응용통계연구 제 6권 2호
1993년 9월 pp.265-275.

Consistency of M-estimators in Nonlinear Regression Model

Kyeong Hee Kim[1] In Hwan Chung[2]

Abstract

This paper deals with the M-estimators in regression model. The class of M-estimators is defined on nonlinear regression model and the conditions to hold the consistency of the considered estimators are suggested when the parameter space of the model is compact.

1. Introduction

In the location problem, Huber (1964) introduced the class of robust estimators which are called M-estimators. This procedure is finding the solutions $\hat{\theta}$ of equations of the form

$$\sum_{i=1}^{n} \psi(X_i - \hat{\theta}) = 0, \tag{1.1}$$

where $X_1 = \theta + \varepsilon_1, \cdots, X_n = \theta + \varepsilon_n$ and $\varepsilon_1, \cdots, \varepsilon_n$ are unknown independent, identically distributed errors which have a distribution function F which is symmetric about 0. If F has a density f which is smooth and if f is known, then the maximum likelihood estimators are obtained by taking $\psi = -f'/f$. Under successively milder regular conditions on ψ and F, Huber (1964) showed that $\hat{\theta}$ satisfying (1.1) were consistent and asymptotically normal.

In this paper we consider M-estimators for nonlinear regression models. Details of the model and the estimators are to be found in Section 2. Statements and proofs of the consistency of the M-estimators appear in Section 3. Finally some remarks on scale are given in Section 4.

2. Model and Estimators

The class of M-estimators can be extended to the regression model. We consider the following nonlinear regression model,

$$y_j = g(x_j, \theta) + \varepsilon_j, \quad j = 1, \cdots, n, \tag{2.1}$$

where $x_j \in \Xi \subset R^m$ denotes the j-th fixed known input vector, $\theta \in R^p$ is the parameter vector from the parameter space $\Theta \subset R^p$ and $g: \Xi \times \Theta \to R^1$ is a known measurable function on Ξ for each $\theta \in \Theta$. The random errors $\varepsilon_1, \cdots, \varepsilon_n$ are independent identically distributed

1) Department of Mathematics, Yonsei University, Shinchondong, Seoul 120-749
2) Department of Mathematics, Yonsei University, Shinchondong, Seoul 120-749

(i.i.d.) random variables which have a distribution function F. We shall write $g(x_j,\theta)$ by $g_j(\theta)$.

The problem of interest is making inference about θ in some optimal way, on the basis of observations on y_j and x_j, $j=1,\cdots,n$.

Let

$$Q_n(\theta)=\sum_{j=1}^{n}\rho(y_j-g_j(\theta))$$

and consider the problem of finding θ which minimizes $Q_n(\theta)$ when ρ is differentiable. If $\psi=\rho'$, the solution minimizing $Q_n(\theta)$ must satisfy the equations

$$\sum_{j=1}^{n}\psi(y_j-g_j(\theta))\frac{\partial}{\partial\theta_i}g_j(\theta)=0,\ \ i=1,\cdots,p.$$

If ρ is convex, the two approaches are equivalent.

Let

$$R_j(t)=y_j-g_j(t)\ \text{ if }\ t=(t_1,\cdots,t_p).$$

Then the errors are given by

$$R_j(\theta)=y_j-g_j(\theta),\ j=1,\cdots,n,$$

if θ is true.

An M-estimator is defined quite naturally as a solution $\hat\theta=(\hat\theta_1,\cdots,\hat\theta_p)$ of the system of equations

$$\sum_{j=1}^{n}\left(\frac{\partial}{\partial\theta_i}g_j(\hat\theta)\right)\psi(R_j(\hat\theta))=0,\ \ i=1,\cdots,p. \tag{2.2}$$

Again if $\psi=-f'/f$ where f is a density of F, these are the likelihood equations, and if $\psi(t)=t$, $\hat\theta$ is the least squares estimator. Many authors have provided conditions which insure the existence, consistency and asymptotic normality of the nonlinear least squares estimator. Jennrich (1969), Malinvaud (1970) and Wu (1981) proved asymptotic normality or consistency of the nonlinear least squares estimator when the errors are i.i.d. random variables. When $\rho(u)=|u|$, the consistency of $\hat\theta$ minimizing $Q_n(\theta)$ which is called the minimum L_1-norm estimator was given by Oberhofer (1982).

We shall give some properties of M-estimators. This clearly requires some conditions on the model (2.1). We define, following Jennrich (1969), a tail product.

Definition 2.1. We shall say that the sequence $\{h_j\}_{j=1}^{\infty}$ where

$$h_j=[h_{j1},\cdots,h_{jk}]:\Theta\rightarrow R^k$$

has a finite tail product if

$$\frac{1}{n}\sum_{j=1}^{n}h_j(\alpha)^T h_j(\beta)$$

converges uniformly in $(\alpha,\beta)\in\Theta\times\Theta$ as $n\to\infty$ (T denotes transpose). In this case, the limit is called the tail product of $\{h_j\}_{j=1}^{\infty}$.

We use the notation

$$Dg_j(\theta)=[\frac{\partial}{\partial\theta_1}g_j(\theta),\cdots,\frac{\partial}{\partial\theta_p}g_j(\theta)],$$

which will be used to give the condition on the model (2.1).

Remark 2.1.

In the general linear model we observe $y=(y_1,\cdots,y_n)$ where

$$y_j=\sum_{i=1}^{p}x_{ij}\beta_i+\varepsilon_j,\quad 1\le j\le n,$$

and ε_j are random errors having distribution function F, the β_i unknown regression parameters and $X=\{x_{ij}\}$, the design matrix.

The classical estimator of $\beta=(\beta_1,\cdots,\beta_p)$ is the least squares estimator, defined as the vector $\hat{\beta}$ which minimizes $\|y-\beta X\|$ (here $\|\cdot\|$ denotes the Euclidean norm).

When F has heavy tails, $\hat{\beta}$ is not efficient and may not even be consistent. Hence, other ''robust'' estimators of β have been proposed. In particular Relles (1968), Huber (1973) and Yohai (1972) have studied the family of Huber M-estimators, defined as solutions of

$$\sum_{j=1}^{n}\rho(y_j-\sum_{i=1}^{p}x_{ij}\beta_i)=minimum,\tag{2.3}$$

where the function ρ is chosen conveniently.

If ρ is convex with derivative ψ, then (2.3) is equivalent to

$$\sum_{j=1}^{n}x_{ij}\psi(y_j-\sum_{i=1}^{p}x_{ij}\beta_i)=0,\quad 1\le i\le p.\tag{2.4}$$

The consistency of these estimators was proved by Relles (1968) when ψ belongs to the Huber family

$$\psi(x,k)=\min(|x|,k)sgn(x),$$

and Yohai (1972) proved the consistency for ψ monotone. When the number of parameters goes to infinity, Portnoy (1984) suggested the conditions for the consistency.

The asymptotic normality was proved by Huber (1973), Yohai and Maronna (1979) when ψ is monotone.

3. Consistency

We shall prove the consistency of the M-estimators under the following conditions.

In the definitions and arguments which follow we shall assume that all probabilities and expectations are calculated under assumption that $\theta_0=(\theta_{01},\cdots,\theta_{0p})$ is the true parameter unless the contrary is specifically indicated.

Condition A.

A1. The parameter space Θ is a compact subset of R^p and θ_0 is an interior point of Θ.

A2. $\{Dg_j(\theta)\}_{j=1}^{\infty}$ has a finite tail product and

$$\Sigma(\theta_0)=\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}(Dg_j(\theta_0))^T(Dg_j(\theta_0))$$

is positive definite. We shall denote $\Sigma(\theta_0)$ by Σ.

A3. If t_n is a sequence in R^p such that $\|t_n\|=O(1/\sqrt{n})$, then

$$\lim_{n\to\infty}\max_{1\le j\le n}|g_j(t_n+\theta_0)-g_j(\theta_0)|=0.$$

We shall also need some conditions on ψ.

Condition B.

B1. ψ is nondecreasing. And there exists a positive number d such that

$$\psi'(x)\ge d$$

for all $x\in S(F)$ where $S(F)$ is the support of F.

B2. $E_F(\psi^2(u))<\infty$.

Remark 3.1.

The conditions A1 and A2 were also used for the asymptotic normality of the nonlinear least squares estimator. In particular, the compactness of the parameter space guarantees the existence of the estimator $\hat{\theta}$ which minimizes $Q_n(\theta)$ (see Lemma 2 in Jennrich (1969)). A3 is a stronger condition than the continuity of the regression function $g_j(\theta)$ since

$$\lim_{n\to\infty}|g_j(t_n+\theta_0)-g_j(\theta_0)|=0,$$

where $g_j(\theta)$ is continuous and $\|t_n\|=O(1/\sqrt{n})$.

Now we suggest some examples related to the conditions above.

Example 3.1. Consider the exponential model with the regression function $g_j(\theta)=\theta_1 e^{\theta_2 x_j}$ where $\theta=(\theta_1,\theta_2)$ ranges over the unit rectangle $\Theta=[0,1]\times[0,1]$ and $x_1,x_2,\cdots$ is a bounded sequence of real numbers whose sample distribution function is G_n converging to a distribution G weakly. Assume that the true parameter $\theta_0=(\alpha_1,\alpha_2)$ is an interior point of Θ and that G is not degenerate. We have

$$Dg_j(\theta)=[e^{\theta_2 x_j},x_j\theta_1 e^{\theta_2 x_j}].$$

Since Θ is compact, $\frac{1}{n}\sum_{j=1}^{n}(Dg_j(\theta))^T(Dg_j(\theta))$ converges uniformly (see Theorem 1 in

Jennrich (1969)) to

$$\Sigma=\begin{pmatrix} \int e^{2\theta x}dG(x) & \int x\theta_1 e^{2\theta x}dG(x) \\ \int x\theta_1 e^{2\theta x}dG(x) & \int x^2\theta_1^2 e^{2\theta x}dG(x) \end{pmatrix}$$

It remains to show that $\Sigma(\theta_0)$ is positive definite. For any $\beta=(\beta_1,\beta_2)\in R^2$,

$$\beta\Sigma(\theta_0)\beta^T=\int(\beta_1+\beta_2\alpha_1x)^2e^{2\alpha x}dG(x).$$

Then $\beta\Sigma(\theta_0)\beta^T=0$ only if $\beta_1+\beta_2\alpha_1x=0$ on a set of x with probability one. Since $\alpha_1\neq 0$ and G is not degenerate this can happen only when $\beta_1=\beta_2=0$. Thus $\Sigma(\theta_0)$ is positive definite. And hence the conditions A1 and A2 are satisfied. Finally the condition A3 holds clearly since $\{x_j\}$ is a bounded sequence of real numbers.

Example 3.2. If $\hat{\theta}$ satisfying (2.2) is the least squares estimator then $\psi(t)=t$. Thus the condition B1 holds for all distribution functions F and the condition B2 simply states the finite condition of the second moment of F.

Lemma 3.1. Let B be a $p\times p$ positive definite matrix with eigenvalues $\lambda_1\geq\cdots\geq\lambda_p>0$ and associated normalized eigenvectors $e_1,\cdots,e_p$. Then

$$\min_{x\neq 0}\frac{xBx^T}{xx^T}=\lambda_p$$

attained when $x=e_p$.

Proof See Rao (1973).

Lemma 3.2. Suppose the conditions A2 and B2 are satisfied, let

$$A_n=\sup_{\|\theta\|=1}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi(\varepsilon_j)Dg_j\left(\frac{L}{\sqrt{n}}\theta+\theta_0\right)\theta^T\right|,$$

then for any $\delta>0$, there exists $L_1>0$ (which does not depend on n) such that

$$P[A_n\geq L_1]\leq\delta.$$

Proof Observe that

$$\begin{aligned}&\sup_{\|\theta\|=1}\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi(\varepsilon_j)Dg_j\left(\frac{L}{\sqrt{n}}\theta+\theta_0\right)\theta^T\right|^2\\ \leq\ &\sup_{\|\theta\|=1}\frac{1}{n}\sum_{j=1}^{n}\left\|\psi(\varepsilon_j)Dg_j\left(\frac{L}{\sqrt{n}}\theta+\theta_0\right)\right\|^2\|\theta^T\|^2\end{aligned}\tag{3.1}$$

by the Schwarz inequality. The last term in (3.1) is

$$\sup_{\|\theta\|=1}\frac{1}{n}\sum_{j=1}^{n}\|\psi(\varepsilon_j)Dg_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\|^2$$
$$=\sup_{\|\theta\|=1}\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{n}\psi^2(\varepsilon_j)\left\{\frac{\partial}{\partial\theta_i}g_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\right\}^2.$$

So

$$E[A_n]^2\le E[\psi^2(\varepsilon_1)]v_n,$$

where

$$v_n=\sup_{\|\theta\|=1}\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{n}\{\frac{\partial}{\partial\theta_i}g_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\}^2.$$

By A2, there exists a $v<\infty$ such that

$$v_n<v,$$

for all n. By the Chebyshev's inequality,

$$P[A_n\ge L_1]\le E[A_n^2]/L_1^2.$$

Hence

$$P[A_n\ge L_1]\le E[\psi^2(\varepsilon_1)]v/L_1^2\ .$$

By the condition B2 and the appropriate choice of L_1, the result follows.

Define, for $\theta\in R^p$ and $L>0$,

$$U(\theta,L)=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi(y_j-g_j(L\theta+\theta_0))Dg_j(L\theta+\theta_0)\theta^T.$$

Lemma 3.3. Suppose that $U(\theta,L)$ is monotone decreasing in $L>0$, then we have

$$P[\sqrt{n}\|\hat\theta-\theta_0\|\ge L]\le P[\sup_{\|\theta\|=1}U(\theta,L/\sqrt{n})\ge 0].$$

Proof Observe that

$$U(\frac{\hat\theta-\theta_0}{\|\hat\theta-\theta_0\|},\|\hat\theta-\theta_0\|)$$
$$=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi[y_j-g_j(\hat\theta)]Dg_j(\hat\theta)(\hat\theta-\theta_0)^T/\|\hat\theta-\theta_0\|$$
$$=0.$$

Thus we have

$$\sup_{\|\theta\|=1}U(\theta,\|\hat\theta-\theta_0\|)\ge 0.$$

Since $U(\theta,L)$ is monotone decreasing in $L>0$, $\|\hat\theta-\theta_0\|\ge L/\sqrt{n}$ implies that

$$\sup_{\|\theta\|=1}U(\theta,L/\sqrt{n})\ge 0\ .$$

So the Lemma follows.

Theorem 3.4. Under the conditions A and B suppose that $U(\theta,L)$ is monotone decreasing in $L>0$. And let $\hat{\theta}$ satisfy (2.2), then $\hat{\theta}$ is $\sqrt{n}$-consistent for θ_0, that is,

$$\sqrt{n}(\hat{\theta}-\theta_0)=O_p(1).$$

Proof By the Lemma 3.3, it suffices to prove that, given $\delta>0$, there exists L such that

$$\limsup_{n\to\infty} P[\sup_{\|\theta\|=1} U(\theta, L/\sqrt{n}) \geq 0] < \delta.$$

Now fix $L>0$ and decompose

$$\begin{aligned} U(\theta,L/\sqrt{n}) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi(\varepsilon_j)Dg_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\theta^T \\ &-\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\psi(\varepsilon_j)-\psi[y_j-g_j(\frac{L}{\sqrt{n}}\theta+\theta_0)]\}Dg_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\theta^T. \end{aligned}$$

Put

$$A_n=\sup_{\|\theta\|=1}|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\psi(\varepsilon_j)Dg_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\theta^T|,$$

and

$$B_n=\inf_{\|\theta\|=1}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\{\psi(\varepsilon_j)-\psi[y_j-g_j(\frac{L}{\sqrt{n}}\theta+\theta_0)]\}Dg_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\theta^T.$$

By Lemma 3.2, there exists $L>0$ such that

$$P[A_n\geq Ld\lambda_p]\leq\delta/2, \tag{3.2}$$

where d is the positive number in the condition B1 and λ_p is the smallest eigenvalue of Σ.

Note that

$$B_n=\inf_{\|\theta\|=1}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{\psi(\varepsilon_j)-\psi[y_j-g_j(\frac{L}{\sqrt{n}}\theta+\theta_0)]}{z_j(n)}Dg_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\theta^T z_j(n),$$

where

$$z_j(n)=g_j(\frac{L\theta}{\sqrt{n}}+\theta_0)-g_j(\theta_0).$$

Then, by A3 and B1, there exists n_1 such that

$$B_n\geq\inf_{\|\theta\|=1}\frac{d}{\sqrt{n}}\sum_{j=1}^{n}Dg_j(\frac{L\theta}{\sqrt{n}}+\theta_0)\theta^T z_j(n), \tag{3.3}$$

for all $n\geq n_1$ with probability 1.

By A_1, there exists a neighborhood $N_1\subset\Theta$ centered at θ_0, and there exists n_2 such that

$$\frac{L}{\sqrt{n}}\theta+\theta_0\in N_1,$$

for all $n\geq n_2$. Thus, by the multivariate version of the Mean Value Theorem,

$$z_j(n)=\sum_{i=1}^{p}\Big(\frac{\partial}{\partial\theta_i}g_j(\theta^L)\Big)\frac{L}{\sqrt{n}}\theta_i,\quad \text{if } n\ge n_2,$$

where θ^L lies in the interior of the line segment joining θ_0 and $\frac{L}{\sqrt{n}}\theta+\theta_0$ which lies in N_1. Besides, we have

$$Dg_j(\frac{L\theta}{\sqrt{n}}+\theta_0)\theta^T=\sum_{k=1}^{p}\Big(\frac{\partial}{\partial\theta_k}g_j(\frac{L\theta}{\sqrt{n}}+\theta_0)\Big)\theta_k,$$

hence

$$\frac{d}{\sqrt{n}}\sum_{j=1}^{n}Dg_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\theta^T z_j(n)$$
$$=\frac{Ld}{n}\sum_{j=1}^{n}\sum_{i=1}^{p}\sum_{k=1}^{p}\Big(\frac{\partial}{\partial\theta_k}g_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\Big)\theta_k\Big(\frac{\partial}{\partial\theta_i}g_j(\theta^L)\Big)\theta_i\ ,$$

for all $n\ge n_2$. By (3.3)

$$B_n\ge Ld\inf_{\|\theta\|=1}\sum_{i=1}^{p}\sum_{k=1}^{p}\theta_i\Big(\frac{1}{n}\sum_{j=1}^{n}\Big(\frac{\partial}{\partial\theta_i}g_j(\theta^L)\Big)\Big(\frac{\partial}{\partial\theta_k}g_j(\frac{L}{\sqrt{n}}+\theta_0)\Big)\Big)\theta_k,$$

for all $n\ge\max(n_1,n_2)$ with probability 1.

By A2 and Lemma 3.1, there exists n_3 such that

$$\inf_{\|\theta\|=1}\sum_{i=1}^{p}\sum_{k=1}^{p}\theta_i\Big(\frac{1}{n}\sum_{j=1}^{n}\Big(\frac{\partial}{\partial\theta_i}g_j(\theta^L)\Big)\Big(\frac{\partial}{\partial\theta_k}g_j(\frac{L}{\sqrt{n}}\theta+\theta_0)\Big)\Big)\theta_k\ \ge\lambda_p$$

$$\text{if } n\ge n_3\ .$$

Letting $n_4=\max(n_1,n_2,n_3)$, then

$$B_n\ge Ld\lambda_p,\quad \text{if}\quad n\ge n_4, \tag{3.4}$$

with probability 1. Therefore, for all $n\ge n_4$, we have

$$P[\sup_{\|\theta\|=1}U(\theta,L/\sqrt{n})\ge 0]$$
$$\le P[A_n\ge Ld\lambda_p]+P[B_n<Ld\lambda_p]$$
$$<\delta,$$

by (3.2) and (3.4).
This finishes the proof.

Remark 3.2.

The monotonicity of $U(\theta,L)$ requires the conditions on both (2.1) and ψ, while Condition A is only concerned with (2.1) and Condition B with ψ. In Example 3.1, $U(\theta,L)$ is monotone decreasing in $L>0$ if $\{x_j\}$ is a bounded sequence of positive real numbers.

By Theorem 3.4, we get the following result immediately.

Corollary 3.5. Under the conditions of the Theorem 3.4, $\hat{\theta}$ satisfying (2.2) is a weakly consistent estimator of θ_0.

4. Estimation of scale

For the linear regression model, typically the estimators obtained from (2.4) are not scale equivariant, i.e., they do not satisfy

$$\hat{\theta}(ay_1, \cdots, ay_n) = a\hat{\theta}(y_1, \cdots, y_n), \quad a > 0.$$

If ρ is convex and $\rho' = \psi$, scale equivariant estimators may be obtained as solutions of

$$\sum_{j=1}^{n} x_{ij}\psi_{\hat{\sigma}}(y_j - \sum_{i=1}^{p} x_{ij}\beta_i) = 0, \quad 1 \le i \le p,$$

where $\psi_\sigma(x) = \psi(x/\sigma)$ and $\hat{\sigma}(y_1, \cdots, y_n)$ is a scale equivariant estimator of a scale parameter σ.

In nonlinear regression model, suppose that $\hat{\theta}$ and $\hat{\sigma}$ are obtained by solving the simultaneous equations

$$\sum_{j=1}^{n} \psi\left(\frac{R_j(\theta)}{\sigma}\right)\frac{\partial}{\partial \theta_i} g_j(\theta) = 0, \quad i = 1, \cdots, p \tag{4.1}$$

$$\sum_{j=1}^{n} \chi\left(\frac{R_j(\theta)}{\sigma}\right) = 0 \tag{4.2}$$

for θ and σ.

Example 4.1. Assume that the observations have a probability density of the form

$$\frac{1}{\sigma} f\left(\frac{R_j(\theta)}{\sigma}\right),$$

then (4.1) and (4.2) give the maximum likelihood estimators if

$$\psi(x) = -f'(x)/f(x),$$
$$\chi(x) = x\psi(x) - 1.$$

Example 4.2. If we take $\chi(x) = sgn(|x| - 1)$, median absolute residuals as the scale estimator are obtained by (4.2).

References

[1] Gallant, A. R. (1987), *Nonlinear statistical models,* John Wiley and Sons, New York

[2] Huber, P. J. (1964), "Robust estimation of a location parameter," *The Annals of Mathematical Statistics* **35**, 73-101

[3] Huber, P. J. (1973), "Robust regression: Asymptotics, conjectures and Monte Carlo," *Annals of Statistics* **1**, 799-821

[4] Jennrich, R. J. (1969), "Asymptotic properties of nonlinear least squares estimators," *The Annals of Mathematical Statistics* **40**, 633-643

[5] Malinvaud, E. (1970), "The consistency of nonlinear regression," *Annals of Mathematical Statistics* **41**, 956-969

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

[7] Rao, C. R. (1973), *Linear statistical inference and its applications,* John Wiley and Sons, New York

[8] Portnoy, S. (1984), "Asymptotic behavior of M-estimators of p regression parameters when p^2/n is large. I: Consistency," *Annals of Statistics* **12**, 1298-1309

[9] Wu, C. F. (1981), "Asymptotic theory of nonlinear least squares estimator," *Annals of Statistics* **9**, 501-513

[10] Yohai, V. J. (1974), "Robust estimation in the linear model," *Annals of Statistics* **2**, 562-567

[11] Yohai, V. J. and Maronna, R. A. (1979), "Asymptotic behavior of M-estimators for the linear model," *Annals of Statistics* **7**, 258-268

비선형 회귀모형에서 M-추정량의 일치성

김경희[1], . 정 인 환[2]

요 약

본 논문에서는 회귀모형의 M-추정량을 다루고 있다. 비선형 회귀모형에서 M-추정량을 정의하고 모형의 모수공간이 콤팩트일 때 일치성이 성립하기 위한 조건을 제시하였다.

1) (120-749) 서울특별시 서대문구 신촌동 134 연세대학교 수학과
2) (120-749) 서울특별시 서대문구 신촌동 134 연세대학교 수학과