Journal of the Korean
Data & Information Science Society
2000, Vol. 11, No. 2, pp. 235 ~ 245

비선형최소분위추정량의 점근적 성질

최 승 회 [1] · 김 태 수 [2] · 박 경 옥 [3]

요 약

두 변수간의 함수관계를 연구하는 회귀분석에서 모수를 추정하기 위하여 가장 널리 사용되는 방법은 최소자승법이다. 그러나 최소자승법은 표본 평균처럼 약간의 이상치에도 민감하게 반응하여 강인성(robustness)을 만족하지 못함으로 새로운 추정량이 필요하다. 본 논문에서는 최소분위추정량과 최소분위추정량에 근거한 일차결합추정량의 점근적 성질을 연구하였다. 또한 최소자승추정량에 대해 제시된 추정량의 점근적 효율성을 구하고 모의실험을 통하여 최소분위추정량의 효율성을 조사하였다.

주제어: 비선형회귀모형, 최소분위추정량, 정규성, 효율성.

1. 서 론

본 논문에서 연구할 비선형 회귀모형은 다음과 같다.

$$y_t = f(x_t, \theta_0) + \epsilon_t\ ,\quad t = 1, \cdots, n. \tag{1.1}$$

여기서 x_t는 t번째의 고정된 입력벡터이고, 모수 θ_0는 모수공간 $\Theta \subset R^p$에 속하는 $p \times 1$ 미지열벡터이다. 또 반응함수 f는 $R^q \times R^p$에서 R^1으로 가는 연속함수이며, 오차항 ϵ_t는 서로 독립이고 같은 분포를 따르는 확률변수 로서 평균은 0이고 유한한 분산을 가지며 연속인 밀도함수 $g(x)$을 가진다.

비선형회귀모형에 관한 통계적인 문제는 독립변수 (x_t)와 종속변수 (y_t)의 관찰치를 기초로 하여 최적의 방법으로 미지모수 θ를 추정 또는 검정하는 것이다. 회귀분석에서 모수의 실제값 θ_0를 추정하는 가장 일반적인 방법인 최소제곱(Least Squares : LS)추정량은

$$L_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}(y_t - f(x_t, \theta))^2$$

[1]경기도 고양시 화전동 한국항공대학교 교양학부 조교수
[2]경기도 안산시 사동 한양대학교 산업공학과 연구교수
[3]경기도 수원시 매탄3동 삼성전자(주) 디지털프린팅(사) 선임연구원

의 값을 최소로 하는 모수의 값으로 정의된다. 비선형회귀모형에 LS추정량($\tilde{\theta}_n$)은 Jennrich (1969)와 Wu (1981)에 의하여 소개되었고 이 두 사람은 최소제곱추정량에 대한 여러가지 성질을 연구하였다. 여러 문헌에서 언급된것 처럼 오차분포가 정규분포일 경우 최소제곱추정량은 효율면에서 최적이다. 하지만 표본 평균에 근거한 최소제곱추정량은 오차항이 몇개의 이상값(outlier)을 갖거나 정규분포의 조건에서 조금이라도 위배될 경우 매우 민감하여 강인성을 만족하지 못한다. 이런 단점을 극복하기 위해 Oberhofer(1982)는

$$A_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}|y_t - f(x_t,\theta)|$$

을 최소로 하는 최소절대편차(Least Absolute Deviation : LAD) 추정량을 제시하였다. Wang(1995)과 Kim 과 Choi(1995)등이 비선형회귀모형에 대한 LAD추정량($\check{\theta}_n$)에 대한 연구을 하여 오차의 확률분포가 코시분포, 이중지수분포, 로지스틱분포 등과 같이 꼬리부분이 비교적 두꺼운 경우이거나 중앙값 근방에서 밀도가 적게 밀집된 상태로 주어지는 경우에 최소절대편차추정량은 최소제곱추정량보다 상대적으로 더 효율적임을 설명하였다. 그러나 비선형 LAD추정량의 점근적 성질을 만족시켜 주는 충분 조건중 하나인 $G(0) = \frac{1}{2}$는 오차항의 구조가 0을 중심으로 대칭인 분포는 만족하지만 분포함수가 다음과 같은 이중지수 (double exponential)분포

$$g(x) = \frac{1}{2}\exp[-|x-1|], -\infty < x < \infty$$

과 같을 경우는 $G(0) = \frac{1}{2e} \neq \frac{1}{2}$이므로 LAD추정양의 비효율성은 적당한 방법으로 보정하여 주어야 한다. 비대칭분포에 대한 LAD추정량을 보완하기 위해 Koenker와 Basset(1978)는 표본분위수를 선형회귀모형에 일반화 한 회귀분위(Regression Quantile : RQ) 추정량을 소개하였고 다음 함수를

$$Q_n(\theta : \beta) = \frac{1}{n}\sum_{t=1}^{n}\rho_\beta(y_t - f(x_t,\theta)) \tag{1.2}$$

최소화 하는 값을 β-회귀분위추정량($\beta - RQ$)이라 하였다. 여기서 함수 ρ_β는

$$\rho_\beta(\omega) = \begin{cases} \beta\omega, & \omega \geq 0 \\ (\beta-1)\omega, & \omega < 0 \end{cases}$$

이고 $0 < \beta < 1$ 이다. 선형회귀모형에 대한 RQ추정량의 성질은 Basset과 Koenker(1986), Powell(1986), 그리고 Portnoy(1991)등 여러 사람에 의하여 연구 되었다.

본 논문에서는 RQ추정량($\hat{\theta}_n(\beta)$)의 일차 결합 형태인

$$C_n(\beta) = \sum_{i=1}^{M}\pi(\beta_i)\hat{\theta}_n(\beta_i)$$

의 점근적 성질을 만족시키는 충분 조건에 대하여 연구한다. 여기서

$$\pi(\beta) = (\pi(\beta_1), \pi(\beta_2), \cdots, \pi(\beta_M))' \tag{1.3}$$

이고, $\pi(\beta_i)$ 는 이산이고 대칭인 확률측도이며 $0 < \beta_1 < \beta_2 < \cdots < \beta_M < 1$ 이다. 식(1.3)에서 $\pi(\frac{1}{2}) = 1$인 경우 $C_n(\beta)$는 LAD추정량이고 $\pi(\beta_i) = 1$인 경우는β_i -RQ추정량임을 알 수 있다.

본 논문의 구성은 2절에서 최소분위추정량과 최소분위추정량에 근거한 일차결합추정량의 점근적 성질을 알아 보고, 제시된 추정량의 효율성을 유도하였다. 지금까지 제시된 LS 추정량, LAD 추정량, RQ 추정량, 그리고 $C_n(\beta)$추정량에 대한 모의실험 결과를 3절에서 소개하였다. 그리고 4절 부록에서는 2절에서 소개된 정리들에 대한 증명을 하였다.

2. 점근적 성질

본 절에서는 식 (1.2)에서 정의된 최소분위추정량에 근거한 일차결합 추정량의 점근적 정규성을 만족하는 충분조건에 대하여 알아보자. $(\Gamma, \mathcal{A}, P_x)$을 R^q위에서 정의된 확률공간이라 하고 $G(x)$을 P_x에 의하여 유도된 분포함수라 하자. 본 논문에서 사용되는 기호를 간단히 하기 위하여 다음과 같이 정의하자.

$$f_t(\theta) = f(x_t, \theta), \nabla f_t(\theta) = \left[\frac{\partial}{\partial \theta_i} f_t(\theta)\right]_{(p\times 1)}, \nabla^2 f_t(\theta) = \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} f_t(\theta)\right]_{(p\times p)}.$$

본 논문에서 다음과 같은 가정이 비선형모형 (1.1)에 성립한다고 가정하자.

Assumption A

A_1 : 모수공간 Θ 는 R^p의 컴팩트 (compact) 부분집합이고 Γ는 R^q의 유계집합이다.

A_2 : 모든 θ 에 대해 $\nabla f_t(\theta)$ 와 $\nabla^2 f_t(\theta)$ 가 존재하고, $\nabla f_t(\theta)$는 $\Gamma \times \Theta$에서 연속이다 .

A_3 : 모수공간에 속하는 임의의 $\theta_1 \neq \theta_2$에 대하여 $P_x\{x \in \Gamma : f(x, \theta_1) \neq f(x, \theta_2)\} > 0$ 이다.

Assumption B

B_1 : 오차의 밀도함수 $g(x)$는 0 근방에서 양수이다.

B_2 : 행렬 $V_n(\theta_0) = \frac{1}{n}\sum_{t=1}^{n} \nabla f_t(\theta_0) \nabla f_t^T(\theta_0)$는 어떤 양의 정칙행렬 $V(\theta_0)$로 수렴한다

오차의 분포함수가 $G(\eta_{\beta_i}) = \beta_i$를 만족하면 식 (1.2)를 최소로 하는 θ값은 다음 목적함수

$$R_n(\theta : \beta) = \frac{1}{n}\sum_{t=1}^{n} \rho_\beta(y_t + \eta_{\beta_i} - f(x_t, \theta)) \tag{2.1}$$

을 최소로 하는 값과 같다. Choi, Kim 그리고 Park(1999)에 의해서 증명된 다음 보조정리는 β_i -RQ추정량의 강 수렴성과 정규성을 설명한다.

정리 2.1 비선형모형 (1.1)이 가정 A 와 가정 B을 만족하면

(1) $\hat{\theta}_n(\beta_i) \xrightarrow{a.s.} \theta_0$,

(2) $\sqrt{n}(\hat{\theta}_n(\beta_i) - \theta_0) \xrightarrow{L} N\left(0, \frac{\beta_i(1-\beta_i)}{(g(\eta_{\beta_i}))^2} V^{-1}(\theta_0)\right)$

가 성립한다. 여기서 $\xrightarrow{a.s.}$와 $\xrightarrow{L}$ 는 각 각 강 수렴과 분포수렴을 표시한다.

정리 2.1 에서는 한개의 회귀분위 추정량에 대한 점근적 성질을 연구하였다. β -RQ추정량에 근거한 일차결합 추정량의 성질을 알아 보기위해 m개의 회귀분위 추정량에 대한 성질을 다음 정리에서 설명한다.

정리 2.2 가정 A 와 가정 B가 비선형모형 (1.1)에서 성립하면

$$\sqrt{n}(\hat{\theta}_n(\beta_1) - \theta_0, \cdots, \hat{\theta}_n(\beta_m) - \theta_0) \xrightarrow{L} N(0, \Omega \otimes V^{-1}(\theta_0))$$

이다. 단,

$$\Omega = \begin{bmatrix} g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_1})\beta_1(1-\beta_1) & \cdots & g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_m})\beta_1(1-\beta_m) \\ \vdots & \ddots & \vdots \\ g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_m})\beta_1(1-\beta_m) & \cdots & g^{-1}(\eta_{\beta_m})g^{-1}(\eta_{\beta_m})\beta_m(1-\beta_m) \end{bmatrix}$$

이고, $\otimes$는 크로네커 곱(kronecker multiplication)이다.

증명 먼저 $D_i = (d_{i1}, \cdots, d_{ip})$라 하고 $\underline{D} = (D_1, \cdots, D_m)$는 $R^{mp\times 1}$의 원소라 하자. 부록에서 $\sum_{i=1}^m C_i\sqrt{n}(\check{\theta}_n(\beta_i) - \theta_0)$가 평균이 0이고 분산-공분산 행렬이

$$\underline{D} \begin{bmatrix} g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_1})\beta_1(1-\beta_1) & \cdots & g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_m})\beta_1(1-\beta_m) \\ \vdots & \ddots & \vdots \\ g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_m})\beta_1(1-\beta_m) & \cdots & g^{-1}(\eta_{\beta_m})g^{-1}(\eta_{\beta_m})\beta_m(1-\beta_m) \end{bmatrix} \otimes V^{-1}(\theta_0)\underline{D^T}$$

인 분포로 수렴함을 자세히 보인다.

다음 정리는 정리 2.2을 이용하여 LAD추정량과 RQ 추정량을 특별한 경우로 가지는 회귀분위추정량의 일차 결합 형태인$C_n(\beta)$의 정규성을 보여주며 증명은 부록으로 넘긴다.

정리 2.3 정리 2.2와 같은 조건에서 비선형 RQ추정량에 근거한 일차결합 추정량 $C_n(\beta)$은 평균이 0이고 분산-공분산 행렬이$\pi'\Omega\pi V^{-1}(\theta_0)$인 분포로 수렴한다. 즉

$$\sqrt{n}(Q_n(\pi) - \theta_0) \xrightarrow{L} N\left(0, \pi'\Omega\pi V^{-1}(\theta_0)\right)$$

이다.

선형회귀모형에 대한 RQ추정량을 연구한 Koenker와 Basset(1978)는 정리 2.3에서 소개된 일차결합추정량 $C_n(\beta)$의 가중함수 (weighted function) 로서

$$(\pi(\frac{1}{3}), \pi(\frac{1}{2}), \pi(\frac{2}{3})) = (0.3, 0.4, 0.3)$$

와

$$(\pi(\frac{1}{4}), \pi(\frac{1}{2}), \pi(\frac{3}{4})) = (0.25, 0.5, 0.25)$$

을 제시하였다. 이와 같은 가중함수를 이용하여 다음절에 모의실험을 실시하였다.

정리 2.3에서 $\pi(\beta_i) = 1$인 경우는 정리 2.1의 결과와 일치함을 알 수 있다. 또한

$$e_1 = \sum_{i=1}^{m} \frac{\beta_i(1-\beta_i)}{g(\eta_{\beta_i})g(\eta_{\beta_i})}\pi^2(\beta_i), \quad e_2 = 2\sum_{i<j} \frac{\beta_i(1-\beta_j)}{g(\eta_{\beta_i})g(\eta_{\beta_j})}\pi(\beta_i)\pi(\beta_j)$$

라 두면, Choi, Kim 그리고 Park(1999)의 4절 결과로 부터 최소제곱(LS)추정량과 최소절대편차 (LAD)추정량에 대한 비선형RQ추정량에 근거한 일차결합 추정량 $C_n(\beta)$의 점근적 효율성은 각 각

$$\frac{e_1 + e_2}{\frac{1}{4g^2(0)}}, \quad \frac{e_1 + e_2}{\sigma^2}$$

이다. 따라서 오차항의 구조가 평균이 1이고 분산이 1인 정규분포 N(1,1)와 평균이 1이고 분산이 2인 이중지수분포 DE(1,2)와 같이 비 대칭인 경우에는 RQ추정량이 더 효율점임을 알 수 있다. 다음 절에서 이와 같은 결과에 대한 모의실험을 제시한다.

3. 모의실험

본 절에서는 최소자승추정량, 최소절대편차추정량, 최소분위추정량, 최소분위추정량에 근거한 일차결합추정량에 효율성을 모의실험을 통하여 알아본다. 식 (1.1)에 제시된 비선형회귀모형에 대한 예로 지수함수 $\alpha e^{\beta x_t}$을 이용하였고 오차항은 다음과 같은 분포함수에서 표본을 사용하였다.

(1) 표준정규분포(standard normal distribution, N(0,1)),

(2) 대칭인 이중지수분포(symmetric double exponential distribution, DE(0, 3)),

(3) 비대칭인 이중지수분포 (asymmetric double exponential distribution, DE(1, 3)).

또한, 모수는 $\alpha = 1, \beta = 1,000$이고 독립변수는 $x_t = t$인 경우에 한하여 30개의 표본을 사용하여 LS, LAD, RQ, 최소분위추정량에 근거한 일차결합추정량(LCRQ)의 추정치와 참값과 추정치에 대한 평균제곱오차(MSE)을 구하였다. 다음 표 3.1은 지수함수의 모형에서 비선형성을 잘 선명하여주는 α에 대한 결과이다.

Table 3.1. 추정량과 평균제곱오차

Distribution	N(0,1)	DE(0,3)	DE(1,3)
LS	**1.000008**	1.000224	0.994256
	(0.000004)	(0.000072)	(0.000102)
LAD	1.000026	**1.000183**	0.993965
	(0.000006)	(0.000072)	(0.000118)
RQ	1.000206	1.001506	0.993378
	(0.000016)	(0.000218)	(0.00033)
LCRQ1	0.999966	1.000528	0.994316
	(0.0000005)	(0.00073)	(0.000107)
LCQR2	1.000029	1.000826	**0.994336**
	(0.000005)	(0.000073)	(0.000105)

표 3.1에서 RQ추정량은 1.000번을 실험할 때 마다 난수를 이용하여 β 값을 임의로 정한 결과이고 LCRQ1는 가중함수를 (0.3, 0.4, 0.3)로 LCQR2 (0.25, 0.5, 0.25)로 주었을 경우의 추정량이다.

표 3.1에서 확인한것 처럼 정규분포의 경우는 최소자승법이 또한 대칭인 이중지수분포의 경우는 최소절대편차추정량이 참값에 가장 근사한 값을 가진다. 그리고 오차항의 분포가 비대칭이며서 꼬리부분이 비교적 두꺼운 경우에 는 최소분위추정량에 근거한 일차결합추정량이 다른 추정량보다 참값을 잘보여주고 있다.

만약 오차항 분포의 $G(0)$값을 알 경우는 LCRQ추정량 보다는 최소분위추정량($RQ(\beta)$)이 더 효율적이다. 표 3.1에서 주어진 비대칭인 이중지수분포는 $G(0) = \frac{1}{3}$을 만족한다. β가 $\frac{1}{3}$인 최소분위추정량($RQ(\frac{1}{3})$)의 추정치는 0.997165이고 평균제곱오차는 0.000104이므로 다른 추정량보다 더 참값에 근접함을 알 수 있다.

결론적으로, 표준정규분포일 경우는 최소제곱추정량이, 꼬리가 두껍고 대칭인 경우는 최소분위추정량이 효율적이며 비대칭이면서 두꺼운 꼬리를 가질 경우는 최소분위추정량이 효율적이다.

4. 부록

정리 2.2의 증명 식 (2.1)에서 정의된 $Q_n(\theta,\beta_i)$는 θ에 관한 일계 도함수는 연속이 아니므로, 점근적으로는 $\rho_\beta(x)$와 유사한 다음과 같이

$$s_\beta(x) = \begin{cases} -\frac{1}{2}\alpha_n^2x^3 + \frac{\alpha_n}{4}x^2 + \beta x + \frac{1}{4\alpha_n} & |x| \le \frac{1}{\alpha_n} \\ \beta x & x > \frac{1}{\alpha_n} \\ (\beta-1)x & x < -\frac{1}{\alpha_n}. \end{cases}$$

함수를 생각하자. 여기서 α_n은 0보다 큰 어떤 δ에 대하여 $\frac{n^2}{\alpha_n^3}$ 와 $\frac{\alpha_n}{n}$ 그리고 $\frac{n^{\frac{1+\delta}{2}}}{\alpha_n}$ 을 만족한다. 특히 $\alpha_n = \frac{1}{n^p}(\frac{2}{3} < p < 1)$이면 위와 같은 조건을 만족한다. 그리고 $R_n^*(\theta,\beta_i)$를

$$R_n^*(\theta,\beta_i) = \frac{1}{n}\sum_{t=1}^{n} S_{\beta_i}(y_t - \eta_{\beta_i} - f(x_t,\theta_0))$$

로 정의하면 $R_n^*(\theta,\beta_i)$는 2계까지 미분가능한 함수로서 점근적으로는 $R_n(\theta,\beta_i)$와 같다.

Choi, Kim 그리고 Park(1999)의 정리3.1의 증명으로 부터 우리는 다음과 같은 사실을 알 수 있다.

(1) $\sqrt{n}(\hat\theta_n(\beta_i) - \theta_0) = -[\nabla^2 R_n^*(\bar\theta_n(\beta_i))]^{-1}\sqrt{n}\nabla R_n^*(\theta_0)$

(2) $\nabla^2 R_n^*(\bar\theta_n(\beta_i)) \xrightarrow{p} g(\eta_{\beta_i})V(\theta_0)$

(3) $$\begin{aligned}\sqrt{n}\nabla R_n^*(\theta_0) = &- \frac{1}{n}\sum_{t=1}^{n}\sqrt{n}\left[\frac{\alpha_n}{2}(r_t(\theta_0,\beta_i)) + (\beta_i - \frac{1}{2})\right] I_{\{|r_t(\theta_0,\beta_i)|\le\frac{1}{\alpha_n}\}}\nabla f_t(\theta_0) \\ &- \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left[\beta_i I_{\{r_t(\theta_0,\beta_i)>\frac{1}{\alpha_n}\}} + (\beta_i-1)I_{\{r_t(\theta_0,\beta_i)<-\frac{1}{\alpha_n}\}}\right]\nabla f_t(\theta_0)\end{aligned}$$

(4) $-\frac{1}{n}\sum_{t=1}^{n}\sqrt{n}\left[\frac{\alpha_n}{2}(r_t(\theta_0,\beta_i)) + (\beta_i - \frac{1}{2})\right] I_{\{|r_t(\theta_0,\beta_i)|\le\frac{1}{\alpha_n}\}}\nabla f_t(\theta_0) \xrightarrow{p} 0$

단, $r_t(\theta,\beta_i) = \lambda - d_t(\theta) - \eta_{\beta_i}$ 이고 $|\bar\theta_n(\beta_i) - \theta_0| \le |\hat\theta_n(\beta_i) - \theta_0|$ 이다.

임의의 벡터 D_i에 대해

$$\sum_{i=1}^{m} D_i\sqrt{n}(\hat\theta_n(\beta_i) - \theta_0) = \sum_{i=1}^{m}\left[D_i\sqrt{n}(\hat\theta_n(\beta_i) - \tilde\theta_n(\beta_i)) + D_i\sqrt{n}(\tilde\theta_n(\beta_i) - \theta_0)\right]$$

이 성립하고 정리 2.1과 Cramer-Wold의 법칙으로 부터 각 i에 대해 다음 식을 유도할 수 있다.

$$D_i\sqrt{n}(\hat\theta_n(\beta_i) - \theta_0) \xrightarrow{L} N(0, D_i g^{-1}(\eta_{\beta_i})g^{-1}(\eta_{\beta_i})\beta_i(1-\beta_i)V^{-1}(\theta_0)D_i')$$

그리고, 만약 $i \ne j$ 라면 식 (1),(2),(3),(4)에 의해 $\sqrt{n}(\hat\theta_n(\beta_i) - \theta_0)$ 와 $\sqrt{n}(\hat\theta_n(\beta_j) - \theta_0))$의 공분산은

$$g^{-1}(\eta_{\beta_i})V^{-1}(\theta_0)\times\Big[\Big\{\tfrac{1}{\sqrt{n}}\textstyle\sum_{t=1}^{n}\Big(\beta_i I_{\{r_t(\theta_0,\beta_i)>\frac{1}{\alpha n}\}}+(\beta_i-1)I_{\{r_t(\theta_0,\beta_i)<-\frac{1}{\alpha n}\}}\Big)\nabla f_t(\theta_0)\Big\}\times$$
$$\Big\{\tfrac{1}{\sqrt{n}}\textstyle\sum_{s=1}^{n}\Big(\beta_j I_{\{r_s(\theta_0,\beta_j)>\frac{1}{\alpha n}\}}+(\beta_j-1)I_{\{r_s(\theta_0,\beta_j)<-\frac{1}{\alpha n}\}}\Big)\nabla f_s^T(\theta_0)\Big\}\Big]\times V^{-1}(\theta_0)g^{-1}(\eta_{\beta_j})\quad(4.1)$$

이다. 식 (4.1)을 $t=s$인 경우에 대하여 계산하면

$$\frac{1}{n}\Big[\beta_i\beta_j\int_{\frac{1}{\alpha n}+\eta_{\beta_j}}^{\infty}dG(\lambda)\;+\;\beta_i(\beta_j-1)\int_{\frac{1}{\alpha n}+\eta_{\beta_i}}^{-\frac{1}{\alpha n}+\eta_{\beta_j}}dG(\lambda)$$
$$+\;(\beta_i-1)(\beta_j-1)\int_{-\infty}^{-\frac{1}{\alpha n}+\eta_{\beta_i}}dG(\lambda)\Big]\nabla f_t(\theta_0)\nabla f_t^T(\theta_0)\quad(4.2)$$

이다. 식 (4.2)을 간단히 하면

$$\frac{1}{n}\nabla f_t(\theta_0)\nabla f_t^T(\theta_0)\beta_i(1-\beta_j)+o_p(1)$$

와 같이 된다. 식 (4.1)을 $t\neq s$인 경우에 대하여 계산하면

$$\frac{1}{n}E\;\Big[\;\beta_i I_{\{r_t(\theta_0,\beta_i)>\frac{1}{\alpha n}\}}+(\beta_i-1)I_{\{r_t(\theta_0,\beta_i)<-\frac{1}{\alpha n}\}}\Big]\nabla f_t(\theta_0)$$
$$\times\;E\Big[\beta_j I_{\{r_s(\theta_0,\beta_j)>\frac{1}{\alpha n}\}}+(\beta_j-1)I_{\{r_s(\theta_0,\beta_j)<-\frac{1}{\alpha n}\}}\Big]\nabla f_s^T(\theta_0)\quad(4.3)$$

이 된다. 한편, 오차항 ϵ_t가 서로 독립이고 같은 분포를 따르는 확률변수임으로 다음 식

$$\begin{aligned}E\Big[\beta_i I_{\{r_t(\theta_0,\beta_i)>\frac{1}{\alpha n}\}}\;&+\;(\beta_i-1)I_{\{r_t(\theta_0,\beta_i)<-\frac{1}{\alpha n}\}}\Big]\\&=\;\beta_i\textstyle\int_{\eta_{\beta_i}+\frac{1}{\alpha n}}^{\infty}dG(\lambda)+(\beta_i-1)\int_{-\infty}^{\eta_{\beta_i}-\frac{1}{\alpha n}}dG(\lambda)\end{aligned}$$

는 0으로 확률수렴한다. 따라서 식 (4.3)은 0으로 확률수렴한다. 그러므로 $\sqrt{n}(\hat{\theta}_n(\beta_i)-\theta_0)$와 $\sqrt{n}(\hat{\theta}_n(\beta_j)-\theta_0))$ 의 공분산은 $g^{-1}(\eta_{\beta_i})g^{-1}(\eta_{\beta_j})V^{-1}(\theta_0)\beta_i(1-\beta_j)$ 으로 확률수렴 한다. 위 식을 이용하여 다음 결과를 유도할 수 있다.

$$Var\left\{\sum_{i=1}^{m}C_i\sqrt{n}(\breve{\theta}_n(\beta_i)-\theta_0)\right\}$$
$$=\underline{D}\begin{bmatrix}g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_1})\beta_1(1-\beta_1)&\cdots&g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_m})\beta_1(1-\beta_m)\\\vdots&\ddots&\vdots\\g^{-1}(\eta_{\beta_1})g^{-1}(\eta_{\beta_m})\beta_1(1-\beta_m)&\cdots&g^{-1}(\eta_{\beta_m})g^{-1}(\eta_{\beta_m})\beta_m(1-\beta_m)\end{bmatrix}$$
$$\otimes V^{-1}(\theta_0)\underline{D}^T$$

따라서 Cramér-Wald의 법칙으로 부터

$$\sqrt{n}(\hat{\theta}_n(\beta_1) - \theta_0, \cdots, \hat{\theta}_n(\beta_M) - \theta_0) \longrightarrow N(0, \Omega \otimes V^{-1}(\theta_0))$$

임을 알 수 있다.

정리 2.3의 증명 증명을 간단히 하기 위하여 다음과 같은 기호를 정의하자.

$$\pi^* = [\pi_k^*](k = 1, \cdots, m), \pi_k^* = [\varphi_{ij}]_{p \times p}$$

여기서

$$\varphi_{ij} = \begin{cases} 0 & i \neq j \\ \pi(\beta_k) & i = j. \end{cases}$$

이고 π^* 는 $p \times pm$행렬이다. 그러면

$$\begin{aligned} \sqrt{n}(C_n(\pi) - \theta_0) &= \sqrt{n}\left(\sum_{i=1}^m \pi(\beta_i)\hat{\theta}_n(\beta_i) - \theta_0\right) \\ &= \sum_{i=1}^m \pi(\beta_i)\sqrt{n}(\hat{\theta}_n(\beta_i) - \theta_0) \\ &= \pi^* \begin{pmatrix} \sqrt{n}(\hat{\theta}_n(\beta_1) - \theta_0) \\ \vdots \\ \sqrt{n}(\hat{\theta}_n(\beta_m) - \theta_0) \end{pmatrix} \end{aligned}$$

이므로 정리 2.2의 증명과 비슷한 방법으로

$$\sqrt{n}(Q_n(\pi) - \theta_0) \xrightarrow{L} N\left(0, \pi^*\Omega \otimes V^{-1}(\theta_0)\pi^{*\prime}\right)$$

이 성립함을 증명할수 있다.

한편, $\frac{\beta_i(1-\beta_j)}{g(\eta_{\beta_i})g(\eta_{\beta_j})} = \sigma_{ij}$ 라 하면

$$\pi^*\Omega \otimes V^{-1}(\theta_0) = \left[\sum_{i=1}^m \pi(\beta_i)\sigma_{i1}V^{-1}(\theta_0), \cdots, \sum_{i=1}^m \pi(\beta_i)\sigma_{im}V^{-1}(\theta_0)\right]$$

이다. 따라서 우리는 다음과 같은

$$\begin{aligned} \pi^*\Omega \otimes V^{-1}(\theta_0)\pi^{*\prime} &= \left[\sum_{k=1}^m \sum_{i=1}^m \pi(\beta_k)\pi(\beta_i)\sigma_{ik}\right] V^{-1}(\theta_0) \\ &= \pi'\Omega\pi V^{-1}(\theta_0) \end{aligned}$$

결과를 얻는다.

참고문헌

1. Bassett, G., and Koenker, R., (1986). Strong consistency of regression quantile and related emprical process, *Econometric Theory* , 2, 633-643.

2. Choi S. H., Kim H. K., and Park, K. O., (1999). Nolinear regression quantiles estimation, *Journal of the Korean Statistical Society* , (to appear).

3. Jennrich, R., (1969). Asymptotic properties of nonlinear least squares estimations, *The Annal of Mathematical Statistics* , 40, 633-643.

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

5. Koenker,R., and Bassett, G., (1978). Regression Quantiles, *Econometrica* , 46, 33-50.

6. Oberhofer, W., (1982). The consistency of nonlinear regression minimizing the L_1-norm, *The Annals of Statistics* , 10, 316-319.

7. Portnoy, S., (1991). Asymptotic behavior of regression quantiles in non-stationary dependent cases, *Journal of Multivariate Analysis* , 50, 100-113.

8. Powell, J. L., (1986). Censored regression quantiles, *Journal of Econometrics* , 32, 143-155.

9. Wang, J., (1995). Asymptotic normality of L_1-estimators in nonlinear regression, *Journal of Multivariate Analysis* , 54, 227-238.

10. Wu, C. F., (1981). Asymptotic theory of nonlinear least squares estimation, *The Annals of Statistics* , 501-513.

Asymptotic Properties of Regression Quanties Estimators in Nonlinear Models

Seunghoe Choi [4] · Taesoo Kim [5] · Kyungok, Park [6]

Abstract

In this paper, we consider the Regression Quantiles Estimators in nonlinear regression models. This paper provides the sufficient conditions for strong consistency and asymptotic normality of proposed estimation and drives asymptotic relative efficiency of proposed estimatiors with least square estimation. We give some examples and results of Monte Carlo simulation to compare least square and regression quantile estimators.

Key Words and Phrases: Nonlinear Regression Model, Regression Quantiles Estimators, Consistency, Normality, Efficiency.

[4]Assistant Professor, Department of General Studies, Hankuk Aviation University, Koyang, Kyungkido, 412-791, Korea

[5]Research Professor, Department of Industrial Engineering, Hangyang University, Ansan, Kyungkido, 425-791, Korea

[6]Senior Engineer, Disital Printing Dvision, Samsung Electronics Co. Ltd, Suwon, Kyungkido, 442-742, Korea