1999 VOL. X(XVIII)

HAJIYEV A.H.

STATISTICAL ESTIMATORS IN REGRESSION MODELS WITH DEPENDENT ERRORS OF OBSERVATIONS

Abstract

The non-linear regression models with increasing number of unknown parameters and unknown dispersions of errors of observation were considered. The properties of estimations of least squares (l.e.s) and M - estimators are investigated. The estimations of elements of covariational matrix of remainders vector, which were used for construction of confidence stripe for unknown function in regression model, were constructed using l.e.s. and M - estimators.

I. Introduction. In present paper the properties of estimators of the elements of covariance matrix in nonlinear regression models where the number of unknown parameters increase when the number of observation goes up are investigated. There are a lot of papers where the linear regression models both with finite [1-5], and with growth number of unknown parameters [6-8] are investigated. The extension of such investigations is working out the methods for investigation of nonlinear regressions. In [9-11] were investigated properties of least square estimators (l.s.e) in nonlinear regression models with finite umber of unknown parameters. The main attention in this paper is making for investigation of properties of l.s.e and M - estimators of unknown parameter. However, for more complete investigation of regression dependence the essential role plays investigators of estimators of elements of covariance matrix of remains vector.

Some results for the case of linear regression models were obtained in [12,13].

II. The statement of problem. Consider regression model

$$y_i = \eta(x_i, \theta) + \varepsilon_i , \tag{1}$$

where

$y = (y_1, y_2, ..., y_N)^T$ is the vector of observations

$\theta = (\theta_1, \theta_2, ..., \theta_m)^T$ is the vector of unknown parameters

$\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_N)^T$ is the vector of errors

x_i is the point of observation

$\eta(x_i, \theta)$ is a nonlinear by θ function

θ^* is the real value of parameter θ

We suppose:

the number of unknown parameters m depends on N;

$\eta(x,\theta), \dfrac{\partial \eta(x,\theta)}{\partial \theta}, \dfrac{\partial^2 \eta(x,\theta)}{\partial \theta_i \partial \theta_j}$ $i, j = \overline{1,m}$ are continuous by (x,θ) and bounded functions; (2)

$\theta \in \Theta$ is a compact set (3)

$E\varepsilon_i = 0, E\varepsilon_1^2 = \sigma_i^2$ are unknown, $\sigma_i^1 \le \sigma_0^2$;

$E\varepsilon_i E\varepsilon_j = r_{ij}; \quad E\varepsilon_i^4 \le C < \infty$ (4)

Denote $\hat{\theta}$ least square estimator $f_{ij}(\theta)=\dfrac{\partial\eta(x_j,\theta)}{\partial\theta_i}, i=\overline{1,m}; j=\overline{1,N}$; $F(\theta)$ is the matrix with elements $f_{ij}(\theta)$;

$O<\lambda_1^{(N)}(\theta)\le\lambda_2^{(N)}(\theta)\le\ldots\le\lambda_m^{(N)}(\theta)$ are eigenvalues of matrix $\left[\dfrac{F^T(\theta)F(\theta)}{N}\right]$, $B(r)$ is the sphere with radius $r>0$ as a center at the point θ^*. For determination of l.s.e of unknown parameter θ^* we can use iteration procedure [14, p.97].

$$\theta(s+1)=\theta(s)+\left[F^T(\theta(s))\cdot F(\theta(s))\right]^{-1}F^T(\theta(s))(y-\eta(x,\theta(s))), \qquad (5)$$

where the quantities in equation (5) were described above. Investigate conditions for which process (5) will be convergent.

III. The properties of l.s.e. To prove convergence of iteration procedure (5) we should first prove some auxiliary statements. Below we will use method introduced in [5] for investigation of generalized regression models.

Rewrite (5) in the following form

$$\theta_N(s+1)=U(\theta_N(s))=\theta_N(s)+A_N(\theta_N(s))\delta(\theta_N(s)), \qquad (6)$$

where

$$A_N(\theta)=\left[\frac{F^T(\theta)F(\theta)}{N}\right]^{-1}\frac{F^T(\theta)}{N},$$

$$\delta(\theta)=y-\eta(x,\theta);\ \ \delta^*=y-\eta(x,\theta^*)=\varepsilon.$$

Denote by

$$\zeta_N(\theta)=m\frac{\partial A_N(\theta)}{\partial\theta_p}\varepsilon,\quad \theta\in B(r).$$

Let the following matrices will be given

$$An_1=\left(a^{(n_1)}\right)_{ij};\quad B_{n_2}=\left(b^{(n2)}\right)_{ij};\quad i=\overline{1,m};\quad j=\overline{1,N};\quad C=(c)_{i_1j_1};\quad i_1,j_1=\overline{1,m}.$$

Denote by

$$D=A_1A_2^T\ldots A_{k-1}^TA_kC^nB_1^TB_2\ldots B_{i=1}B_i^T,$$

where $k=2k_1+1; l=2l_1+1; n,l_1,k_1$ are integer numbers.

Lemma 1. *If there exists constant value* C_1 *such that*

$$\left|a_{ij}^{(n_1)}\right|\le C_1,\quad \left|b_{ij}^{(n_2)}\right|\le C_1,\quad \max_{1\le i\le m}\sum_{j=1}^m\left|c_{ij}\right|\le C_1.$$

Then

$$E\left(\varepsilon^TD\varepsilon\right)\le C_1^n m^{k_1+l_1+1}\quad N^{k_1+l_1+1}\left(C_1N+\sum_{i\ne j}\left|r_{ij}\right|\right). \qquad (7)$$

Proof. Denote by $d_{ij}, i,j=\overline{1,N}$ the elements of matrix D. Then

$$E\left(\varepsilon^TD\varepsilon\right)=E\sum_{i=1}^N\sum_{j=1}^N d_{ij}\varepsilon_i\varepsilon_j=\sum_{i=1}^N d_{ii}E\varepsilon_i^2+\sum_{i\ne j}d_{ij}E\varepsilon_i\varepsilon_j\le$$
$$\le\sigma_0^2 trD+\sum_{i=j}d_{ij}\left|r_{ij}\right| \qquad (8)$$

Using conditions of Lemma 1, we have

$$d_{ij}=\sum_{i_1=1}^{N}\dots\sum_{i_{k_1}=1}^{N}\sum_{j_1=1}^{N}\dots\sum_{j_{k_1+1}=1}^{m}\sum_{p_1=1}^{m}\dots\sum_{p_n=1}^{m}\sum_{r_1=1}^{N}\dots\sum_{r_{l_1}=1}^{N}\times$$

$$\times\sum_{s_1=1}^{m}\dots\sum_{s_{l_1}}^{m}a_{ij_1}^{(1)}a_{j_1i_1}^{(2)}\dots a_{i_{k_1}}^{(k_1)}p_1^C p_1p_2\dots c_{p_{n-1}p_n}\times$$

$$\times b_{p_nr_1}^{(1)^T}b_{r_1s_1}^{(2)}\dots b_{s_{l_1}}^{(i)}j\le C_1^nN^{k_1}m^{k_1+1}N^{l_1}m^{l_1+1}=$$

$$=C_1N^{k_1+l_1}m^{k_1+l_1+2}$$

$$trD=\sum_{i=1}^{N}d_{ii}\le C_1^{n+1}N^{k_1+i_1}m^{k_1+l_1+2}N$$

Taking into account (8), we obtain the statement of Lemma 1.

Further we will suppose $\theta\in B(r)$.

Theorem 1. *Let for the model (1) conditions (2)-(4) are hold. If*

$$\frac{m^4\sqrt{m}}{N(\lambda_1(\theta))^3}+\frac{m^4}{N^2(\lambda_1(\theta))^3}\sum_{i\neq j}|r_{ij}|\to 0\quad for\quad N\to\infty, r\to 0, \tag{9}$$

$$\frac{m^5}{N(\lambda_1(\theta))^4}+\frac{m^5}{N^2(\lambda_1(\theta))^4}\sum_{i\neq j}|r_{ij}|\to 0\quad for\quad N\to\infty, r\to 0. \tag{10}$$

Then

$$\zeta_N(\theta)\xrightarrow{P}0\quad \text{for}\quad N\to\infty,\ r\to 0.$$

Proof. From matrix analyses [16, p.156] it follows

$$\frac{\partial A_N(\theta)}{\partial\theta_p}=\frac{1}{N}\left(\frac{F^T(\theta)F}{N}\right)^{-1}\frac{\partial F^T(\theta)}{\partial\theta_p}+\frac{1}{N}\left(\frac{F^T(\theta)F}{N}\right)^{-2}\times$$

$$\times\left[\left(\frac{\partial F^T(\theta)}{\partial\theta_p}F(\theta)+F^T(\theta)\frac{\partial F(\theta)}{\partial\theta_p}\right)\Big/N\right]F^T.$$

(11)

Denote:

$\bar f_{ij}(\theta)$- the elements of matrix $\left(\frac{F^T(\theta)F(\theta)}{N}\right)^{-1}$, $i,j=\overline{1,m}$;

$\partial f_{kl}(\theta)$ - the elements of matrix $\frac{\partial F(\theta)}{\partial\theta_p}$, $k=\overline{1,m}$; $l=\overline{1,N}$.

Then from (11) it follows that

$$E\left\|m\frac{\partial A_N(\theta)}{\partial\theta_p}\varepsilon\right\|=m\varepsilon\left(\varepsilon^T\frac{\partial A_N^T(\theta)}{\partial\theta_p}\frac{\partial A_N(\theta)}{\partial\theta_p}\varepsilon\right)=$$

$$=E\left[\varepsilon^T\frac{\partial F(\theta)}{\partial\theta_p}\left(\frac{F^TF}{N}\right)^{-1}\left(\frac{F_N^T(\theta)F(\theta)}{N}\right)^{-1}\frac{\partial F^T(\theta)}{\partial\theta_p}\varepsilon\right]\frac{m}{N^2}+$$

$$+\frac{m}{N^2}E\left[\varepsilon^T\frac{\partial F(\theta)}{\partial\theta_p}\left(\frac{F^TF}{N}\right)^{-3}\left(\frac{\partial F^T}{\partial\theta_p}F+F^T(\theta)\frac{\partial F(\theta)}{\partial\theta_p}\right)F^T(\theta)\varepsilon\right]+$$

$$\times\left(\frac{\partial F^T(\theta)}{\partial\theta_p}F(\theta)+F^T(\theta)\frac{\partial F(\theta)}{\partial\theta_p}\right)F^T(\theta)\varepsilon\Bigg],\quad \theta(r)\ . \tag{12}$$

Denote by $I_i(\theta)$ $i=\overline{1,4}$ the i th member in (12). Using relation [17, p.18]

$$\max_{1\le i\le m}\sum_{j=1}^{m}\left|f_{ij}(\theta)\right|\le\sqrt{m}/\lambda_1^{(N)}(\theta),\quad \theta\in B(r) \tag{13}$$

and Lemma 1, we have

$$I_1\le m\frac{C}{N^2}m\left(\frac{\sqrt{m}}{\lambda_1^{(N)}(\theta)}\right)^2\left(N+\sum_{i\ne j}|r_{ij}|\right)=C_1\left(\frac{m^3}{N(\lambda_1)^2}+\right.$$

$$\left.+\frac{m^3}{N^2(\lambda_1(\theta))^2}\sum_{i\ne j}|r_{ij}|\right)\to 0,\quad \theta\in B(r)\quad \text{for}\quad N\to\infty, r\to 0$$

by virtue of condition (9).

Open the brackets in expression for $I_2(\theta)$ and denote by $J_{2,i}(\theta)$ the i th member in $I_2(\theta)$

$$J_{2,1}(\theta)=\frac{m}{N^2}E\left[\varepsilon^T\frac{\partial F(\theta)}{\partial\theta_p}\left(\frac{F^T(\theta)F(\theta)}{N}\right)^{-3}\frac{\partial F^T(\theta)}{\partial\theta_p}F(\theta)F^T(\theta)\varepsilon\right].$$

So as for I_1, we have

$$J_{2,1}(\theta)\le\frac{C}{N^2m^3}\left(\frac{\sqrt{m}}{\lambda_1(\theta)}\right)^3\left(N+\sum_{i\ne j}|r_{ij}|\right)=0C\frac{m^4\sqrt{m}}{N(\lambda_1(\theta))^3}+$$

$$+\frac{m^4\sqrt{m}}{(\lambda_1(\theta))^3}\frac{1}{N^2}\sum_{i\ne j}|r_{ij}|\to 0,\quad \theta\in B(r)\quad \text{при}\quad N\to\infty, r\to 0$$

by virtue condition (9).

By the same way it is proved that $J_{2,2}(\theta)\to 0$ and $I_3(\theta)\to 0$ for $N\to\infty, r\to 0$. Open the brackets in $I_4(\theta)$. Denote by $J_{4,i}(\theta)$ the i th number in $I_4(\theta)$.

$$I_{4,1}(\theta)=\frac{m}{N^2}E\left[\varepsilon^T F(\theta)\frac{\partial F(\theta)}{\partial\theta_p}F(\theta)\left(\frac{F^T(\theta)F(\theta)}{N}\right)^4\frac{\partial F^T(\theta)}{\partial\theta_p}F(\theta)F^T(\theta)\varepsilon\right]$$

$\theta\in B(r)$.

According to the relation (13) and Lemma 1, we have

$$J_{4,1}(\theta)=\frac{1}{N^2}Cm^3\left(\frac{\sqrt{m}}{\lambda_1(\theta)}\right)^4\left(N+\sum_{i\ne j}|r_{ij}|\right)=\frac{C}{N}\frac{m^5}{\lambda_1^4(\theta)}+$$

$$+\frac{m^5}{N^2\lambda(\theta)}\sum_{i\ne j}|r_{ij}|\to 0,\quad \text{for}\quad N\to\infty, r\to 0$$

by virtue of condition of Theorem 1.

The rest members $J_{4,i}(\theta)$ $i=2,3,4$; have the same form as $J_{4,1}(\theta)$. Therefore $I_4(\theta)\to 0$ for $N\to\infty, r\to 0$.

Then from Tchebyshev's inequality it follows

$$P\left\{\left\| m\frac{\partial A(\theta)}{\partial\theta_p}\varepsilon\right\| > \alpha\right\} \le \frac{E\left\{\varepsilon^T\frac{\partial A_N^T(\theta)}{\partial\theta_p}\frac{\partial A_N(\theta)}{\partial\theta_p}\varepsilon\right\}}{a^2} \le$$

$$\le \frac{I_1+I_2+I_3+I_4}{a^4} \to 0 \quad \text{for } N\to\infty,\ r\to 0$$

Now we pass to the proof of convergence of iteration procedure (5). Denote by

$$L_p = \frac{\partial u(\theta)}{\partial\theta_p};\ \tau(r) = \max_{p=1,\dots,m}\sup_{\theta\in B(r)}\|L_p\|,\quad \rho(\theta)=U(\theta)-\theta,\quad \rho^* = \rho(\theta^*).$$

Theorem 2. *Let*

$$\tau(r) + \frac{\|\rho^*\|}{r} < 1.$$

Then under conditions of theorem 1 there exists such random variable $\theta_N(\infty)$, that $(\theta_N(\infty)-\theta_N(s))\xrightarrow{P}0$ for $s\to 0$.

Proof. As $\frac{\partial U(\theta)}{\partial\theta}$ exists by virtue of condition (3), then by Lagrange's formula we have

$$\|U(\theta)-U(\theta^*)\| \le \left\|\frac{\partial U}{\partial\theta}\right\|\cdot\|\theta-\theta^*\| .$$

Consider

$$\left\|\frac{\partial U(\theta)}{\partial\theta}\right\| = \left\|\sum_{p=1}^{m}\frac{\partial U(\theta)}{\partial\theta_p}\right\| \le m\sup_{1\le p\le m}\left\|\frac{\partial U(\theta)}{\partial\theta_p}\right\| . \tag{14}$$

As

$$\frac{\partial\delta(\theta)}{\partial\theta_p} = -\frac{\partial\eta(x,\theta)}{\partial\theta_p} = -F(\theta)$$

then taking into account relation (6), we have

$$m\frac{\partial U(\theta)}{\partial\theta_p} = \left[I_p + \frac{\partial A_N(\theta)}{\partial\theta_p}\delta(\theta) + A_N(\theta)\frac{\partial\delta(\theta)}{\partial\theta_p}\right]m =$$

$$= m\left[\frac{\partial A_N(\theta)}{\partial\theta_p}\delta(\theta^*) + \frac{\partial A_N(\theta)}{\partial\theta_p}\left(\delta(\theta)-\delta(\theta^*)\right)\right] = \tag{15}$$

$$= \frac{\partial A_N(\theta)}{\partial\theta_p}\varepsilon m + \frac{\partial A_N(\theta)}{\partial\theta_p}\Delta\eta(x,\theta,\theta^*)m ,$$

where $\Delta\eta(x,\theta,\theta^*) = \eta(x,\theta^*) - \eta(x,\theta)$.

First member in (15) for $N\to\infty$ tends to zero according to Theorem 1. Second member also tends to zero by virtue of the fact that $\Delta\eta(x,\theta,\theta^*)\to\infty$ for $r\to 0$ uniformly by x. Then taking into account (14) and (15) we have that starting from some r and s the mapping $u(0)$ is compressed. From the other hand,

$$\|u(\theta)-\theta^*\| \le \|u(\theta)-u(\theta^*)\| + \|u(\theta^*)-\theta^*\| \le \tau(r)r + \|\rho^*\| \le r,\ \theta\in B(r) \tag{16}$$

according by Theorem 2. Thus, the mapping $u(\theta)$ maps the sphere $B(r)$ into itself. So as $B(r)$ is complete metric space, then according to the method of compressed mappings [18, p.75] for all $\theta(0)\in B(r)$, $s\to\infty$ equation

$$U(\theta)=\theta$$

has the unique solution which we denote by $\theta_N(\infty)$. So as

$$u(\theta_N(s-1))=\theta_N(s)$$

and

$$\|u(\theta_N(s))-u(\theta_N(\infty))\|<r\to 0,$$

then it follows

$$\|\theta_N(s)-\theta_N(\infty)\|\xrightarrow{P}0 \quad \text{for } s\to\infty$$

and this fact proves theorem 2.

Theorem 3. *Let*

$$\frac{m}{\lambda_1^{(N)}(\theta)}+\frac{m}{\left(\lambda_1^{(N)}(\theta)\right)^2}\sum_{i\neq j}|r_{ij}|<\infty,\ N\to\infty,\quad r\to 0\,.$$

Then under conditions of Theorem 2 the random variable

$$\sqrt{N}(\theta_N(\infty)-\theta_N(s)) \ \textit{for}\ N\to\infty$$

is bounded in probability.

Proof.

$$NE\|\rho\|=E(u(\theta)-\theta)^T(u(\theta)-\theta)=NE\varepsilon^T A_N^T(\theta)A_N(\theta)\varepsilon=$$

$$=E\varepsilon^T F(\theta)\left(\frac{F^T(\theta)F(\theta)}{N}\right)^{-1}\left(\frac{F^T(\theta)F(\theta)}{N}\right)^{-1}\frac{F^T(\theta)}{N}\varepsilon$$

Using the boundness of elements of matrix $F(\theta)$, the relation (13), Lemma 1 and Theorem 3 we obtain

$$E\|\rho(\theta)\|\le\sigma_0^2\frac{1}{N}tr\left(\frac{F^T(\theta)F(\theta)}{N}\right)^{-1}+\frac{m}{N^3\lambda_1^2(r)}\sum_{i\neq j}|r_{ij}|=$$

$$=\frac{C_1 m}{N\lambda_1(\theta)}+\frac{m}{N^2\lambda_1^2(\theta)}\sum_{i\neq j}|r_{ij}| \qquad \forall\theta\in B(r)$$

But as $\theta^*\in B(r)$, we have

$$NE\|\rho^*\|<A\,.$$

Using Tchebyshev's inequality, we obtain that random value $\sqrt{N}\rho^*$ is bounded in probability.

Let choose

$$\tau(r)=\frac{1}{2};\quad r=\frac{2k}{\sqrt{N}};\quad \|\rho^*\|<\frac{k}{\sqrt{N}}$$

then

$$\tau(r)+\frac{\|\rho^*\|}{r}<\frac{1}{2}+\frac{\|\rho^*\|\sqrt{N}}{2k}<1\,.$$

Taking into account

$$\|\theta_N(\infty)-\theta^*\|\le\sup_{\theta(s)\in B(r)}\|\theta_N(\infty)-\theta(s)\|,$$

and using theorem 2 we obtain

$$\lim_{k\to\infty}\lim_{N\to\infty} P\left\{\left\|\theta_N(\infty)-\theta^*\right\|>\frac{k}{\sqrt{N}}\right\}=0 .$$

Remark 1. The random variable $\theta(\infty)$(index N for simplicity let down) satisfies to the equality

$$A(\infty)\delta(\infty)=0$$

which is normal equation for l.s.e.

As

$$\theta(s+1)=u(\theta(s))=\theta(s)+A(\theta(s))\delta(\theta(s)).$$

then for $s\to\infty$ we have

$$\theta(\infty)=\theta(\infty)+A(\theta(\infty))\delta(\theta(\infty)).$$

Form here it follows

$$A(\infty)\delta(\infty)=0 .$$

Theorem 4. *Let ε_i are independent identically distributed random variables. Then under conditions of Theorem 3 it holds*

$$\sqrt{N}\left(\hat{\theta}-\theta^*\right)\to N\left(O,\Sigma\left(\theta^*\right)\right),$$

where

$$\Sigma\left(\theta^*\right)=\left(\frac{F^T\left(\theta^*\right)F\left(\theta^*\right)}{N}\right)^{-1}\frac{F^T\left(\theta^*\right)I\left(\sigma^2\right)F\left(\theta^*\right)}{N}\left(\frac{F^T\left(\theta^*\right)F\left(\theta^*\right)}{N}\right)^{-1},$$

$$I\left(\sigma^2\right)=E\varepsilon\varepsilon^T .$$

Proof. According to the statement I

$$A(\theta(\infty))\delta(\theta(\infty))=0 . \tag{17}$$

Then

$$\begin{aligned}A(\theta(\infty))\delta(\theta(\infty))&=\left[A\left(\theta^*\right)+A(\theta(\infty))-A\left(\theta^*\right)\right]\left[y-\eta\left(x,\theta^*\right)+\right.\\&+\eta\left(x,\theta^*\right)-\eta(x,\theta(\infty))\left.\right]=A\left(\theta^*\right)\left(y-\eta\left(x,\theta^*\right)\right)+A\left(\theta^*\right)\left(\eta\left(x,\theta^*\right)-\right.\\&-\eta(x,\theta(\infty))\left.\right)+\left[A(\theta(\infty))-A\left(\theta^*\right)\right]\left(y-\eta\left(x,\theta^*\right)\right)+\left[A(\theta(\infty))-A\left(\theta^*\right)\right]\times\\&\times\left(\eta\left(x,\theta^*\right)-\eta(x,\theta(\infty))\right)=\rho^*+\left[A(\theta(\infty))-A\left(\theta^*\right)\right]\delta^*+\\&+A(\theta(\infty))\Delta\eta\left(x,\theta^*,\theta(\infty)\right)=0\end{aligned} \tag{18}$$

by virtue of relation (17).

For $\left[A(\theta(\infty))-A\left(\theta^*\right)\right]$ and $\Delta\eta\left(x,\theta^*,\theta(\infty)\right)$ the simple identifies are true

$$A(\theta(\infty))-A\left(\theta'^*\right)=\left[\frac{\partial A\left(\theta^*\right)}{\partial\theta}+\int_0^1\left(\frac{\partial A\left(\theta^*+z\left(\theta(\infty)-\theta^*\right)\right)}{\partial\theta}-\right.\right.$$

$$\left.\left.-\frac{\partial A\left(\theta^*\right)}{\partial\theta}\right)dz\left(\theta(\infty)-\theta^*\right)\right]$$

$$\eta(\theta(\infty))-\eta\left(\theta'^*\right)=\left[\frac{\partial\eta\left(\theta^*\right)}{\partial\theta}+\int_0^1\left(\frac{\partial\eta\left(\theta^*+z\left(\theta(\infty)-\theta^*\right)\right)}{\partial\theta}-\right.\right.$$

$$\left.\left.-\frac{\partial\eta\left(\theta^*\right)}{\partial\theta}\right)dz\right]\left(\theta(\infty)-\theta^*\right). \tag{19}$$

Substituting (19) into (18), we obtain

$$\rho^* + \frac{\partial A(\theta^*)}{\partial\theta}\delta^*(\theta(\infty)-\theta^*) + \int_0^1\left(\frac{\partial A(\theta^* + z(\theta(\infty)-\theta^*))}{\partial\theta}\delta^*(\theta(\infty)-\theta^*) - \right.$$

$$-\frac{\partial A(\theta^*)}{\partial\theta}\delta^*(\theta(\infty)-\theta^*) + A(\theta(\infty))\frac{\partial\eta(\theta^*)}{\partial\theta}(\theta(\infty)-\theta^*) +$$

$$+ A(\theta(\infty))\int_0^1\left[\frac{\partial\eta}{\partial\theta}(\theta^* + z(\theta(\infty)-\theta^*))(\theta(\infty)-\theta^*) - A(\theta(\infty))\frac{\partial\eta(\theta(\infty))}{\partial\theta}\times\right.$$

$$\times(\theta(\infty)-\theta^*) - A(\theta(\infty))\frac{\partial\eta(\theta^*)}{\partial\theta}(\theta(\infty)-\theta^*) = 0. \tag{20}$$

As

$$A(\theta(\infty))\frac{\partial\eta(\theta^*)}{\partial\theta} = I_m$$

then from (20) it follows

$$\sqrt{N}(\theta(\infty)-\theta^*) = \sqrt{N}\rho^* + \sqrt{N}M_1 + \sqrt{N}M_2, \tag{21}$$

where

$$M_1 = \int_0^1 \frac{\partial A(\theta^* + z(\theta(\infty)-\theta^*))}{\partial\theta}\delta^* dz(\theta(\infty)-\theta^*)$$

$$M_2 = A(\theta(\infty))\int_0^1 \frac{\partial\eta(\theta^* + z(\theta(\infty)-\theta^*))}{\partial\theta}dz(\theta(\infty)-\theta^*).$$

According to Theorem 3 the random variable $\sqrt{N}(\theta(\infty)-\theta^*)$ is bounded in probability, and by Theorem 1

$$\frac{\partial A_N(\theta^*)}{\partial\theta} \xrightarrow{P} 0 \quad N \to \infty.$$

Then we obtain that

$$\sqrt{N}M_1 \xrightarrow{P} 0 \quad \text{for } N \to \infty.$$

By the same way

$$\sqrt{N}M_2 \xrightarrow{P} 0 \quad \text{for } N \to \infty.$$

Let's show that

$$\sqrt{N}\rho^* \to N(O, \Sigma(\theta^*)), \quad N \to \infty.$$

Remark 2. Matrix $\Sigma(\theta^*)$ is unknown. In 5[th] item we will evaluate the asymptotic estimators of elements of matrix $\Sigma(\theta^*)$.

Introduce non-zero vector of dimension m,

$$l = (l_1, l_2, \ldots, l_m)^T.$$

Consider

$$\sqrt{N}l^T\rho^* = \sqrt{N}l^T\left(\frac{F^T(\theta^*)F(\theta^*)}{N}\right)^{-1}\frac{F^T(\theta^*)}{N}(y - \eta(x,\theta^*)).$$

Then

$$\sqrt{N}l^T\rho^* = \frac{1}{N}\sum_{k=1}^N\sum_{l=1}^m\sum_{j=1}^m \sqrt{N}l_i x_{ij} f_{jk}\varepsilon_k$$

where f_k are the elements of matrix $F^T(\theta^*)$, $j=\overline{1,N}$; $k=\overline{1,m}$; x_{ij} are elements of $\left(\frac{F^T(\theta^*)F(\theta^*)}{N}\right)^{-1}$, $i,j=\overline{1,m}$.

Then

$$E\sqrt{N}l^T\rho^*=0$$

by virtue of condition (4).

Consider

$$E\left(\sqrt{N}l^T\rho^*\right)^2=\frac{1}{N}\sum_{i=1}^{N}\left(\sum_{i=1}^{m}\sum_{j=1}^{m}l_i x_{ij} f_{jk}\right)^2\sigma_i^2\le$$

$$\le\frac{\sigma_0^2}{N}\left(\frac{\sqrt{m}}{\lambda_1(\theta)}\right)^2C_1^2N\left(\sum_{i=1}^{m}l_iI_m\right)^2\le C_1^2\sigma_0^2\frac{m^3}{\lambda_1^2(\theta)}$$

it is bounded according to Theorem 4. According to central limit theorem, the random variable $\sqrt{N}l^T\rho^*$ has asymptotically Gaussian distribution

$$E\left\|\sqrt{N}l^T\rho^*\right\|=NEl^T\rho^*\rho^{*T}l=l^T\left(\frac{F^T(\theta^*)F(\theta^*)}{N}\right)^{-1}\times$$

$$\times\frac{F^T(\theta^*)I(\sigma^2)F(\theta^*)}{N}\left(\frac{F^T(\theta^*)F(\theta^*)}{N}\right)^{-1}l=l^T\Sigma(\theta^*)l.$$

According to criterion of normality [19, p.352] it follows that

$$\sqrt{N}\rho^*\xrightarrow{P}N(0,\Sigma(\theta^*)).$$

IV. Properties of M-estimator.

Remind that M-estimator $\overset{M}{\theta}$ of θ in regression model (I) was introduced by P.Houber and determined as a solution of the equation [20, p.51]

$$\min_{\theta\in\Theta}\sum_{i=1}^{N}\rho(y_i-\eta(x_i,\theta))=\sum_{i=1}^{N}\rho\left(y_i-\eta\left(x_i,\overset{M}{\theta}\right)\right),$$

where $\rho(x)$ is a some given function.

We will suppose that

$$\frac{d_\rho(x)}{dx}\Big/x=\psi(x)/x>0. \qquad (22)$$

Expand function $\eta(x,\theta)$ into Taylor series at the vicinity of point $\theta^*\in B(r)$ and consider

$$O=\sum_{i=1}^{N}\frac{\psi(y_i-\eta(x_i,\theta))}{y_i-\eta(x_i,\theta)}(y_i-\eta(x_i,\theta))-\sum_{i=1}^{N}\frac{\partial\eta(x_i,\theta^*)}{\partial\theta}\times$$
$$\times\frac{\psi(y_i-\eta(x_i,\theta))}{y_i-\eta(x_i,\theta)}(\theta^*-\theta)+o(\|\theta^*-\theta\|),\quad\theta\in B(r) \qquad (23)$$

Further we will suppose that r is enough small and we'll neglect by member of order $o(r)$.

Denote

$$\omega_{ii}(\theta)=\left[\frac{\psi(y_i-\eta(x_i,\theta))}{y_i-\eta(x_i,\theta)}\right]^{\frac{1}{2}},\quad \omega_{ij}=0,\quad i\neq j$$

$$W=W(\theta)_{ij},\quad i,j=\overline{1,N};\quad y_i(\theta)=\omega_{ii}(\theta)(y_i-\eta(x_i,\theta))$$

$$y(\theta)=(y_1(\theta),\ldots,y_N(\theta))^T;\quad g_{ij}(\theta)=\frac{\partial\eta(x_i,\theta)}{\partial\theta_j}\omega_{ii}(\theta);$$

$$i=\overline{1,N};\quad j=\overline{1,m};\quad G(\theta)=(g(\theta))_{ij}.$$

Then taking into account the relations (22), (23) can be rewritten in vector from

$$G^T(\theta)y(\theta)=\left(G^T(\theta)G(\theta)\right)\left(\overset{M}{\theta}-\theta^*\right). \tag{24}$$

But (24) is a normal equation for determining $\overset{M}{\theta}$. As θ^* is unknown then we obtain iteration procedure for $\overset{M}{\theta}$:

$$\overset{M}{\theta}_N(s+1)=\overset{M}{\theta}_N(s)+\left(G^T\theta(s)G\theta_N(s)\right)^{-1}G^T\theta_N(s)\left(y(\theta_N(s))-\eta(x,\theta_N(s))\right). \tag{25}$$

Iteration procedure (25) coincides with (5), if matrix $G(s)$ is replaced by matrix $F(\theta)$ and is required validity of the condition:

$$\omega_{ii}(\theta);\quad \frac{\partial\omega_{ii}(\theta)}{\partial\theta}\ \text{are bounded functions.}$$

Remark 3. *If $\rho(x)=x^2$, then $\omega_{ii}=2$; $G(\theta)=F(\theta)$ and therefore (25) coincides with (5), i.e. M - estimator coincides with l.s.e.*

Denote

$$0<\mu_1(\theta)\le\mu_2(\theta)\le\ldots\le\mu_m(\theta)$$

eigenvalues of matrix $\left[G^T(\theta)G(\theta)/N\right.$.

Determine

$$\hat{\rho}(\theta)=u(\hat{\theta})-\theta,$$

where

$$u(\hat{\theta})=\hat{\theta}+\left(G^T(\hat{\theta})G(\hat{\theta})\right)^{-1}G^T(\hat{\theta})\left(y-\eta(x,\hat{\theta})\right)$$

$$\hat{\tau}(r)=\sup_{\rho=1,\ldots,m}\sup_{B(r)}\left\|\frac{\partial u(\hat{\theta})}{\partial\theta}\right\|$$

Now we can carry these results for cases of M-estimators. Below we will formulate only basic results and we will indicate the analogous of which theorems they are.

Theorem 5. *Let* $\overset{M}{\theta}(O)\in B(r)$

$$\left[\sqrt{m}m^4/\left(N\mu_1^3(\theta)\right)\right]\left(1+\sum_{i\neq j}|r_{ij}|/N\right)\to 0\quad\text{for}\quad N\to\infty,\quad r\to 0$$

$$\left[m^5/\left(N\mu^4(\theta)\right)\right]\left(1+\sum_{i\neq j}|r_{ij}|/N\right)\to 0\quad\text{for}\quad N\to\infty,\quad r\to 0$$

Then there exists the random variable $\overset{M}{\theta}(\infty)$ such that

$$\left|\overset{M}{\theta}_N(\infty)-\overset{M}{\theta}_N(s)\right|\overset{P}{\to}0,\quad s\to\infty.$$

(The analogue of Theorem 2).

Theorem 6. *If*

$$\tau(r)+\frac{\|\rho^*\|}{r}<1,$$

then under conditions of theorem 5 the random variable $\sqrt{N}\left(\overset{M}{\theta}(\infty)-\theta^*\right)$ *is bounded in probability for* $N\to 0$.

(The analogue of Theorem 3).

Theorem 7. *If the random variable* $\sqrt{N}\left(\overset{M}{\theta}(\infty)-\theta^*\right)$ *is bounded in probability, then under conditions of theorem 6*

$$\sqrt{N}\left(\overset{M}{\theta}_N(\infty)-\theta^*\right)\to N\left(O,\hat{\Sigma}(\theta^*)\right),\quad N\to\infty.$$

V. Estimation of elements of covariance matrix and construction of a confident band.

Consider covariance matrix of the vector $\sqrt{N}\left(\hat{\theta}-\theta^*\right)$, which is denoted by

$$C_N=NE\left(\hat{\theta}-\theta^*\right)\left(\hat{\theta}-\theta^*\right)^T,\ c_{kl}\text{ are elements of }C_N,\quad k,l=\overline{1,m}.$$

From the relation (21) it follows that

$$\sqrt{N}\left(\hat{\theta}-\theta^*\right)=\sqrt{N}\rho^*+\sqrt{N}M_1+\sqrt{N}M_2$$

where $\sqrt{N}M_1$ and $\sqrt{N}M_2$ tends to zero for $N\to\infty$, and the random variable $\sqrt{N}\rho^*$ is bounded in probability. Therefore suppressing members that tends to zero for $N\to\infty$ for matrix C_N we can write:

$$C_N NE^*\left(\rho^*\right)^T=\left(\frac{F^T(\theta^*)F(\theta^*)}{N}\right)^{-1}\frac{F^T(\theta^*)I(\sigma^2)F(\theta^*)}{N}\left(\frac{F^T(\theta^*)F(\theta^*)}{N}\right)^{-1}$$

where $I\left(\sigma^2\right)=E\varepsilon\varepsilon^T$.

The value θ^* is unknown, but by Ttheorem 5 it is known $\sqrt{N}$-consistent estimator. For the elements of the matrix $\left[F^T\left(\hat{\theta}\right)I\left(\sigma^2\right)F\left(\hat{\theta}\right)/N\right]$ according to Theorem 1 there exist asymptotically unbiased and consistent estimators.

Thus, if conditions of theorems 2 and 4 hold, then we can estimate the elements of covariance matrix C_N. Denote the elements of matrix $B=\dfrac{F^T\left(\hat{\theta}\right)I\left(\sigma^2\right)F\left(\hat{\theta}\right)}{N}$ by b_{kl}. Let

$O<\hat{\lambda}_1\le\hat{\lambda}_2\le\ldots\le\hat{\lambda}_m$ be the eigenvalues of matrix $\dfrac{F^T\left(\hat{\theta}\right)F\left(\hat{\theta}\right)}{N}$.

$\hat{y}_i=\eta\left(x_i,\hat{\theta}\right),\quad \hat{I}_{kl}\left(\hat{\theta}\right)=\delta_{ij}f_{ki}\left(\hat{\theta}\right)f_{ij}\left(\hat{\theta}\right)$

$\hat{y}=\left(\hat{y}_{12},\hat{y}_2,\ldots,\hat{y}_N\right)^T$

$\hat{b}_{kl}=\left(y-\hat{y}\right)^T\hat{I}_{kl}\left(\hat{\theta}\right)\left(y-\hat{y}\right)/N$

$\hat{B}=\left(\hat{b}\right)_{kl};\quad k,l=\overline{1,m}$

$$\hat{C}_N = \left(\frac{\hat{F}^T \hat{F}}{N}\right)^{-1} \hat{B} \left(\frac{\hat{F}^T \hat{F}}{N}\right)^{-1}$$

$\hat{c}_{kl}$ are elements of $\hat{C}_N$.

Now we can use theorem 2 from [12] requiring validity of its conditions.

Theorem 8. *Let*

$$\begin{aligned} m\sqrt{m}\big/\left(N\hat{\lambda}_1\right) &\to 0 \quad \text{for} \quad N \to \infty \\ m\sqrt{m}\big/\left(N\hat{\lambda}_1^2\right) & \quad \text{is bounded} \end{aligned} \tag{26}$$

Then

$$E(\hat{c}_{kl} - c_{kl}) \to 0 \quad (\hat{c}_{kl} - c_{kl}) \xrightarrow{P} 0 \quad \text{for} \quad N \to \infty.$$

Proof. As all conditions of Theorem 2 are hold, then

$$E\left(\hat{b}_{kl} - b_{kl}\right) \to 0 \quad \left(\hat{b}_{kl} - b_{kl}\right) \xrightarrow{P} 0 \quad \text{for} \quad N \to \infty.$$

According to Theorem 6 we have

$$\left(\hat{\theta} - \theta^*\right) \xrightarrow{P} 0$$

and as all elements of matrix F are continuous functions, then passing to the limit and taking into account condition (26), we obtain

$$E\left(\hat{c}_{kl} - c_{kl}\right) \to 0 \quad \left(\hat{c}_{kl} - c_{kl}\right) \xrightarrow{P} 0 \quad \text{for} \quad N \to \infty.$$

Theorem 8 is an analogue of Theorem 2 and describes case of noncorrelated random variables $\varepsilon_i, \varepsilon_j, i \neq j$. Denote by $E\varepsilon_i\varepsilon_j = r_{ij}$, then the following theorem is true.

Theorem 9. *Let*

$$\frac{m\sqrt{m}}{\hat{\lambda}_1}\left(\sigma_0^2 + \sum_{i \neq j} \left|r_{ij}\right| \big/ N\right) \to 0 \quad \text{for} \quad N \to \infty$$

$$\frac{1}{N^2} \sum_{1 \le i_1 \le i_2 \le i_3 \le i_4 \le N} E\left(\varepsilon_{i_1}\varepsilon_{i_2}\varepsilon_{i_3}\varepsilon_{i_4} - r_{i_1 i_2} r_{i_3 i_4}\right) \to 0 \quad \text{for} \quad N \to \infty$$

Then

$$E(\hat{c}_{kl} - c_{kl}) \to 0 \quad (\hat{c}_{kl} - c_{kl}) \xrightarrow{P} 0 \quad \text{for} \quad N \to \infty.$$

It is obvious that for $r_{ij} = 0$ *the conditions of Theorem 8 and 9 coincide.*

Estimating elements of covariance matrix C_N and using statement of Theorem 7 we obtain, that quadratic form $\left(\hat{\theta} - \theta^*\right)^T \hat{C}_N^{-1} \left(\hat{\theta} - \theta^*\right)$ has asymptotic *chi*-square distribution with m degree of freedom. We can construct confidence ellipsoid and using this ellipsoid we can determine low and upper bounds of confidence band for function $\eta(x)$.

Note that mentioned-above results for estimating of elements of covariance matrix easily applies to the cases of M-estimator.

Author express thanks to prof. Jn.K.Belyaev and M.B.Malyutov for discussing and useful advices .

References

[1]. Anderson *T.W.*, Taylor J.B. *Strong consistency of least squares estimates in normal linear regression //Ann. statist. - 1976, 4, -p. 788-790.*

[2]. Drygas H. *BLUE in linear regression models.* //Probab. mathem. statist. -1985, Fasc. 2. -P,

177-186.

[3]. Ching-Shui Cheng, Ker-Chan Li. *The strong consistency of M-estimators in linear models.* //J. Multiv. Anal. -1984, 15. -P. 91-98.

[4]. Chen Chi-Zing, lai T.L., Wei C.Z. *Convergence systems and strong consistency of least squares estimates in regression models.* //J. Multiv. Anal. -1981, 11. -P. 319-333.

[5]. Novikov A.A. *Consistency of least squares estimates in regression model with martingals.* //In: Statistical control of random processes. Proc. Steklov Instit. Seminars. N-Y, Optimization Software. -1985, v.1. -P.79-93.

[6]. Гирко В.Л. *Введение в общий статистический анализ.* //Теор. Вер. Примен. -1987, 32, вып. 2. -С. 252-265.

[7]. Portnoy S. *Asymptotic behavior of M-estimators of p-regression parameters when p/n is large. I Consistency* //Ann. Stat. -1984, 12, -P. 1298-1309.

[8]. Pousse A. *Extensions and asymptotics studies in factor analyses.* //В сб. Всемирный конгресс общества матем. стат. и теории вероятн. им. Бернулли, Т.И. -М. Наука. -1986. -P. 200-202.

[9]. Jenrich R.I. *Asymptotic properties of non-linear least squares estimators.* //Ann. Math. Statist. -1969, 40, -P. 633-643.

[10]. Malinvaud E. *The consistency of non-linear reqressions.* //Ann. Math. Statist. -1970, 41. -P 956-969.

[11]. Pazman A. *Some statistical consequences of the geometry of nonlinear regression.* //Proce. 5 Pannovion Symp. Math. Stat. Viseqrad, Hungary. -1985. -P. 273-283.

[12]. Гаджиев А.Г. *Статистические оценки в линейных регрессионных моделях с зависимыми ошибками наблюдений.* //Теор. вер. примен. -1989. т. 34. вып. 2. -С. 764- 768.

[13]. Гаджиев А.Г. *Статистические оценки в регрессионных моделях* //ДАН СССР. -1990. т.313, №2. -С. 265-268.

[14]. Демиденко Е.З. *Линейная и нелинейная регрессия.* М., Финансы и статистика, 1981, 301 с.

[15]. Малютов М.Б. *Нижние границы для средней длительности последовательно планируемого експеримента.* //Изв. Вузов. Математика. -1983, 11. -С. 19-41.

[16]. Хорн Р., Джонсон И. *Матричный анализ.* М.: Мир, 1989. 654 с

[17]. Крылов В.И., Бобков В.В., Монастырский П.И. *Начало теории вычислительных методов.* -Минск, Наука и техника, 1985. -277 с.

[18]. Колмогоров А.Н., Фомин С.В. *Элементы теории функций и функционального анализа.* -М.: Наука, 1981. -542 с.

[19]. Рао С.Р. *Линейные статистические методы и их применения.* -М.: Наука, 1968. -547с.

[20]. Хьюбер П. *Робастность в статистике.* -М.: Мир. 1984. -304 с.

[21]. Гаджиев А.Г. *Об оценивании в линейных регрессионных моделях.* //Вестик МГУ, сер. матем. мех. -1989, №1, -с. 8-13.

Hajiyev A.H.
Institute Cybernetics of AS Azerbaijan.
9, F. Agayev str., 370141, Baku, Azerbaijan.

Received May 20, 1999; Revised July 9, 1999.
Translated by Panarina V.K.