한국시뮬레이션학회 '00추계학술대회 논문집 2000.11.11 아주대학교

비선형 회귀모형 추정량들의 몬테칼로 시뮬레이션에 의한 비교

김 태수*·이 영해**

Monte Carlo simulation of the estimators for nonlinear regression model

Tae Soo Kim*, Young Hae Lee**

Key Words: regression estimators, consistency, asymptotic normality, Monte Carlo simulation

Abstract

In regression model we estimate the unknown parameters using various methods. There are the least squares method which is the most general, the least absolute deviation, the regression quantile and the asymmetric least squares method. In this paper, we will compare each others with two case: to begin with the theoretical comparison in the asymptotic sense, and then the practical comparison using Monte Carlo simulation for a small sample size.

1. Introduction

Generally the nonlinear regression model is

$$y_t = f(x_t, \theta_o) + \varepsilon_t, \qquad (1.1)$$

where $t = 1, 2, \cdots, T$ and $f(x_t, \theta_o)$ is a real valued nonlinear function defined on $R^{p_1 + p_2}$, x_t is a $(1 \times p_2)$ observed vector, the error terms ε_t are independent and identically distributed(i.i.d.) with finite variance. The parameter vector θ_o which is interior point in a compact parameter space $\Theta \subset R^{p_1}$ is unknown and to be estimated.

We consider the objective function which defined by the following formula.

$$S_T(\theta; \alpha, \beta) = \frac{1}{T} \sum_{t=1}^{T} \varphi_\beta^\alpha(y_t - f(x_t, \theta)), \qquad (1.2)$$

where $\alpha = 1$ or 2, $0 < \beta < 1$, and $\varphi_\beta^\alpha(\lambda)$ is called a check function which is defined

$$\varphi_\beta^\alpha(\lambda) = \begin{cases} \beta\lambda^\alpha, & \lambda \geq 0, \\ (1-\beta)|\lambda|^\alpha, & \lambda < 0. \end{cases}$$

And now let the minimizer $\widehat{\theta_T}(\alpha, \beta)$ of the objective function $S_T(\theta; \alpha, \beta)$ such that

$$\widehat{\theta_T}(\alpha, \beta) = \underset{\theta \in \Theta}{\arg\min}\ S_T(\theta; \alpha, \beta).$$

Then the $\widehat{\theta_T}(\alpha, \beta)$ is called as follows,

for $\beta = 0.5$, $\alpha = 2$, the least squares estimator (LSE),

for $\beta = 0.5$, $\alpha = 1$, the least absolute deviation estimator (LAD),

for $\beta \neq 0.5$, $\alpha = 2$, the asymmetric least squares estimator (ALS),

---

* 한양대학교 산업공학과 연구교수
** 한양대학교 산업공학과 교수

for $\beta \neq 0.5$, $\alpha = 1$, the regression quantile estimator (RQE).

Jennrich(1969) first rigorously proved the existence of the LSE and showed the strong consistency and asymptotic normality of the LSE under the several assumptions. Wu(1981) gave some conditions such that as a Lipschitz type condition on the sequence $f(x_t, \theta)$ to prove the asymptotic properties of LSE. And Oberhofer(1982) studied the weak consistency about the LAD. Kim H.K(1995) proved the asymptotic properties of the LAD. Koenker, R. and Bassett, G.(1978) discussed the RQE. And Whitney K. Newey and James L. Powell(1987) studied the ALS. On the other hand, the concept of the periodicity in time series is of fundamental interest, since it provides a means for formalizing the notions of dependence or correlation between adjacent points. In this paper we think about a sum of sinusoidal components:

$$f(x_t, \theta_o) = \sum_{r=1}^{q} \{A_{ro}\cos(\omega_{ro}t) + B_{ro}\sin(\omega_{ro}t)\},$$

where $\theta_o = (A_{1o}, B_{1o}, \omega_{1o}, \cdots, A_{qo}, B_{qo}, \omega_{qo})$, for $q \geq 1$, A_{ro}, B_{ro}'s are some fixed unknown constants, ω_{ro} is unknown frequency lying between 0 to π $(1 \leq r \leq q)$ and in this case the observed value x_t means t. But the above formula does not satisfy Jennrich(1969)'s assumption nor Wu's Lipschitz type condition, the method which are proposed by Jennrich(1969) and Wu(1981) is not available. So, Walker(1971) obtained the asymptotic properties of an approximate LSE. Kundu(1993) and Kundu and Mitra(1996) gave the direct proof of the strong consistency and asymptotic normality and observed that the approximate LSE and the LSE are asymptotically equal. Also the asymptotic properties of LAD is proved by Kim T.S., Kim H.K. and Choi S.H.(2000). Kim T.S. and Kim H.G(2000) and Kim T.S., Kim H.G. and Hur S.(2000) proved the asymptotic properties of RQE. The asymptotic normality of ALS is discussed by Jang(1999).

The results are discussed in section 2. And then, for the asymptotic state which means the large sample size, we are able to check the relative efficiency of the four estimators. It is given in section 3. But they are adjusted in the theoretical sense. But in the practical phenomenon, we deals with the finite data. So, we will check the prpperties of the four estimates and efficiency by the Monte Carlo simulation in section 4.

2. The Asymptotic Properties

2.1 The Least Squares Estimators

Theorem 2.1 : Let

$$y_t = \sum_{r=1}^{q} \{A_{ro}\cos(\omega_{ro}t) + B_{ro}\sin(\omega_{ro}t)\} + \varepsilon_t,$$

where $0 < \omega_{ro} < \pi$, for $r = 1, 2, \cdots, q$, and ε_t are distributed independently and identically with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = \sigma^2 < \infty$. And assume that $\lim_{T \to \infty} \min_{1 \leq r \neq s \leq q} (T \mid \omega_{ro} - \omega_{so} \mid) = \infty$. Then we have the LSE is a strongly consistent estimator of θ_o and $P(\widehat{\theta_T}) = (P_1(\widehat{\theta_{1T}}), P_2(\widehat{\theta_{2T}}), \cdots, P_q(\widehat{\theta_{qT}})$, where for $r = 1, 2, \cdots, q$, $P_r(\widehat{\theta_{rT}}) = (\sqrt{T}(\widehat{A_{rT}} - A_{ro}), \sqrt{T}(\widehat{B_{rT}} - B_{ro}), \sqrt{T^3}(\widehat{\omega_{rT}} - \omega_{ro}))$ converges in law $N(0_{3q \times 1}, \sigma^2 \Sigma^{-1})$, where $\Sigma = (\Sigma_{rs})_{3q \times 3q}$, for $r, s = 1, 2, \cdots, 0$, and

$$\Sigma_{rs} = \begin{cases} 0, & \text{if } r \neq s \\ \begin{pmatrix} \frac{1}{2} & 0 & \frac{B_{ro}}{4} \\ 0 & \frac{1}{2} & \frac{-A_{ro}}{4} \\ \frac{B_{ro}}{4} & \frac{-A_{ro}}{4} & \frac{A_{ro}^2 + B_{ro}^2}{6} \end{pmatrix}, & \text{if } r = s. \end{cases}$$

2.2 The Least Absolute Deviation Estimators

Theorem 2.2 : Under the same conditions of Theorem 2.1, the LAD is a strongly consistent estimator of θ_o and $P(\widehat{\theta_T})$ converges in law

$N(0_{3q\times 1}, \frac{1}{\{2g(0)\}^2}\Sigma^{-1})$. where $g(\varepsilon_t)$ is a continuous probability density function of ε_t.

2.3. The Regression Quantile Estimators

Theorem 2.3 : Under the same conditions of Theorem2.1 except $E(\varepsilon_t)=0$ and with the additional condition $G(0)=\beta$, $(0<\beta<1, \beta\neq 0.5)$, where $G(\varepsilon_t)$ is a distribution function of the error term ε_t, The RQE is a strongly consistent estimator of θ_o and $P(\widehat{\theta_T})$ converges in law $N(0_{3q\times 1}, \frac{\beta(1-\beta)}{\{g(0)\}^2}\Sigma^{-1})$. where $g(\varepsilon_t)$ is a continuous probability density function of ε_t.

2.4 The Asymmetric Least Squares Estimators

Theorem 2.4 : Under the same conditions of Theorem2.3, the ALS is a strongly consistent estimator of θ_o and $P(\widehat{\theta_T})$ converges in law $N(0_{3q\times 1}, \frac{(1-2\tau)d+\tau^2(\sigma^2+\mu^2)}{[\beta+\tau(1-2\beta)]^2}\Sigma^{-1})$. where $g(\varepsilon_t)$ is a continuous probability density function of ε_t, and $\tau=\frac{b}{2b-\mu}$, $\mu=E(\varepsilon_t)$, $b=\int_{-\infty}^{0} xg(x)dx$, $d=\int_{-\infty}^{0} x^2 g(x)dx$.

3. The Asymptotic Relative Efficiency

Since the assumptions of the LSE are equal to the LAD, so the asymptotic efficiency of LSE relative to LAD estimators is

$$\lim_{T\to\infty} eff\,(\widehat{\theta_T}(2,1/2)|\widehat{\theta_T}(1,1/2))=\{2g(0)\}^2\sigma^2.$$

It does implies that the LSE is asymptotically more efficient than LAD in the sinusoidal model for the normal distribution, but for the heavy-tailed error distributions more than normal distribution likewise the Laplace distribution, t-distribution and logistic distribution, ets, the LAD is asymptotically more efficient than the LSE. On the other hand, the asymptotic efficiency of ALS relative to RQE

$$\lim_{T\to\infty} eff\,(\widehat{\theta_T}(2,\beta)|\widehat{\theta_T}(1,\beta))\ ,\ \text{where}\ \ \beta\neq 0.5,$$

is $\frac{g^2(0)[(1-2\tau)d+\tau^2(\sigma^2+\mu^2)]}{\beta(1-\beta)[\beta+\tau(1-2\beta)]^2}$.

4. Monte Carlo Simulation

We performed some Monte Carlo simulations

	T=10		T=15		T=25	
	LSE	LAD	LSE	LAD	LSE	LAD
0.25 π	.7843	.7810	.7852	.7852	.7842	.7833
	.0052	.0098	.0016	.0028	.0004	.0006
	.1352	.1395	.0699	.0715	.0361	.0371
	.901	.842	.932	.887	.932	.854
0.50 π	1.5693	1.5718	1.5700	1.5690	1.5705	1.5701
	.0052	.0099	.0016	.0028	.0003	.0006
	.1284	.1335	.0732	.0751	.0345	.0350
	.904	.863	.928	.883	.941	.882
.75 π	2.3555	2.3583	2.3566	2.3576	2.3563	2.3586
	.0051	.0096	.0016	.0028	.0003	.0006
	.1213	.1260	.0767	.0779	.0339	.0350
	.922	.877	.918	.859	.947	.888

Table3.1 The distribution of error is $N(0,1)$

	T=10		T=15		T=25	
	LSE	LAD	LSE	LAD	LSE	LAD
0.25 π	.7535	.7674	.7576	.7744	.7741	.7742
	.1041	.0550	.0293	.0146	.0062	.0032
	.5535	.5773	.3060	.3057	.1468	.1495
	.881	.890	.853	.907	.849	.902
0.50 π	1.5346	1.5732	1.5505	1.5593	1.5554	1.5563
	.1018	.0549	.0289	.0150	.0063	.0032
	.5469	.5714	.3051	.3115	.1478	.1495
	.848	.925	.879	.912	.859	.883
0.75 π	2.3507	2.3494	2.3322	2.3499	2.3483	2.3519
	.1047	.0522	.0296	.0146	.0063	.0031
	.5507	.5550	.3080	.3087	.1465	.1455
	.847	.910	.872	.922	.897	.937

Table3.2 The Distribution of error is the Laplace with a parameter $\phi=3$

to compare the four different estimators. We will check the behaviour of the estimators for small sample sizes since they are asymptotically founded the exact relations. Numerical results are reported for $T=10, 15, 25$ and $\omega=0.25\pi(\approx 0.785398)$, $0.5\pi(\approx 1.570796)$, $0.75\pi(\approx 2.356194)$. For a particular T and ω, 1000's different sets of data were generated. The two parameters A and B are taken as a 1.5 each. Under the given each data sets, we estimated the nonlinear parameter ω by the four methods, and described the average estimates, average mean squared error and the average length of 95% confidence interval over 1000 simulations. And the last figures in each tables is the coverage probability.

	T=10		T=15		T=25	
	ALS	RQE	ALS	RQE	ALS	RQE
0.25 π	.7844	.7854	.7854	.7819	.7857	.7857
	.0090	.0135	.0025	.0036	.0005	.0007
	.1583	.1631	.0790	.0815	.0411	.0414
	.933	.897	.938	.900	.950	.885
0.50 π	1.5694	1.5662	1.5712	1.5703	1.5708	1.5704
	.0091	.0135	.0025	.0037	.0005	.0008
	.1509	.1549	.0832	.0848	.0395	.0400
	.948	.905	.939	.893	.951	.892
0.75 π	2.3553	2.3551	2.3566	2.3551	2.3562	2.3563
	.0091	.0135	.0025	.0036	.0005	.0008
	.1447	.1503	.0857	.0870	.0384	.0390
	.946	.908	.931	.870	.960	.898

Table3.3 The distribution of error is $N(1,1)$

	T=10		T=15		T=25	
	ALS	RQE	ALS	RQE	ALS	RQE
0.25 π	.7699	.7792	.7666	.7718	.7117	.7801
	.1163	.1045	.0327	.0285	.0073	.0055
	.5494	.56619	.3045	.3097	.1519	.1513
	.860	.893	.784	.916	.566	.915
0.50 π	1.5748	1.5672	1.5580	1.5675	1.5315	1.5636
	.1216	.1071	.0335	.0286	.0071	.0055
	.5637	.5745	.3115	.3155	.1499	.1487
	.841	.896	.803	.920	.601	.938
0.75 π	2.3621	2.3557	2.3432	2.3531	2.3324	2.3557
	.1190	.1046	.0329	.0294	.0071	.0058
	.5528	.5706	.3084	.3151	.1491	.1483
	.812	.892	.836	.928	.633	.934

Table3.4 The distribution of error is the Laplace with a parameter $\phi=3$ and $\beta=0.4$

5.Conclusions

Likewise the asymptotic cases, we have a same results under the small sample size for the most part. So in nonlinear regression analysis, do not depend only the least squares methods, but we should determine the use of the estimating method among the four different ways after the check of the distribution of the error for the perfect data analysis.

References

[1] Debasis Kundu(1993), Asymptotic theory of least squares estimator of a particular nonlinear regression model, Statistical & Probability Letters, 18, 13-17.

[2] Debasis Kundu and Amit Mitra.(1996), Asymptotic theory of least squares estimator of a nonlinear time series regression model, Commun. Statist.-Theory Meth. 23(1), 133-141.

[3] Jang S.E. (1999), Nonlinear Regression Analysis for Asymmetric Error Distribution, Yonsei University.

[4] Jennrich, R.(1969), Asymptotic properties of nonlinear least squares estimations, Ann. Math.

Stat. 40, 633-643.

[5] Kim T.S. and Kim H.K. (2000), Consistency of RQE for the nonlinear time series regression model, Bull. Korean Math, submitted.

[6] Kim T.S., Kim H.K and Hur S.(2000), Asymptotic properties of nonlinear regression quantile estimation, Statistics and probability letters, submitted.

[7] Kim T.S., Kim H.K. and Choi S.H., (2000), Asymptotic Properties of LAD Estimators of a Nonlinear Time Series Regression Model, Journal of the korean statistical society, vol29, no2, 187-199.

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

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

[10] Walker, A.M.(1971), On the estimation of a harmonic component in a time series with stationary independent residuals, Biometrika, 58, 21-36.

[11] White, H.(1981), Nonlinear regression on cross-section data, Econometrica 48, 721-746.

[12] Whitney K. Newey and James L. Powell(1987), Asymmetric least equares estimation and testing, Econometrica, Vol 53, No 4, 819-847.

[13] Wu, C. F.(1981), Asymptotic theory of nonlinear least squares estimation, The Annals of Statistics 9, 501-513.