where $\underline{y} = (y_1, \cdots, y_n)'$, X is n times p matrix of known constants whose i-th row is $\underline{x_i}'$, $\underline{\beta} = (\beta_0, \beta_1, \cdots, \beta_{p-1})'$ is a vector of unknown parameters, and $\underline{z} = (z_1, \cdots, z_n)'$ is a vector of independent, identically distributed random variables with unknown distribution function F.

The traditional way of estimating regression slope is the method of least squares. Despite the obvious advantage of least squares(L.S) estimator of regression slope this estimator performs poorly when underlying error distribution F has tails heavier than those of normal or asymmetric error distributions. Many authors like Johns(1974) and Sacks(1975) considered asymptotically efficient L-estimate in the location model. In this paper, we extend the result of Ruppert and Carroll(1980) for estimating regression slope to asymptotically efficient slope estimator using regression quantile when trimming is given.

2. The proposed estimators and large sample properties

Before starting this section, we introduce some notation and assumption which are imposed for all theorems in this section. Although $\underline{y}$, X, and $\underline{z}$ in (1.1) depends on n,

1) Department of Computer Science and Statistics, Seoul City University, 90 Jeonnong Dong, Dongdaemoon Ku, Seoul, KOREA.

this is not made explicit in the notation. Let $\underline{e}=(1,0,\ldots,0)'$ be $(p\times 1)$, and let I_p be the $(p\times p)$ identity matrix. Whenever r is scalar, $\underline{r}=r\underline{e}$. For $0<p_i<1$, define $\xi_i=F^{-1}(p_i)$. Let $N_p(\underline{\mu},\Sigma)$ denote the p-variate normal distribution with mean vector $\underline{\mu}$ and variance-covariance matrix Σ. We also make the following assumptions about the family $\wedge$ of distributions in what follows.

A1. F has a continuous density f and $f(x)>0$ for all $F\in\wedge$.

A2. Let $\underline{x_i}'=(x_{i1},x_{i2},\ldots,x_{ip})$ be the i-th row of X and $x_{i1}=1$, $i=1,2,\ldots,n$ and

$$\sum_{i=1}^{n} x_{ij}=0,\ j=2,3,..,p.$$

A3. $\lim_{n\to\infty}\left(\max_{j\le p, i\le n} n^{-\frac{1}{2}}\,|x_{ij}|\right)=0.$

A4. There exists a positive definite matrix Q such that $\lim_{n\to\infty} n^{-1}(X'X)=Q.$

A5. Assuming $\underline{\widehat{\beta_0}}$ be preliminary fit, then $\sqrt{n}(\underline{\widehat{\beta_0}}-\underline{\beta}-\theta\underline{e})=0_p(1)$ for some constant θ.

A6. The density f and its first three derivatives exist and continuous; moreover they are uniformly bounded and bounded away from zero for all $F\in\wedge$.

A7. The Fisher information $I(F,0,0)$ is positive and finite; that is,

$$0<\int_{-\infty}^{\infty}[f'(x)/f(x)]^2 f(x)dx<\infty.$$

Motivation of our estimator is as follows; Assume $0<p_0<p_1<\ldots<p_k=q_0<1$. Moreover, Let $K(p_i)$ be the corresponding p_i-th regression quantiles. Then for $i=1,2,\ldots,n$, define

$$a_1=\begin{cases}1 & \text{if } \underline{x_i}'K(p_0)<y_i\le\underline{x_i}'K(p_1)\\ 0 & \text{if } otherwise\end{cases}$$
$$a_2=\begin{cases}1 & \text{if } \underline{x_i}'K(p_1)<y_i\le\underline{x_i}'K(p_2)\\ 0 & \text{if } otherwise\end{cases}$$
$$\vdots \qquad\qquad (2.1)$$
$$a_k=\begin{cases}1 & \text{if } \underline{x_i}'K(p_{k-1})<y_i\le\underline{x_i}'K(p_k)\\ 0 & \text{if } otherwise\end{cases}$$

Let $L(\underline{p_1}),L(\underline{p_2}),\cdots,L(\underline{p_k})$ be the least squares estimators based on those observations with $a_1=1,a_2=1,\cdots,a_k=1$ respectively. That is,

$$\begin{aligned} L(\underline{p_1}) &= (X'A_1X)^{-1}X'A_1\underline{y}\ , \\ L(\underline{p_2}) &= (X'A_2X)^{-1}X'A_2\underline{y}\ , \\ &\vdots \\ L(\underline{p_k}) &= (X'A_kX)^{-1}X'A_k\underline{y}\ , \end{aligned} \tag{2.2}$$

where $A_i=$ diag($a_i, a_i, \cdots a_i$) for $i=1,2,\cdots,k$ and X design matrix with rows $\underline{x_i}'$ for $i=1,2,\cdots,n.$

Then our estimator has the following form:

$$B_k=w_1L(\underline{p_1})+w_2L(\underline{p_2})+\ldots+w_kL(\underline{p_k}), \text{ with } \sum_{i=1}^{k}w_i=1.$$

Therefore, our slope estimator can be obtained by just deleting the intercept part from $L(\underline{p_1}),\cdots,L(\underline{p_k})$. Let us denote these $(p-1)$ dimension estimators as $L_0(\underline{p_1}),\cdots,L_0(\underline{p_k})$. Then our slope estimator has the following form:

$$C_k=w_1L_o(\underline{p_1})+w_2L_0(\underline{p_2})+\ldots+w_kL_0(\underline{p_k}), \text{ with } \sum_{i=1}^{k}w_i=1. \tag{2.3}$$

However, we state necessary theorem to prove asymptotic property of our estimator.

Theorem 2.1 Fix k and $p_0,p_1,\cdots,p_k$ such that $p_1-p_0=p_2-p_1=\ldots=p_k-p_{k-1}=q_k$ then

$$\sqrt{n}[B_k-\underline{\beta}-\sum_i w_i\underline{\delta}(\underline{p_i})] \xrightarrow{D} N_p(0,(\underline{w}'\Sigma\underline{w})Q^{-1})$$

where $\underline{w}=(w_1,w_2,\cdots,w_k)'$, and for $i=1,2,\cdots,k$ and $i\le j\le k$

$$\underline{\delta}(\underline{p_i})=\left(q_k^{-1}\int_{\xi_{i-1}}^{\xi_i}xdF(x),0,\cdots,0\right)'$$

$$\Sigma=(\sigma_{ij})_{k\times k}\ ,$$

with

$$\sigma_{ii} = q_k^{-2}\left\{-2\int_{\xi_{i-1}}^{\xi_i} xF(x)dx + 2\xi_i \int_{\xi_{i-1}}^{\xi_i} F(x)dx - [\int_{\xi_{i-1}}^{\xi_i} F(x)dx]^2\right\},$$

$$\sigma_{ij} = \sigma_{ji} = q_k^{-2}\left\{\xi_j \int_{\xi_{i-1}}^{\xi_i} F(x)dx - 2\xi_i \int_{\xi_{i-1}}^{\xi_i} F(x)dx - \int_{\xi_{i-1}}^{\xi_i} F(x)dx \int_{\xi_{j-1}}^{\xi_j} F(x)dx\right\}\Bigg\},$$

Proof See proof of theorem 2.1 of Han(1993).

From theorem 2.1 we deduce the fact that $\sqrt{n}(C_k - \underline{\beta_0}) \xrightarrow{D} N_{p-1}(0,(\underline{w}'\Sigma\underline{w})Q_0^{-1})$, where $\underline{\beta_0}$ is obtained from $\underline{\beta}$ by deleting first component and Q_0^{-1} is obtained from Q^{-1} by deleting first row and column. Next we want to find minimizing weight $\underline{w}$ of the asymptotic variance of C_k. But usually it is very difficult to estimate the asymptotic variance of C_k. So we adopt the similar idea of Johns(1974) to approximate variance-covariance structure Σ and then find minimizing weight $\underline{w}$. We approximate. for $i = 1,2,\cdots,k,$ as follows;

$$\sigma_{ij} \sim q_k^{-2} b_i(1-b_i) d_i^2 \tag{2.7}$$

$$\sigma_{ij} \sim q_k^{-2} b_i(1-b_i) d_i d_j \tag{2.8}$$

where $d_i = \xi_i - \xi_{i-1}$, $b_i = p_{i-1} + \frac{1}{2} q_k$. Then the approximated variance-covariance structure of Σ is given as follows;

$$B = q_k^{-2} D \begin{bmatrix} b_1(1-b_1) & b_1(1-b_2) & \cdots & b_1(1-b_k) \\ b_2(1-b_2) & b_2(1-b_2) & \cdots & b_2(1-b_k) \\ \vdots & \vdots & & \vdots \\ b_1(1-b_k) & b_2(1-b_k) & \cdots & b_k(1-b_k) \end{bmatrix} D,$$

where D = diag($d_1, d_2, \cdots, d_k$).

Next we want to minimize the following quantity;

$$\min_{\underline{w}'\underline{1}=1} \{ (\underline{w}' B \underline{w}) Q_0^{-1} \}$$

where $\underline{1} = (1,1,\cdots,1)'$. Because Q_0^{-1} is fixed, we want to minimize $\underline{w}'B\underline{w}$ under the constraint $\underline{w}'\underline{1} = 1$. By a straightforward Lagrange Multiplier argument, we establish that minimizing weight is

$$\underline{w} = (\underline{1}' B^{-1} \underline{1})^{-1} B \tag{2.9}$$

and

$$\min_{\underline{w}'\underline{1}=1}\{\underline{w}'B\underline{w}\}=(\underline{1}'B^{-1}\underline{1})^{-1} \tag{2.10}$$

With the minimizing weight given above, we define our slope estimator as given by (2.3). If we define

$$e_1=1/d_1(b_2/b_1d_1-1/d_2)q_k,$$
$$e_i=1/d_i(2/b_i-1/d_{i-1}-1/d_{i+1})q_k, i=2,3,\ldots,k-1,$$
$$e_k=1/d_k(-1/d_k-(1-b_{k-1})/(1-b_k)d_k)q_k,$$

then by straightforward calculation, we have the minimizing weight as

$$w_i=e_i/\left\{\sum_{i=1}^{k}e_i\right\},\quad i=1,2,\ldots,k.$$

Let e_i^* be a consistent estimator of e_i by simply substituting d_i for corresponding consistent estimate $\widehat{d_i}$. Then if we define $w_i^*=e_i^*/\left\{\sum_{i=1}^{k}e_i^*\right\}$, $(i=1,2,\ldots,k)$, then our estimator is of the form

$$\widehat{C_k}=w_1^*L_0(\underline{p_1})+w_2^*L_0(\underline{p_2})+\ldots+w_k^*L_0(\underline{p_k}). \tag{2.11}$$

Han(1993) shows that for all k, $\sqrt{n}(C_k-\underline{\beta_0})$ and $\sqrt{n}(\widehat{C_k}-\beta_0)$ have same limiting distribution by using residuals after fitting preliminary fit $\underline{\widehat{\beta_0}}$ which satisfies A5. Next theorem gives proof of asymptotic efficiency of our slope estimator C_k hence of $\widehat{C_k}$ when trimming is given.

Before starting the theorem, we briefly mention Rao-Cramer Lower bound $I(F,p_0,p_k)$ when $[(n+1)p_0]$ and $[(n+1)(1-p_k)]$ observations has been trimmed from left and right respectively. This lower bound is proposed by Placket(1958) and defined by

$$I(F,p_0,p_k)=\int_{\xi_0}^{\xi_k}(f'/f)^2 f dx+\frac{f^2(\xi_0)}{p_0}+\frac{f^2(\xi_k)}{(1-p_k)}\ .$$

Theorem 2.2 Let $p_1-p_0=p_2-p_1=\ldots=p_k-p_{k-1}=q_k$ Then, for given $\varepsilon>0$, there exists N such that for all $k\geq N$ and $F\in\Lambda$,

$$n^{1/2}(C_k-\underline{\beta_0})\xrightarrow{d}N(0,\sigma^2(F)Q_0^{-1}),$$

where $\sigma^2(F)\leq[I(F,p_0,p_k)]^{-1}+\varepsilon$,

Proof Let's write $\underline{w}'\Sigma\underline{w}=\underline{w}'B\underline{w}+\underline{w}'R\underline{w}$. We first show that with this minimizing weight $\underline{w}$ given in (2.9) , $\underline{w}'B\underline{w}\rightarrow[I(F,p_0,p_k)]^{-1}$ for k sufficiently large and then we show that $\underline{w}'R\underline{w}\to 0$ for sufficiently large k. By straightforward calculation, B^{-1} is

$$B^{-1}=q_k^2D^{-1}GD^{-1}$$

where $G=(g_{ij})_{k\times k}$ and

$$\begin{aligned} g_{i,i+1}=g_{i+1,i}&=\frac{-1}{(b_{i+1}-b_i)},i=1,2,\dots,k-1,\\ g_{11}&=\frac{b_2}{b_1(b_2-b_1)},\\ g_{ii}&=\frac{(b_{i+1}-b_{i-1})}{(b_i-b_{i-1})(b_{i+1}-b_i)},i=2,3,\dots,k-1,\\ g_{kk}&=\frac{1-b_{k-1}}{(1-b_k)(b_k-b_{k-1})}, \end{aligned} \tag{2.12}$$

and all other $g_{ij}=0$. Also by strightforward calculation, the minimum of $\underline{w}'B\underline{w}$ given in (2.10) is

$$(\underline{1}'B^{-1}\underline{1})^{-1}=q_k\left\{\frac{1}{d_1}\left(\frac{b_2}{b_1d_1}-\frac{1}{d_2}\right)+\frac{1}{d_2}\left(\frac{2}{d_2}-\frac{1}{d_1}-\frac{1}{d_3}\right)+\cdots+\frac{1}{d_k}\left[-\frac{1}{d_k}-\frac{(1-b_{k-1})}{(1-b_k)d_k}\right]\right\}. \tag{2.13}$$

Let $F^{-1}(u)=G(u)$ and $r_i=p_0+(i-1)q_k$, $i=1,2,\dots,k$. Then we have

$$d_i=G(r_i+q_k)-G(r_i),i=1,2,\dots,k,$$

$$d_{i+1}=G(r_i+2q_k)-G(r_i+q_k),i=1,2,\dots,k-1$$

$$d_{i-1}=G(r_i)-G(r_i-q_k),i=2,3,\dots,k.$$

Now let G_i',G_i'', and G_i''' represent the first three derivatives of G evaluated at r_i. Then, by a straight forward Taylor's series expansion in terms of q_k, we obtain

$$
\begin{aligned}
d_i^{-1} &= q_k^{-1}(G_i')^{-1} - \frac{1}{2}(G_i')^{-2}G_i'' + q_k[\frac{1}{4}(G_i'')^2(G_i)^{-3} \\
&\quad - \frac{1}{6}G_i'''(G_i')^{-2}] + \xi q_k^2, \\
d_{i+1}^{-1} &= q_k^{-1}(G_i')^{-1} - \frac{3}{2}(G_i')^{-2}G_i'' + q_k[\frac{9}{4}(G_i'')^2(G_i')^{-3} \\
&\quad \frac{-7}{6}G_i'''(G_i')^{-2}] + \xi q_k^2, \\
d_{i-1}^{-1} &= q_k^{-1}(G_i')^{-1} + \frac{1}{2}(G_i')^{-2}G_i'' + q_k[\frac{1}{4}(G_i'')^2(G_i')^{-3} \\
&\quad - \frac{1}{6}G_i'''(G_i')^{-2}] + \xi q_k^2,
\end{aligned}
\tag{2.14}
$$

where ξ is a generic uniformly bounded function of i, k and $F \in \wedge$. Hence for $i = 2, 3, \ldots, k-1$,

$$
\begin{aligned}
e_i &= \frac{1}{d_i}(\frac{2}{d_i} - \frac{1}{d_{i-1}} - \frac{1}{d_{i+1}}) \\
&= [2(G_i'')^2(G_i')^{-4} - G_i'''(G_i')^{-3}] + \xi q_k.
\end{aligned}
\tag{2.15}
$$

Therefore, if f_i, f_i' and f_i'' are f_i and its first two derivatives evaluated at $G(r_i)$, from (2.15), we have

$$
e_i = -[f_i''(f_i)^{-1} - (f_i')^2(f_i)^{-2]} + \xi q_k, \qquad i = 2, 3, \ldots, k-1 \tag{2.16}
$$

Note also that $d_k = G(p_k) - G(p_k - q_k)$ and $d_{k-1} = G(p_k - q_k) - G(p_k - 2q_k)$.

Therefore, from (2.14), we have

$$
e_k = q_k^{-1}(f_k' + \frac{f_k^{\,2}}{(1-p_k)} + \xi q_k) . \tag{2.17}
$$

Similarly , from (2.14), we have

$$
e_1 = q_k^{-1}(f_0' + \frac{f_0^{\,2}}{p_0} + \xi q_k) . \tag{2.18}
$$

Therefore, combining (2.16),(2.17) and (2.18) and using the definition of the Riemann integral, we have

$$
\begin{aligned}
q_k \sum_{i=1}^{k} e_i \rightarrow & \int_{p_0}^{q_0} \left\{ \frac{f'(G(u))}{f(G(u))} - [\frac{f'(G(u))}{f(G(u))}]^2 \right\} du + \\
& \frac{f_1^{\,2}}{p_0} + \frac{f_k^{\,2}}{(1-p_k)} + f_k' - f_1' \\
& - \int_{\xi_0}^{\xi_k} (\frac{f'}{f})^2 f dx + \frac{f_1^{\,2}}{p_0} + \frac{f_k^{\,2}}{(1-p_k)} \\
& = I(F, p_0, p_k)
\end{aligned}
\tag{2.19}
$$

for sufficiently large k. Hence, we have $\underline{w}'B\underline{w} \to [I(F,p_0,p_k)]^{-1}$ for sufficiently large k.

Next, we want to show that $\underline{w}'R\underline{w} \to 0$ for sufficiently large k.

Let $R=(r_{ij})_{k\times k}$. Then we can write

$$r_{ii}=q_k^{-2}\Big\{2G(r_i+q_k)\int_{G(r_i)}^{G(r_i+q_k)}F(x)dx-2\int_{G(r_i)}^{G(r_i+q_k)}xF(x)dx-\Big[\int_{G(r_i)}^{G(r_i+q_k)}F(x)dx\Big]^2-b_i(1-b_i)d_i^{\,2}\Big\}$$

and

$$r_{ij}=r_{ji}=q_k^{-2}\Big\{-G(r_j)\int_{G(r_i)}^{G(r_i+q_k)}F(x)dx+G(r_j+q_k)\int_{G(r_i)}^{G(r_i+q_k)}F(x)dx-\int_{G(r_i)}^{G(r_i+q_k)}F(x)dx\int_{G(r_j)}^{G(r_j+q_k)}F(x)dx-b_i(1-b_j)d_id_j\Big\}. \quad (2.20)$$

Expanding Taylor series of the above r_{ii} and r_{ij} with respect to q_k and noting that $b_i=r_i+.5q_k$, we get

$$\begin{aligned}r_{ii}=q_k^{-2}\{&(2G_i+2G_i'q_k+\xi q_k^2)(r_iG_i'q_k+\frac{(r_iG_i''+G_i')}{2}q_k^2+\xi q_k^3)-2r_iG_iG_i'q_k-\\&[r_i(G_i')^2+G_iG_i'+r_iG_iG_i'']q_k^2+r_i^2(g_i')^2q_k^2-\\&r_i(1-r_i)(G_i')^2q_k^2+\xi q_k^3\},\end{aligned} \quad (2.21)$$

where ξ is a generic uniformly bounded functions for all i, k and $F\in\wedge$.

Therefore by simplifying r_{ii} we get $r_{ii}=0(q_k)$ uniformly for all i, k and $F\in\wedge$. Similary , we have

$$r_{ij}=q_k^{-2}\{r_iG_i'G_j'q_k^2-r_ir_jG_i'G_j'q_k^2-r_i(1-r_j)G_i'G_j'q_k^2+\xi q_k^3\}. \quad (2.22)$$

Therfore by simplifying r_{ij} we get $r_{ij}=0(q_k)$ uniformly for all i, j , k and $F\in\wedge$.

Note that $\underline{1}'B^{-1}\underline{1}\to I(F,p_0,p_k)$ for sufficiently large k. Also by straightforward calculation, we have $B^{-1}\underline{1}=q_k\underline{e}$ and we deduce that with the minimizing weight given in (2.9), $w_1=w_k=0(1), w_2=w_3=\ldots=w_{k-1}=0(q_k)$. Therefore $\underline{w}'R\underline{w}\to 0$ for sufficiently large k. This completes the proof of theorem 2.2.

References

[1] Andrews, D.F. , Bickel , P.J. , Hampel, F.R. , Huber, P.J., Rogers, W.H. and Tukey, J.W.(1972). *Robust Estimates of Location: Survey and Advances,* Princeton University Press.

[2] Han,S.M.(1993). Adaptive L-Estimation For Regression Slope under Asymmetric Error Distributions, *The Korean Journal of Applied Statistics,* Vol. 6 , 79-93

[3] Johns, M.V. (1974). Nonparametric Estimation of Location, *Journal of the American Statistical Association,* Vol. 69, 453-460

[4] Koenker, R., and Bassett,G. (1978). Regression Quantiles, *Econometrica,* Vol. 46, 33-50.

[5] Placket, R.L. (1958), Linear estimation from censored data, *Annals of Mathematical Statistics,* Vol. 29, 131-142.

[6] Ruppert, D., and Carroll, R.J. (1978). Robust regression by trimmed least squares estimation, *Institute of Statistics Mimeo Series,* 1186, University of North Carolina at Chapel Hill, Dept. of Statistics.

[7] ________(1980). Trimmed least squares estimation in the linear model, *Journal of the American Statistical Association,* Vol. 75, 828-838.

[8] Sacks, J. (1975). An asymptotically efficient sequence of estimators of a location parameter, *Annals of Statistics,* Vol. 3, 285-298.

절사가 주어질때 회귀기울기의 점근적 최량 L-추정법

한상문[2)]

요약

Han (1993)의 임의의 오차분포하에서 회귀모형에의 기울기 추정법을 응용하여 회귀분위선(regression quantile)에 의해 적당한 상·하위절사가 주어질때 점근적으로 최량의 회귀모형에서의 기울기 추정량을 구성할 수 있음을 보였다.

2) (130-743) 서울시 동대문구 전농동 90 서울시립대학교 전산통계학과.