33

PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 10, Fasc. 1 (1989), p. 1–10

CONSISTENCY OF M-ESTIMATORS IN A LINEAR MODEL, GENERATED BY NON-MONOTONE AND DISCONTINUOUS ψ-FUNCTIONS

BY

JANA JUREČKOVÁ (PRAGUE)

Abstract. Existence of a $\sqrt{n}$-consistent M-estimator of regression parameter vector and its asymptotic normality are proved in a general situation including redescending estimators generated by possibly discontinuous ψ-functions.

1. Introduction. Consider the linear regression model

$$Y_n = X_n \beta + E_n, \tag{1.1}$$

where Y_n and E_n are $(n \times 1)$ random vectors, X_n is an $(n \times p)$ design matrix and the coordinates $E_1, \ldots, E_n$ of E_n are independent and identically distributed random variables with a distribution function (d.f.) F. Let x_i' denote the ith row of X_n, $i = 1, \ldots, n$.

M-estimators of $\boldsymbol{\beta}$ became a part of a general statistical consciousness. For a function ψ: $R^1 \to R^1$ such that

$$\int \psi(x)\, dF(x) = 0 \tag{1.2}$$

the M-estimator M_n of $\boldsymbol{\beta}$ is usually defined as a solution of the system of equations

$$\sum_{i=1}^{n} x_i \psi(Y_i - x_i' t) = O \tag{1.3}$$

with respect to $t \in R^p$.

The asymptotic behavior of M_n (as $n \to \infty$) has been studied by many authors. The basic references can be found e.g. in Huber [4] and Hampel et al. [2]. The question of primary interest is that of conditions under which there exists a solution M_n of (1.3) for which

$$n^{1/2} \|M_n - \boldsymbol{\beta}\| = O_p(1) \quad \text{as } n \to \infty, \tag{1.4}$$

and if this is the case then there is a question of the asymptotic distribution of $n^{1/2}(\boldsymbol{M}_n - \boldsymbol{\beta})$.

If ψ is nondecreasing, then such an $\boldsymbol{M}_n$ exists and is unique in probability under general conditions. This question was studied, among others, by Huber [3] and Yohai and Maronna [7] and not only for fixed p but also when p is permitted to grow as $n \to \infty$.

If ψ is not monotone, there could be more solutions of the system (1.3) and some of them possibly inconsistent; Freedman and Diaconis [1] found a family of distributions with inconsistent M-estimators of location. For an appropriate family of distributions, there typically exists at least one root of (1.3) which satisfies (1.4). For sufficiently smooth ψ and under some restrictions on the design matrix $\boldsymbol{X}_n$, this was demonstrated by Portnoy [6] even in the case where $p \to \infty$ and (p. $\log n)/n \to 0$ as $n \to \infty$. He proved it with the aid of results of the general theory of nonlinear equations in several variables.

However, there are still other open questions: (i) what could we say in the case of non-monotone ψ if Portnoy's conditions on $\boldsymbol{X}_n$ are not satisfied and/or ψ is not continuous; (ii) even if we know that there exists a $\sqrt{n}$-consistent solution of (1.3), then, having found one single root, we do not know whether it is just the consistent one.

Some authors recommend using redescending M-estimators generated by the class of functions

$$\Psi = \{\psi : \psi(x) = 0 \text{ for all } |x| \geqslant r,\ r \in R^1\};$$

the main motivation is that these estimators are able to reject extreme outliers entirely (cf. [2]). As examples of such functions we could mention "Tukey's biweight"

$$\psi(x) = x(r^2 - x^2)^2 I[-r \leqslant x \leqslant r], \tag{1.5}$$

further the function generating the "skipped mean" in the location model,

$$\psi(x) = \begin{cases} x & \text{for } |x| \leqslant r, \\ 0 & \text{for } |x| > r \end{cases} \tag{1.6}$$

or the function generating the "skipped median" in the location model

$$\psi(x) = \begin{cases} \operatorname{sign} x & \text{for } |x| \leqslant r, \\ 0 & \text{for } |x| > r. \end{cases} \tag{1.7}$$

Another class is formed by functions which are non-monotone and tend to 0 as $|x| \to \infty$. Such are, e.g., the log-likelihood functions of non-unimodal densities.

This all means that the existence of a consistent M-estimator generated by a non-monotone function deserves a more detailed study. Surprisingly, this question was not yet satisfactorily solved; functions (1.6) and (1.7) do not even comply with Portnoy's conditions.

In the case of non-monotone ψ, it is better to define the M-estimator as a solution of the minimization problem

$$\sum_{i=1}^{n} \varrho(Y_i - x_i' t) := \min \tag{1.8}$$

with respect to $t \in R^p$, where ϱ is an absolutely continuous function and $\psi(x) = d\varrho/dx$ a.e. This is in analogy with the definition of the least-squares estimator. If ϱ is convex, then (1.8) is equivalent to (1.3).

It is the aim of the present study to show that there exists a solution of (1.8) satisfying (1.4) for a general class of ϱ-functions including nonconvex ones with discontinuous derivatives. We shall also prove the asymptotic normality of this solution and its asymptotic representation by a sum of i.i.d. random variables.

2. Assumptions and auxiliary asymptotic linearity result. We impose the following conditions on ϱ, F and on the design matrix X_n:

(A1) ϱ: $R^1 \to R^1$ is an absolutely continuous function, bounded from below and such that the function

$$\lambda(u) = \int \varrho(x-u)\, dF(x) \tag{2.1}$$

has a unique minimum at $u = 0$.

(A2) The function $\psi(x) = d\varrho/dx$ has the form

$$\psi = \psi_1 + \psi_2, \tag{2.2}$$

where ψ_1 is a step function,

$$\psi_1(x) = \begin{cases} \alpha_j & \text{for } s_j < x < s_{j+1},\, j = 0, \ldots, k, \\ \frac{1}{2}(\alpha_{j-1} + \alpha_j) & \text{for } x = s_j,\, j = 1, \ldots, k, \end{cases} \tag{2.3}$$

where $-\infty = s_0 < s_1 < \ldots < s_k < s_{k+1} = \infty$ and $\alpha_0, \ldots, \alpha_k \in R^1$, not all equal; ψ_2 is a sum of two continuous functions, say, $\psi_2 = \psi_2^{(1)} + \psi_2^{(2)}$, such that
$d\psi_2^{(1)}/dx$ is a step-function and $d\psi_2^{(2)}/dx$ absolutely continuous and

$$\int \left(\psi_2^{(i)}(x+u+v) - \psi_2^{(i)}(x+u)\right)^2 dF(x) \leqslant K_1 v^2 \quad (i = 1, 2) \tag{2.4}$$

for $|u| \leqslant \delta$, $|v| \leqslant \delta$, K_1, $\delta > 0$.

(A3) There exist $\gamma_i = \gamma_i(\psi, F)$, $i = 1, 2$, such that

$$\int (\psi_i(x+h) - \psi_i(x))\, dF(x) = h\gamma_i + o(h) \quad \text{as } h \to 0, \tag{2.5}$$

and $\gamma = \gamma_1 + \gamma_2 > 0$.

(A4) F has a positive and bounded derivative f in a neighborhood of $s_1, \ldots, s_k$.

(B1) $\lim_{n\to\infty} Q_n = Q$, where $Q_n = n^{-1} X_n' X_n$ and Q is a positively definite $(p \times p)$-matrix.

(B2) $a_n = \max_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} |x_{ij}| = o(n^{1/2})$ as $n \to \infty$.

(B3) $b_n = \max_{1 \leqslant j \leqslant p} (n^{-1} \sum_{i=1}^{n} x_{ij}^4) = O(1)$ as $n \to \infty$.

We shall start with the following uniform asymptotic linearity result:

LEMMA 2.1. *Let $E_1, E_2, \ldots$ be i.i.d. random variables with the distribution function F. Let ψ: $R^1 \to R^1$ be a function satisfying conditions* (A2) *and* (A3), *let F satisfy* (A4) *and the matrix X_n satisfy* (B1)-(B3). *Then, for any fixed $\tau \leqslant 1/2$ and $C > 0$,*

$$\sup_{\|t\| < C} |n^{-1+\tau} \sum_{i=1}^{n} x_{ij} [\psi(E_i - n^{-\tau} x_i' t) - \psi(E_i) + n^{-\tau} \gamma x_i' t]| \xrightarrow{p} 0 \tag{2.6}$$

as $n \to \infty$, $j = 1, \ldots, p$.

Proof. Write

$$S_j(t) = n^{-1+\tau} \sum_{i=1}^{n} x_{ij} [\psi(E_i - n^{-\tau} x_i' t) - \psi(E_i)] \tag{2.7}$$

and

$$S_j^0(t) = S_j(t) - \mathrm{E} S_j(t), \quad j = 1, \ldots, p;\ t \in R^p. \tag{2.8}$$

Let first $\psi \equiv \psi_1$ (the step-function of (2.3)). Then, for t, $u \in R^p$, $t \leqslant u$ (coordinatewise),

$$\begin{aligned} &\mathrm{E}[S_j^0(u) - S_j^0(t)]^4 \\ &\leqslant n^{-4+4\tau} \{11 \sum_{i=1}^{n} x_{ij}^4 \sum_{\nu=1}^{k} (\alpha_\nu - \alpha_{\nu-1})^4 |F(s_\nu + n^{-\tau} x_i' u) - F(s_\nu + n^{-\tau} x_i' t)| + \\ &\quad + [\sum_{i=1}^{n} x_{ij}^2 \sum_{\nu=1}^{k} (\alpha_\nu - \alpha_{\nu-1})^2 |F(s_\nu + n^{-\tau} x_i' u) - F(s_\nu + n^{-\tau} x_i' t)|]^2\} \\ &\leqslant \|u - t\|\, O(n^{-3+3\tau} a_n) + \|u - t\|^2\, O(n^{-2+2\tau} a_n^2), \end{aligned} \tag{2.9}$$

hence

$$\overline{\lim_{n\to\infty}}\, a_n^{-2} n^{2-2\tau} \mathrm{E}[S_j^0(u) - S_j^0(t)]^4 \leqslant K \|u - t\|^2. \tag{2.10}$$

Moreover, by (A3),

$$(2.11)\qquad |\mathrm{E}\,[S_j(\boldsymbol{u})-S_j(\boldsymbol{t})+\gamma n^{-1}\sum_{i=1}^{n} x_{ij}\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})]|$$

$$= |\mathrm{E}\,\{n^{-1+\tau}\sum_{i=1}^{n} x_{ij}[\psi(E_i-n^{-\tau}\boldsymbol{x}_i'\boldsymbol{u})-\psi(E_i-n^{-\tau}\boldsymbol{x}_i'\boldsymbol{t})+\gamma n^{-\tau}\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})]\}|$$

$$= o\big(n^{-1}\sum_{i=1}^{n}|x_{ij}\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})|\big) = \|\boldsymbol{u}-\boldsymbol{t}\|\,o(1) \quad \text{as } n\to\infty.$$

Combining (2.10) and (2.11), we get

$$(2.12)\qquad \mathrm{E}\,[S_j(\boldsymbol{u})-S_j(\boldsymbol{t})+\gamma n^{-1}\sum_{i=1}^{n} x_{ij}\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})]^4 \leqslant \|\boldsymbol{u}-\boldsymbol{t}\|^2\,o(1)$$

$$\text{as } n\to\infty \text{ for } \boldsymbol{t},\ \boldsymbol{u}\in R^p,\ \boldsymbol{t}\leqslant \boldsymbol{u},\ \|\boldsymbol{t}\|,\ \|\boldsymbol{u}\|\leqslant C.$$

Thus, by results of [5], the proposition (2.6) holds if $\psi \equiv \psi_1$. Let now $\psi \equiv \psi_2$. Write

$$(2.13)\qquad A_{ni}(\boldsymbol{t}) = \psi(E_i-n^{-\tau}\boldsymbol{x}_i'\boldsymbol{t})-\psi(E_i), \quad i=1,\ldots,n.$$

Then, by (2.5),

$$(2.14)\qquad \mathrm{E}\,[A_{ni}(\boldsymbol{u})-A_{ni}(\boldsymbol{t})]^2 \leqslant K_1 n^{-2\tau}\big(\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})\big)^2$$

for both $\psi = \psi_2^{(\nu)}\ (\nu = 1, 2)$ and

$$(2.15)\quad \mathrm{var}\big(S_j(\boldsymbol{u})-S_j(\boldsymbol{t})\big) \leqslant K_1 n^{-2}\sum_{i=1}^{n} x_{ij}^2\big(\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})\big)^2 \leqslant K_1\|\boldsymbol{u}-\boldsymbol{t}\|^2\,O(n^{-1}).$$

Moreover, by (A.3),

$$(2.16)\qquad |\mathrm{E}\,[S_j(\boldsymbol{u})-S_j(\boldsymbol{t})+n^{-1}\gamma\sum_{i=1}^{n} x_{ij}\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})]| \leqslant \|\boldsymbol{u}-\boldsymbol{t}\|\,o(1),$$

hence

$$(2.17)\qquad \mathrm{E}\,[S_j(\boldsymbol{u})-S_j(\boldsymbol{t})+n^{-1}\gamma\sum_{i=1}^{n} x_{ij}\boldsymbol{x}_i'(\boldsymbol{u}-\boldsymbol{t})]^2 \leqslant \|\boldsymbol{u}-\boldsymbol{t}\|^2\,o(1)$$

and, by results of [5], (2.6) holds for $\psi \equiv \psi_2$.

COROLLARY 2.1. *Let $E_1, E_2, \ldots$ be i.i.d. random variables with the distribution function F. Let ϱ: $R^1 \to R^1$, $\psi = \varrho'$ and F satisfy conditions* (A1)-(A4) *and* *let X_n satisfy* (B1)-(B3). *Then, for any fixed $\tau \leqslant 1/2$ and $C > 0$,*

$$\text{(2.18)} \qquad \sup_{\|t\| \leqslant C} \Big| n^{-1+2\tau} \sum_{i=1}^{n} [\varrho(E_i - n^{-\tau} x_i' t) - \varrho(E_i) + n^{-\tau} x_i' t \psi(E_i)] - $$

$$ - (\gamma/2) t' Q_n t \Big| \xrightarrow{P} 0 \quad as \quad n \to \infty.$$

Proof. By Lemma 2.1, if $t_j \geqslant 0$,

$$\text{(2.19)} \qquad \sup_{|t_k| \leqslant C, k=j,\dots,p} \Big| n^{-1+\tau} \int_0^{t_j} \sum_{i=1}^{n} x_{ij} \{\psi(E_i - n^{-\tau} [x_{ij} s +$$

$$+ \sum_{k=j+1}^{p} x_{ik} t_k]) - \psi(E_i) + n^{-\tau} \gamma (x_{ij} s + \sum_{k=j+1}^{p} x_{ik} t_k)\} ds \Big| \xrightarrow{P} 0 \quad \text{as } n \to \infty,$$

and analogously if $t_j < 0$ with the reversed order of integration, $j = 1, \dots, p$. This, in turn, implies (2.18).

3. Consistency and asymptotic normality. Assume that the functions ϱ and F satisfy conditions (A1)-(A4) and X_n satisfies conditions (B1)-(B2) introduced in Section 2. If M_n minimizes $\sum \varrho(Y_i - x_i' t)$, $i = 1, 2, \dots, n$, then

$$\text{(3.1)} \qquad T_n = M_n - \beta$$

also minimizes $\sum (\varrho(E_i - x_i' t) - \varrho(E_i))$, $i = 1, 2, \dots, n$, with respect to $t \in R^p$. By Lemma 2.1, for any fixed $C > 0$,

$$\text{(3.2)} \qquad \min_{\|t\| \leqslant C} \sum_{i=1}^{n} [\varrho(E_i - n^{-1/2} x_i' t) - \varrho(E_i)]$$

$$= \min_{\|t\| \leqslant C} (-t' Z_n + \frac{\gamma}{2} t' Q_n t) + o_p(1) \quad \text{as } n \to \infty,$$

where

$$\text{(3.3)} \qquad Z_n = n^{-1/2} \sum_{i=1}^{n} x_i \psi(E_i) \ \big(= O_p(1)\big)$$

is a random vector of R^p. By (3.2), the minimum of

$$\sum_{i=1}^{n} [\varrho(E_i - n^{-1/2} x_i' t) - \varrho(E_i)]$$

over the sphere $\|t\| \leqslant C$ can be approximated by a minimum of a convex function over the same sphere. However, the minimum of the convex function on the right-hand side of (3.2) over $\|t\| \leqslant C$ in turn coincides with the minimum of the same function over all R^p, provided only $C > 0$ is sufficiently large. Actually, we have

LEMMA 3.1. *Under the conditions of Lemma* 2.1, *given* $\varepsilon > 0$, there exist $C_0 > 0$ *and* n_0 *such that, for* $n \geqslant n_0$ *and* $C \geqslant C_0$,

$$P(B_n(C, \varepsilon)) < \varepsilon, \tag{3.4}$$

where

$$B_n(C, \varepsilon) = \{\omega: |\min_{\|t\| \leqslant C} \sum_{i=1}^{n} [\varrho(E_i - n^{-1/2} x_i' t) - \varrho(E_i)] - \min_{t \in R^p} [-t' Z_n + \frac{\gamma}{2} t' Q_n t]| > \varepsilon\}. \tag{3.5}$$

Proof. Let U_n be the solution of the minimization

$$\min_{t \in R^p} [-t' Z_n + (\gamma/2) t' Q_n t].$$

Then

$$U_n = \gamma^{-1} Q_n^{-1} Z_n (= O_p(1)) \tag{3.6}$$

and hence there exist $C_0 > 0$ and n_0 so that $P(\|U_n\| > C_0) < \varepsilon/2$ for $n \geqslant n_0$. Then, regarding (3.2) and (3.5), for $C \geqslant C_0$ and $n \geqslant n_0$ we have

$$P(B_n(C, \varepsilon)) \leqslant P(B_n(C, \varepsilon) \cap [\|U_n\| \leqslant C]) + (\varepsilon/2)$$
$$\leqslant P\{|\min_{\|t\| \leqslant C} \sum_{i=1}^{n} [\varrho(E_i - n^{-1/2} x_i' t) - \varrho(E_i)] - \min_{\|t\| \leqslant C} (-t' Z_n + (\gamma/2) t' Q_n t)| > \varepsilon\} + (\varepsilon/2) \leqslant \varepsilon.$$

The function

$$g_n(t) = -t' Z_n + (\gamma/2) t' Q_n t \tag{3.7}$$

has a unique minimum over R^p equal to $(-1/2\gamma) Z_n' Q_n^{-1} Z_n$ which is negative with probability 1 starting from some n_0. Hence, by Lemma 3.1,

$$\min_{\|t\| \leqslant C} \sum_{i=1}^{n} [\varrho(E_i - n^{-1/2} x_i' t) - \varrho(E_i)] = (-1/2\gamma) Z_n' Q_n^{-1} Z_n + o_p(1). \tag{3.8}$$

Moreover, the sequence $Z_n' Q_n^{-1} Z_n \sigma_\psi^{-1}$ has asymptotically the χ^2-distribution with p degrees of freedom, where $\sigma_\psi^2 = \int \psi^2(x) dF(x)$; thus, by (3.8), there exist $\delta > 0$ and n_0 to given $\varepsilon > 0$ such that, for $n \geqslant n_0$,

$$P\{\min_{\|t\| \leqslant C} \sum_{i=1}^{n} [\varrho(E_i - n^{-1/2} x_i' t) - \varrho(E_i)] \leqslant -\delta\} > 1 - \varepsilon. \tag{3.9}$$

Now, assume that M_n is not a $\sqrt{n}$-consistent estimator of β; then there exists a $\tau < 1/2$ such that

(3.10) $$\|T_n\| = \|M_n - \beta\| = O_p(n^{-\tau}) \text{ but } \|T_n\| \neq o_p(n^{-\tau})$$

as $n \to \infty$. Then, by Lemma 2.1,

(3.11) $$n^{-1+2\tau} \sum_{i=1}^{n} [\varrho(E_i - x_i' T_n) - \varrho(E_i)]$$
$$= (n^\tau T_n') Q_n (n^\tau T_n) - n^{-(1/2)+\tau} (n^\tau T_n') Z_n + o_p(1)$$
$$= (n^\tau T_n') Q_n (n^\tau T_n) + o_p(1).$$

The sequence $\{H_n\}$ of distribution functions of $(n^\tau T_n') Q_n (n^\tau T_n)$ contains a convergent subsequence $\{H_{n_k}\}$ which converges to a nondegenerate d.f. H, concentrated on the positive half-axis. Hence, given an $\varepsilon > 0$, there exist $\delta_1 > 0$ and an integer k_1 such that

(3.12) $$P\Big\{\sum_{i=1}^{n_k} [\varrho(E_i - x_i' T_{n_k}) - \varrho(E_i)] \geqslant n_k^{1-2\tau} \delta_1\Big\} > 1 - \varepsilon \quad \text{for } k \geqslant k_1.$$

Notice that the lower bound $n_k^{1-2\tau}\delta_1$ in (3.12) is unbounded as $k \to \infty$. This implies that the function $\sum (\varrho(E_i - x_i' t) - \varrho(E_i))$, $i = 1, 2, \ldots, n$, to be minimized could take on positive unbounded values with probability arbitrarily close to 1 if $t = T_n$; while, by (3.9), the minimum of the same function over the sphere $\|t\| \leqslant n^{-1/2} C$ is negative with probability arbitrarily close to 1. Thus, $\tau < 1/2$ cannot be true; this means that $\tau = 1/2$ and $\|T_n\| = O_p(n^{-1/2})$.

We are in a position to formulate the main theorem of the paper.

THEOREM 3.2. *Let $Y_1, \ldots, Y_n$ be independent random variables, Y_i distributed according to the d.f. $F(y - x_i' t)$, $i = 1, \ldots, n$. Let M_n be the point of global minimum of $\sum \varrho(Y_i - x_i' t)$, $i = 1, 2, \ldots, n$, with respect to $t \in R^p$, where the functions ϱ, F and the matrix $X_n = (x_1, \ldots, x_n)'$ satisfy conditions* (A1)-(A4) *and* (B1)-(B2). *Then*

(3.13) $$n^{1/2} \|M_n - \beta\| = O_p(1),$$

(3.14) $$n^{1/2}(M_n - \beta) = n^{-1/2} \gamma^{-1} Q^{-1} \sum_{i=1}^{n} x_i \psi(E_i) + o_p(1) \quad \text{as } n \to \infty,$$

and $n^{1/2}(M_n - \beta)$ has asymptotically p-dimensional normal distribution

(3.15) $$N_p(O, (\sigma_\psi^2/\gamma^2) Q^{-1}).$$

Proof. (3.13) follows from the above considerations. To prove (3.14) and (3.15), it suffices to prove that

(3.16) $$\|n^{1/2} T_n - U_n\| \xrightarrow{P} 0 \quad \text{as } n \to \infty$$

for T_n of (3.1) and U_n of (3.6). Because the function $g_n(t)$ is convex and has a unique minimum at $t = U_n$ with probability 1, we get

$$(3.17)\qquad g_n(t)-g_n(U_n) = \frac{\gamma}{2}(t-U_n)' Q_n(t-U_n) \geqslant \frac{\gamma}{2}\|t-U_n\|^2 \lambda_n^0 \quad \text{for } t \in R^p,$$

where λ_n^0 is the minimal eigenvalue of Q_n. Hence, given $\eta > 0$, there exist $\delta > 0$ and n_0 so that, for $n \geqslant n_0$,

$$(3.18)\qquad P(\|n^{1/2} T_n - U_n\| > \eta) \leqslant P\big(g_n(n^{1/2} T_n) - g_n(U_n) > \delta\big).$$

Being combined with Lemma 2.1 this further implies that, to $\eta > 0$ and $\varepsilon > 0$, there exists an n_1 such that, for $n \geqslant n_1$,

$$\begin{aligned} &P(\|n^{1/2} T_n - U_n\| > \eta) \\ &\leqslant P\Big\{g_n(n^{1/2} T_n) - g_n(U_n) > \delta, \Big|\sum_{i=1}^{n} [\varrho(E_i - n^{-1/2} x_i' U_n) - \varrho(E_i)] - \\ &\qquad - g_n(U_n)\Big| \leqslant \frac{\delta}{4}, \Big|\sum_{i=1}^{n} [\varrho(E_i - x_i' T_n) - \varrho(E_i)] - g_n(n^{1/2} T_n)\Big| \leqslant \frac{\delta}{4}\Big\} + \varepsilon \\ &\leqslant P\Big\{\frac{\delta}{2} < \sum_{i=1}^{n} [\varrho(E_i - x_i' T_n) - \varrho(E_i - n^{-1/2} x_i' U_n)]\Big\} + \varepsilon = \varepsilon. \end{aligned}$$

Propositions (3.13) and (3.14) then follow from (3.3), (3.6) and from assumption (A1).

Remark. The function ϱ does not need to be necessarily convex; however, conditions (2.1) of (A1) and $\gamma > 0$ of (A3) are crucial for the existence of a $\sqrt{n}$-consistent solution of the minimization (1.8).

REFERENCES

[1] D. A. Freedman and P. Diaconis, *On inconsistent M-estimators*, Ann. Statist. 10 (1982), p. 454–461.

[2] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Function*, John Wiley, New York 1986.

[3] P. J. Huber, *Robust regression: Asymptotics, conjectures and Monte Carlo*, Ann. Statist. 1 (1973), p. 799–821.

[4] – *Robust Statistics*, John Wiley, New York 1981.

[5] J. Jurečková and P. K. Sen, *Uniform second order asymptotic linearity of M-statistics in linear models* (to appear).

[6] S. Portnoy, *Asymptotic behavior of M-estimators of p regression parameters when p^2/n is large. I. Consistency*, Ann. Statist. 12 (1984), p. 1298–1309.

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

Charles University
Department of Probability and Statistics
Sokolovská 83
18600 Prague 8, Czechoslovakia

Received on 1. 7. 1987