# ASYMPTOTIC BEHAVIOR OF M-ESTIMATORS FOR THE LINEAR MODEL WITH DEPENDENT ERRORS*

by

B.L.S. Prakasa Rao
Purdue University and
Indian Statistical Institute, New Delhi

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #79-18

August, 1979

* Research partially supported by the National Science Foundation under Grant MCS77-19640 at Purdue University.

# ASYMPTOTIC BEHAVIOR OF M-ESTIMATORS FOR THE LINEAR MODEL WITH DEPENDENT ERRORS*

by

B.L.S. Prakasa Rao
Purdue University and
Indian Statistical Institute, New Delhi

## Abstract

Asymptotic properties of M-estimators for the linear model $\underset{\sim}{Y}_n = \underset{\sim}{X}_n \theta + \underset{\sim}{U}_n$, where $\underset{\sim}{U}_n = (u_1, .., u_n)$ and $\{u_i, i \geq 1\}$ form a stationary $\phi$-mixing process, are investigated.

---

AMS 1970 Subject Classification: Primary 62G35 Secondary 62J05,62M99

Key words and phrases: Robust estimation, linear model, consistency, asymptotic normality, mixing process

* Research partially supported by the National Science Foundation under Grant MCS77-19640 at Purdue University.

## 0. Introduction

Asymptotic theory of maximum likelihood type robust estimators or the so called M-estimators for the linear model has been studied by Huber (1972) and more recently by Yohai and Maronna (1979) under the assumption that the errors are independent and identically distributed. Huber (1972) says that "the assumption of independence is a serious restriction; the assumption that the errors are identically distributed simplifies notations and calculations but could easily be relaxed". Our aim in this paper is to extend the results of Yohai and Maronna (1979) on consistency and asumptotic normality of M-estimators of regression coefficients when the errors form a stationary $\phi$-mixing process. As can be expected, the results are not as sharp as they are in the independent case. Our results include the case when the number p of the parameters increases with the number n of observations.

Asymptotic properties of M-estimators for location parameter families were studied by Deniau, Oppenheim and Viano (1977) for mixing processes and asymptotic theory of M-estimators for Markov processes is investigated in Prakasa Rao (1972) generalizing the work of Huber (1967).

## 1. Preliminaries

Let $\{U_n, n \geq 0\}$ be a real-valued stationary process defined on a probability space $(\Omega, \mathcal{B}, P)$. Denote the $\sigma$-algebra generated by $U_i, k \leq i \leq m$ by $\mathcal{B}_k^m$. Let

$$\phi(n) = \sup[\text{ess sup}\{P(B|\mathcal{B}_0^k) - P(B)\} : B \in \mathcal{B}^{\infty}_{k+n}].$$

$\{U_n, n \geq 1\}$ is said to be $\phi$-mixing with mixing coefficient $\phi(n)$ if $\phi(n) \downarrow 0$ as $n \to \infty$.

Lemma 1.1. Suppose f is $\mathcal{B}_0^i$-measurable, g is $\mathcal{B}_j^\infty$-measurable and $E|f|^2 < \infty$ and $E|g|^2 < \infty$. Then

$$|Efg - Ef\,Eg| \leq 2\phi^{\frac{1}{2}}(|j-i|)\, E^{\frac{1}{2}}|f|^2\, E^{\frac{1}{2}}|g|^2.$$

Lemma 1.2. Suppose the random variables $f_i$ are $\mathcal{B}_0^i$-measurable for $1 \leq i \leq n$ and $E|f_i|^2 < \infty$ for $1 \leq i \leq n$. then

$$\left|\operatorname{Var} \sum_{i=1}^n f_i - \sum_{i=1}^n \operatorname{Var} f_i\right| \leq 4 \left(\sum_{j=1}^n \phi^{\frac{1}{2}}(j)\right) \sum_{i=1}^n \operatorname{Var} f_i.$$

In particular, if $M \equiv \sum_{j=1}^\infty \phi^{\frac{1}{2}}(j) < \infty$, then

$$\operatorname{Var}\left(\sum_{i=1}^n f_i\right) \leq (4M+1) \sum_{i=1}^n \operatorname{Var} f_i.$$

Lemma 1.3. Suppose $\{U_n, n \geq 1\}$ is a stationary process $\phi$-mixing with mixing coefficient $\phi(.)$ satisfying

$$\sum_{j=1}^\infty \phi^{\frac{1}{2}}(j) < \infty.$$

Let $\chi(.)$ be a real valued measurable function such that $E|\chi(U_1)|^2 < \infty$ for some $\delta > 0$. Further suppose that $E[\chi(U_1)] = 0$ and $\chi(U_1)$ is non-degenerate. Then

$$\frac{1}{\sqrt{n}\sigma} \sum_{j=1}^n \chi(U_j) \xrightarrow{L} N(0,1)$$

where $\sigma^2 = E[\chi(U_1)]^2 + 2\sum_{n=1}^\infty E[\chi(U_1)\chi(U_n)]$.

For proofs of Lemmas 1.1-1.3, we refer the reader to Iosifescu and Theodorescu (1969) (cf. Ibragimov (1962)).

Lemma 1.4. Let $\{U_n, n\geq 1\}$ be a stationary $\phi$-mixing processes with mixing coefficient $\phi(.)$ satisfying

$$\sum_{j=1}^{\infty} \phi^{\frac{1}{2}}(j) < \infty.$$

Lef f be a real valued measurable function such that $0<E[f(U_1)]^2< \infty$. Define

$$\sigma^2 = \mathrm{Var}[f(U_1)]+2 \sum_{i=2}^{\infty} \mathrm{Cov}[f(U_1), f(U_i)].$$

Further suppose that $\{\beta_{in}, 1\leq i\leq n, n\geq 1\}$ is a double sequence of real numbers such that

$$\sup_{1\leq i\leq n} n^{\frac{1}{2}}|\beta_{in} - n^{-\frac{1}{2}}| = o(1).$$

Then

$$\sum_{i=1}^{n} [f(U_i)-E(f(U_i))]\beta_{in} \xrightarrow{L} N(0,\sigma^2).$$

Proof: - In view of Lemma 1.3, it is sufficient to prove that

$$R_n \equiv \sum_{i=1}^{n} [f(U_i)-E(f(U_i))]\beta_{in} - n^{-\frac{1}{2}} \sum_{i=1}^{n} [f(U_i)-E(f(U_i))] \xrightarrow{p} 0 \text{ as } n\to\infty.$$

But $E(R_n)=0$ and

$$\text{Var }(R_n) \le 2 \sum_{i=1}^{n} \sum_{j=1}^{n} \phi^{\frac{1}{2}}(|i-j|)\ \text{Var}^{\frac{1}{2}}(f(U_i))\ \text{Var}^{\frac{1}{2}}(f(U_j)).$$

$$(\beta_{in}-\frac{1}{\sqrt{n}})\ (\beta_{jn}-\frac{1}{\sqrt{n}})$$

(by Lemma 1.1)

$$\le 2\ \text{Var }(f(U_1)) \sup_{1\le i\le n} |\beta_{in}-\frac{1}{\sqrt{n}}|^2 . \sum_{i=1}^{n} \sum_{j=1}^{n} \phi^{\frac{1}{2}}(|i-j|)$$

$$=2\ \text{Var }(f(U_1))\{\sup_{1\le i\le n} n|\beta_{in}-\frac{1}{\sqrt{n}}|^2\}\{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \phi^{\frac{1}{2}}(|i-j|)\}.$$

But

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \phi^{\frac{1}{2}}(|i-j|) < \infty$$

since $\sum_{j=1}^{\infty} \phi^{\frac{1}{2}}(j)<\infty$. Furthermore $\sup_{1\le i\le n} n|\beta_{in}-\frac{1}{\sqrt{n}}|^2=o(1)$ by hypothesis.

Hence Var $(R_n)\to 0$ as $n\to\infty$. Since $E(R_n)=0$, we obtain that $R_n \xrightarrow{p} 0$ as $n\to\infty$.

## 2. Asymptotic Theory

Let us consider the general linear model

$$(2.0)\qquad \underset{\sim}{Y}_n = X_n\theta + \underset{\sim}{U}_n$$

where $X_n$ is a given nxp-matrix, $\theta$ is the unknown p-dimensional vector, $\underset{\sim}{U}_n= (u_1,..,u_n)$ is the error vector with $\{u_i, i\ge 1\}$ forming a stationary process which is $\phi$-mixing with mixing coefficient $\phi(u)$ and $\underset{\sim}{Y}_n=(Y_{1n},\ldots,Y_{nn})$ is the vector of observations. Let $\underset{\sim}{X}'_{in}\in R^p$ be the ith row of $X_n$ where p possibly dependent on n.

Let $\chi(.)$ be a non-decreasing function and consider the equation

$$(2.1) \qquad \sum_{i=1}^{n} \chi(Y_{in}-\underset{\sim}{X}'_{in}\theta)\underset{\sim}{X}_{in}=0.$$

Any solution $\hat{\theta}_n$ satisfying (2.1) is called an M-estimator of $\theta$. (cf. Huber (1972)). Assume that

(A0) $\sum_n \phi^{\frac{1}{2}}(n)<\infty$,

and

(A1) $X'_n X_n$ is non-singular for large n (say) $n\geq n_0$.

Hereafter we assume that $n\geq n_0$. Let $M_n$ be any pxp matrix such that $M'_n M_n = X'_n X_n$. Let

$$(2.2) \qquad \theta^*=M_n\theta, \hat{\theta}^*_n=M_n\hat{\theta}_n \text{ and } \underset{\sim}{Z}_{in}=(M'_n)^{-1}\underset{\sim}{X}_{in}.$$

Then $\hat{\theta}^*_n$ is a solution of

$$(2.3) \qquad \sum_{i=1}^{n} \chi(Y_{in}-\underset{\sim}{Z}'_{in}\theta)\underset{\sim}{Z}_{in}=0.$$

Since we are interested in the asymptotic behaviour of the M-estimator $\hat{\theta}_n$ or equivalently $\hat{\hat{\theta}}^*_n$, we assume that $\theta^*=0$ without loss of generality. In this case, we can write (2.3) in the form.

$$(2.4) \qquad \sum_{i=1}^{n} \chi(u_i-\underset{\sim}{Z}'_{in}\theta)Z_{in}=0.$$

In addition to assumption (A1), let us suppose that the following conditions are satisfied.

(A2) $\chi(.)$ is non-decreasing and there exist b>0, c>0 and d>0 such that

$$\frac{\chi(u+\bar{z})-\chi(u)}{z} \geq d \text{ if } |u|\leq c \text{ and } |z|\leq b$$

where $q=F(c)-F(-c)>0$. Here $F(.)$ is the distribution of $u_1$.

(A3) $E_F(\chi^2(u))=v<\infty$, $E_F(\chi(u))=0$.

Note that $\sum_{i=1}^{n} |\underset{\sim}{Z}_{in}|^2 = p$ and $\sum_{i=1}^{n} \underset{\sim}{Z}_{in}\, \underset{\sim}{Z}'_{in} = I$

where $|\underset{\sim}{\alpha}|$ is the Euclidean norm of $\underset{\sim}{\alpha}$ and I is an identity matrix of order pxp.

Lemma 2.1. For any $i_1, i_2, .., i_\ell$ in 1 to n,

$$P(|\sum_{j=1}^{\ell} \chi(u_{i_j})\underset{\sim}{Z}_{jn}|\geq k) \leq 4Mpv/k^2.$$

Proof - Note that $E(\sum_{j=1}^{\ell} \chi(u_{i_j})\underset{\sim}{Z}_{jn})=0$ and

$$\begin{aligned} \text{Var } &(\sum_{j=1}^{\ell} \chi(\chi_{i_j})\underset{\sim}{Z}_{jn}) \\ &\leq (4M+1)\sum_{j=1}^{\ell} \text{Var } [\chi(u_{i_j})]\, \underset{\sim}{Z}'_{jn}\, \underset{\sim}{Z}_{jn} \qquad \text{(by Lemma 1.2)} \\ &= (4M+1)v\sum_{j=1}^{\ell} |\underset{\sim}{Z}_{jn}|^2 \\ &\leq (4M+1)vp \end{aligned}$$

and the result follows by Chebyshev's inequality.

Let

$$D(\varepsilon)= d\sum_{j=1}^{n}\underset{\sim}{Z}_{jn}\, \underset{\sim}{Z}'_{jn}\; I_{[|u_j|\leq c]}\; I_{[|\underset{\sim}{Z}_j|\leq\varepsilon]}$$

and $D_0(\varepsilon)=E\, D(\varepsilon)$ where $I_A$ denotes the indicator function of set A.

For any matrix A, define $||A||^2=\mathrm{trace}(A'A)$. With these notations the following lemmas can be proved. The proofs of these are the same as those in Yohai and Maronna (1979) as the independence of $\{u_i, i\geq 1\}$ is not used in proving these lemmas. We omit the details.

Lemma 2.2. For any $\delta>0$,

$$P(||D(\varepsilon)-D_0(\varepsilon)||\geq\delta)\leq r^2\varepsilon^2 p/\delta^2$$

where $r^2=d^2q(1-q)$.

Lemma 2.3. Let $K_0$ be chosen so that

$$\chi(-\infty)<-K_0<0<K_0<\chi(\infty).$$

Let $J_0$ be a subset of 1 to n with cardinality m. Let $\eta>0$ and define

$$T=\{\underset{\sim}{t}:\eta\leq|t_j|\leq 1, 1\leq j\leq n\}.$$

Then, for any $\delta>0$, there exists $L=L(\eta,\delta,m)$ which does not depend on n such that

$$P(\sup_{\substack{J\subset J_0\\ t\in T}}\sum_{j\in J}[\chi(u_j-Lt_j)t_j+K_0|t_j|]\geq 0)\leq\delta.$$

As a consequence of Lemmas 2.1-2.3, we obtain the following theorem as in Yohai and Maronna (1979):

Theorem 2.1. Assume (A0)-(A3). Then, for any fixed p,

(2.5) $$|\hat{\theta}^*-\theta^*|=O_p(1).$$

If $p=p_n$ depends on n, and $\lim_n p_n \max_{1\leq i\leq n}|\underset{\sim}{Z}_{in}|^2=0$, then

(2.6) $$|\hat{\theta}^*_n-\theta^*|=O_p(p_n^{\frac{1}{2}}).$$

In particular, if p is fixed and the smallest eigen value $\lambda_n$ of $X_n' X_n$ tends to infinity as $n \to \infty$, then $\hat{\theta}_n \xrightarrow{p} \theta$ as $n \longrightarrow \infty$.

Theorem 2.2. In addition to assumptions (A0)-(A3), further suppose that the following conditions hold:

(B1) $\chi(.)$ is three times differentiable with a bounded third derivative i.e., $|\chi'''(x)| < c < \infty$ for all x,

(B2) $E_F|\chi'(u)|$ and $E_F|\chi''(u)|^2$ are finite,

(B3) $E_F\, \chi''(u) = 0$,

(B4) $\lim_n p_n^{3/2}\, \varepsilon_n = 0$ where $\varepsilon_n = \max_{1 \le i \le n} |\underset{\sim}{Z}_{in}|^2$

and

(B5) there exists $\underset{\sim}{\alpha}_n \in R^{p_n}$ with $|\underset{\sim}{\alpha}_n| = 1$ such that

$\sup_i |\sqrt{n}\, \underset{\sim}{Z}_{in} - \underset{\sim}{\alpha}_n| \longrightarrow 0$ as $n \longrightarrow \infty$.

Then

$$\underset{\sim}{\alpha}_n'(\hat{\theta}_n^* - \theta^*) \xrightarrow{L} N(0, \zeta^2) \tag{2.7}$$

where $\zeta^2 = (E_F[\chi^2(u_1)] + 2\sum_{i=2}^{\infty} E_F[\chi(u_1)\chi(u_i)]) / (E_F\chi'(u_1))^2$.

Prob.: Assume that $\theta^*=0$ without loss of generality. Then $\hat{\theta}_n^*$ is a solution of

$$\sum_{i=1}^{n} \chi(u_i - \underset{\sim}{Z}_{in}'\, \theta)\underset{\sim}{Z}_{in} = 0. \tag{2.8}$$

By Taylor's expansion,

$$0 = \sum_{i=1}^{n} \chi(u_i - \underset{\sim}{Z}_{in}'\, \hat{\theta}_n^*)\underset{\sim}{Z}_{in}'\, \underset{\sim}{\alpha}_n \tag{2.9}$$

$$= w_1 - w_2 - \underset{\sim}{W}_3' \hat{\theta}_n^* + \frac{1}{2} \hat{\theta}^{*\prime} W_4 \theta^* - \frac{1}{6} w_5$$

where

(2.10) $w_1 = \sum_{i=1}^{n} \chi(u_i) \underset{\sim}{Z}_{in}' \underset{\sim}{\alpha}_n,$

(2.11) $w_2 = (E_F \chi') \sum_{i=1}^{n} \underset{\sim}{\alpha}_n' \underset{\sim}{Z}_{in} \underset{\sim}{Z}_{in}' \hat{\theta}_n^*,$

(2.12) $\underset{\sim}{W}_3 = \sum_{i=1}^{n} [\chi'(u_i) - E_F \chi'] (\underset{\sim}{\alpha}_n' \underset{\sim}{Z}_{in}) \underset{\sim}{Z}_{in},$

(2.13) $W_4 = \sum_{i=1}^{n} \chi''(u_i) (\underset{\sim}{\alpha}_n' \underset{\sim}{Z}_{in}) \underset{\sim}{Z}_{in} \underset{\sim}{Z}_{in}',$

and

(2.14) $w_5 = \sum_{i=1}^{n} \chi'''(u_i + \eta_{in} \underset{\sim}{Z}_{in}' \hat{\theta}_n^*) (\underset{\sim}{Z}_{in}' \hat{\theta}_n^*)^3 \underset{\sim}{\alpha}_n' \underset{\sim}{Z}_{in}$

with $|\eta_{in}| < 1$.

Since $|\underset{\sim}{\alpha}_n| = 1$ and $\sup_{1 \le i \le n} |\sqrt{n}\, \underset{\sim}{Z}_{in} - \underset{\sim}{\alpha}_n| \longrightarrow 0$, it follows that

$$\sup_{1 \le i \le n} |\sqrt{n}\, \underset{\sim}{Z}_{in}' \alpha_n - \underset{\sim}{\alpha}_n' \alpha_n| \longrightarrow 0 \text{ as } n \to \infty$$

and hence

$$\sup_{1 \le i \le n} \sqrt{n} |\beta_{in} - \frac{1}{\sqrt{n}}| = o(1)$$

where $\beta_{in} = \underset{\sim}{Z}_{in}' \underset{\sim}{\alpha}_n$. Hence, by Lemma 1.4, we obtain that

$$\Sigma \chi(u_i) \beta_{in} \xrightarrow{L} N(0, \sigma^2)$$

where

(2.15) $\sigma^2 = E[\chi^2(u_1)] + 2 \sum_{i=2}^{\infty} E(\chi(u_1) \chi(u_i)].$

Therefore

(2.16) $\quad w_1 \xrightarrow{L} N(0,\sigma_1^2)$.

Note that

(2.17) $\quad w_2=(E_F\chi')\underset{\sim}{\alpha}_n'(\sum_{i=1}^{n}\underset{\sim}{Z}_{in}\,\underset{\sim}{Z}_{in}')\hat{\theta}_n^*$.

and

$$\sum_{i=1}^{n}\underset{\sim}{Z}_{in}\underset{\sim}{Z}_{in}'$$

is identity matrix of order $p_n x p_n$. Observe that $E(\underset{\sim}{W}_3)=0$ and

(2.20) $\quad p_nE|\underset{\sim}{W}_3|^2=p_n\sum_{\ell=1}^{p_n}E|W_{3\ell}|^2$

$$=p_n\sum_{\ell=1}^{p_n}\mathrm{Var}(\sum_{i=1}^{n}\chi'(u_i)\underset{\sim}{\alpha}_n'\underset{\sim}{Z}_{in}Z_{in\ell})$$

$$\leq(4M+1)p_n\sum_{i=1}^{n}\mathrm{Var}\,[\chi'(u_i)](\underset{\sim}{\alpha}_n'\underset{\sim}{Z}_{in})^2|\underset{\sim}{Z}_{in}|^2$$

(by Lemma 1.2)

$$\leq(4M+1)p_n\mathrm{Var}(\chi'(u_1))\varepsilon_n\sum_{i=1}^{n}(\underset{\sim}{\alpha}_n'\underset{\sim}{Z}_{in})^2$$

$$\leq C_1p_n\varepsilon_n$$

for some constant $C_1>0$ since $\sum_{i=1}^{n}(\underset{\sim}{\alpha}_n'\underset{\sim}{Z}_{in})^2=1$. The last term tends to zero as $n\to\infty$ by (B4). On the other hand $||W_4||^2=$ trace $(W_4W_4')$ and $E(\chi''(u_1))=0$.

Hence

$$p_n^2 E||W_4||^2 \leq (4M+1) p_n^2 \mathrm{Var}(\chi''(u_1)) \sum_{i=1}^{n} (\underset{\sim}{\alpha}_n' \underset{\sim}{Z}_{in})^2 |\underset{\sim}{Z}_{in}|^4$$

by arguments analogous to those given above and the last term is bounded by

$$(2.21) \quad C_2 p_n^2 \varepsilon_n^2$$

for some constant $C_2 > 0$ since $\sum_{i=1}^{n} (\underset{\sim}{\alpha}_n' \underset{\sim}{Z}_{in})^2 = 1$. But $p_n \varepsilon_n \to 0$ by (B4). Clearly

$$|w_5| \leq c \sum_{i=1}^{n} (\underset{\sim}{Z}_{in}' \hat{\theta}_n^*)^2 |\hat{\theta}_n^*| |\underset{\sim}{Z}_{in}|^2$$

where c is given by (B1) and hence

$$(2.22) \quad |w_5| \leq c|\hat{\theta}_n^*| \varepsilon_n \sum_{i=1}^{n} (\underset{\sim}{Z}_{in}' \hat{\theta}_n^*)^2$$

$$= c|\hat{\theta}_n^*|^3 \varepsilon_n$$

$$= |p_n^{-\frac{1}{2}} \hat{\theta}_n^*|^3 c p_n^{3/2} \varepsilon_n$$

$$= c p_n^{3/2} \varepsilon_n O_p(1)$$

by theorem 2.1. The last term tends to zero by hypothesis.

Acknowledgement: The author thanks Professor Shanti S. Gupta and the Department of Statistics of Purdue University for inviting him to spend the Summer 1979 during which thime this work was done.

## REFERENCES

[1] Deniau, C. Oppenheim, G., Viano, M.C.(1977) Asymptotic normality of M-estimators on weakly dependent data. In Recent Developments in Statistics (J.R. Barra et. al. editors). North Holland.

[2] Huber, P.J.(1967) The behaviour of maximum likelihood estimates under nonstandard conditions. Proc. Fifth Berkeley Symp. Math Statist. and Prob. 1, 221-233.

[3] Huber, P.J.(1972) Robust regression: Asymptotics, Conjectures and Monte Carlo. Ann. Statistics 1, 799-821.

[4] Ibragimov, I.A.(1962) Some limit theorems for stationary processes. Theory of Probability and its Applications. 7, 349-382.

[5] Iosifescer, M., Theodorescu, R.(1969) Random Processes and Learning, Springer-Verlag, Berlin.

[6] Prakasa Rao, B.L.S.(1972) Maximum likelihood estimation for Markov processes. Ann. Inst. Statist. Math. 24, 333-345.

[7] Yohai, V.J., Maronna, R.(1979) Asymptotic behaviour of M-estimators for the linear model. Ann. Statistics 1, 258-268.

Department of Statistics
Purdue University
West Lafayette, IN 47906