KYBERNETIKA — VOLUME 42 (2006), NUMBER 2, PAGES 181 – 202

THE LEAST TRIMMED SQUARES
Part II: $\sqrt{n}$-consistency

Jan Ámos Víšek

$\sqrt{n}$-consistency of the least trimmed squares estimator is proved under general conditions. The proof is based on deriving the asymptotic linearity of normal equations.

Keywords: robust regression, the least trimmed squares, $\sqrt{n}$-consistency, asymptotic normality

AMS Subject Classification: 62J05, 62F35, 62F12

INTRODUCTION AND NOTATIONS

The paper is a continuation of [4]. That is why only brief introduction of notations will be given. For discussion of the definitions and assumptions see Part I.

Let N denote the set of all positive integers, R the real line and R^p the p-dimensional Euclidean space. Moreover, for any set A let A^o denote the interior of the set (in the topology implied by Euclidean metric). We shall consider for any $n \in N$ the linear regression model

$$Y_i = x_i^{\mathrm{T}} \beta^0 + e_i, \quad i = 1, 2, \ldots, n \tag{1}$$

where Y_i and $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^{\mathrm{T}}$ are values of response and of explanatory variables for the ith case, respectively. β^0 is the vector of regression coefficients and e_i represents random fluctuation (disturbance) of Y_i from the mean value $\mathsf{E}Y_i$. (To be complete, let us add that of course $x_i^{\mathrm{T}}\beta = \sum_{j=1}^{p} x_{ij}\beta_j$.)

Throughout the paper we shall assume that the random variables are defined on a basic probability space $(\Omega, \mathcal{A}, P)$ (other assumptions are given below).

Let us recall that we made (in Part I) one exception from the commonly used notation. Since in what follows we shall use for the description of sets somewhat complicated expressions containing moreover indices, we shall write (in many cases) $I\,\{\textit{property describing the set } A\}$ instead of traditional notation $I_{\{\textit{property describing the set } A\}}$.

In what follows the definition of *the least trimmed squares* will be considered in the form:

Definition 1. For a compact set $\mathcal{K}$ such that the vector of the true regression coefficients $\beta^0 \in \mathcal{K}^o$ the estimator given as

$$\hat{\beta}^{(\mathrm{LTS},n,h)} = \arg\min_{\beta \in \mathcal{K}} \sum_{i=1}^{h} r^2_{(i)}(\beta) \tag{2}$$

will be called *the least trimmed squares* (LTS).

It is clear that for given i the squared residual appears in the sum on the right hand side of (2) iff $r_i^2(\beta) \le r^2_{(h)}(\beta)$, so that we can write equivalently

$$\begin{aligned} \hat{\beta}^{(\mathrm{LTS},n,h)} &= \arg\min_{\beta \in \mathcal{K}} \sum_{i=1}^{n} r_i^2(\beta) \cdot I\left\{ r_i^2(\beta) \le r^2_{(h)}(\beta) \right\} \\ &= \arg\min_{\beta \in \mathcal{K}} \sum_{i=1}^{n} (Y_i - x_i^{\mathrm{T}}\beta)^2 \cdot I\left\{ r_i^2(\beta) \le r^2_{(h)}(\beta) \right\}. \end{aligned} \tag{3}$$

Now, denote $G(z)$ the distribution function of e_1^2. For any $\alpha \in (0,1)$, u_α^2 will be the upper α-quantile of $G(z)$, i. e.

$$P(e_1^2 > u_\alpha^2) = 1 - G(u_\alpha^2) = \alpha. \tag{4}$$

Further, denote by $[a]_{\mathrm{int}}$ the integer part of a and for any $n \in N$ put

$$h_n = [(1-\alpha)n]_{\mathrm{int}}. \tag{5}$$

Moreover, for any $a, b \in R$ we shall denote $(a,b)_{ord} = (\min\{a,b\}, \max\{a,b\})$ and the same will be used for the closed intervals. Finally, put $Q_n = \frac{1}{n}\sum_{i=1}^n x_i x_i^{\mathrm{T}}$ and for an arbitrary $\alpha \in (0,1)$ $Q_n(\alpha) = \frac{1}{n}\sum_{i=1}^n x_i x_i^{\mathrm{T}} I\{r_i^2(\beta^0) \le u_\alpha^2\}$.

Prior to continuing the discussion on *the least trimmed squares* it is useful to give the assumptions which will be used in the most assertions.

Assumptions $\mathcal{A}$

The sequences $\{x_i\}_{i=1}^\infty$ $(x_i \in R^p)$ is a fix sequence of nonrandom vectors from R^p. Further, the sequence $\{e_i\}_{i=1}^\infty$ $(e_i \in R)$ is a sequence of independent and identically distributed random variables. The distribution function $F(z)$ of random fluctuation e_1 is symmetric and absolutely continuous with a bounded density $f(z)$ which is strictly decreasing on R^+. The density is positive on $(-\infty, \infty)$ and has bounded in absolute value the first and the second derivative. The second derivative is further Lipschitz of the first order. Moreover,

$$\sum_{i=1}^{n} \|x_i\|^4 = \mathcal{O}(n) \quad \text{and} \quad \mathsf{E}e_1^4 = \kappa_4 \in (0,\infty). \tag{6}$$

Finally,

$$\lim_{n\to\infty} Q_n = Q \tag{7}$$

where Q is a regular matrix (and convergence is of course assumed coordinatewise).

Alternatively to the Assumptions $\mathcal{A}$, we shall use Assumptions $\mathcal{B}$ (the reasons for it were given in Part I).

Assumptions $\mathcal{B}$

The sequences $\{x_i\}_{i=1}^{\infty}$ $(x_i \in R^p)$ is a fix sequence of nonrandom vectors from R^p. Moreover, (7) holds for some regular matrix Q. Further for any $n \in N$

$$\max_{1 \le i \le n,\ 1 \le j \le p} |x_{ij}| = \mathcal{O}(1). \tag{8}$$

The sequence $\{e_i\}_{i=1}^{\infty}$ $(e_i \in R)$ is a sequence of independent and identically distributed random variables with absolutely continuous symmetric distribution function $F(z)$. There is a neighbourhood of u_α in which the distribution $F(z)$ has a bounded density $f(z)$ which is positive and has bounded in absolute value the first and the second derivative. The second derivative is further Lipschitz of the first order. Moreover, the density $f(z)$ is strictly decreasing on R^+ and $\mathsf{E}e_1^4 = \kappa_4 \in (0, \infty)$.

We have proved (in Part I) that

$$\tilde{\beta}^{(\mathrm{LTS},n,h)} = \arg\min_{\beta \in R^p} \sum_{i=1}^{h} r_{(i)}^2(\beta) \tag{9}$$

can be found among solutions of

$$\sum_{i=1}^{n} \left[(Y_i - x_i^{\mathrm{T}}\beta)\, x_i \cdot I\left\{ r_i^2(\beta) \le r_{(h)}^2(\beta) \right\} \right] = 0, \tag{10}$$

i. e. that at the point given as the solution of the extremal problem (9) the relation (10) holds. Notice please that whenever we prove that the estimator given by (2) is consistent (i. e. exists and converges in probability to β^0), it also solves (10).

Assumptions $\mathcal{C}$

There are distribution functions $H^{(\beta)}(t), t \in R, \beta \in R^p$ such that for any compact set $\mathcal{W} \subset R^p$

$$\sup_{\beta \in \mathcal{W}} \sup_{t \in R} \left| \frac{1}{n} \sum_{i=1}^{n} I\left\{ x_i^{\mathrm{T}}(\beta - \beta^0) \le t \right\} - H^{(\beta)}(t) \right| = \mathcal{O}(n^{-\frac{1}{2}}). \tag{11}$$

Remark 1. Recently it was found that when X_i's are i.i.d. the first supremum in (11) can be taken over R^p, see [5].

$\sqrt{n}$-CONSISTENCY OF THE LEAST TRIMMED SQUARES

Lemma 1. Let $\alpha \in \left(0, \frac{1}{2}\right)$ and let Assumptions $\mathcal{A}$ or $\mathcal{B}$ and $\mathcal{C}$ be fulfilled. Then for any $\varepsilon > 0$ and $\Delta > 0$ there are $\delta = \delta_{\Delta,\varepsilon} > 0$ and $n_{\Delta,\varepsilon} \in N$ such that for all $n > n_{\Delta,\varepsilon}$

$$P\left(\sup_{\beta \in B(\beta^0,\delta)} \left| r^2_{(h_n)}(\beta) - u^2_\alpha \right| < \Delta \right) > 1 - \varepsilon.$$

P r o o f. Let us fix $\varepsilon > 0$ and $\Delta > 0$. Employing Lemma 1 of Part I we can find a constant $K^{(\varepsilon)} < \infty$ and $n^{(1)} \in N$ so that for any $n > n^{(1)}$ we have

$$P\left(\sup_{\beta \in \mathcal{K}} \left| r^2_{(h_n)}(\beta) - u^2_\alpha(\beta)\} \right| < n^{-\frac{1}{2}} K^{(\varepsilon)} \right) > 1 - \varepsilon. \tag{12}$$

Let us find $n_{\Delta,\varepsilon} \geq n^{(1)}$ such that for all $n > n_{\Delta,\varepsilon}$ we have $n^{-\frac{1}{2}} K^{(\varepsilon)} < \frac{1}{2}\Delta$. In the proof of Lemma 2 of Part I we have shown that there is a $\delta \in (0,1)$ so that for all $\beta, \tilde{\beta} \in R^p$, $\|\beta - \tilde{\beta}\| < \delta$ we have

$$\left| u_\alpha(\beta) - u_\alpha(\tilde{\beta}) \right| \leq K \cdot \|\beta - \tilde{\beta}\|^2.$$

Utilizing it for $\tilde{\beta} = \beta^0$, we can find $\delta > 0$ so that for any $\|\beta - \beta^0\| < \delta$ we have

$$\left| u_\alpha(\beta) - u_\alpha(\beta^0) \right| < \frac{1}{2}\Delta. \tag{13}$$

Taking into account that $u_\alpha(\beta^0) = u_\alpha$ and (12) together with (13), we conclude the proof. □

Assertion 1. Let $\{e_i\}_{i=1}^\infty$ $(e_i \in R)$ be a sequence of independent and identically distributed random variables with absolutely continuous distribution function $F(z)$. Then for any $n \in N$ and any $i = 1, 2, \ldots, n$ we have

$$P(r^2_i(\beta^0) = r^2_{(h_n)}(\beta^0)) = \frac{1}{n}.$$

P r o o f. The proof can be found in [1]. Since it is not easy available, let us give it (moreover, it's short). First of all, let us recall that for any $i \neq j$, $i, j = 1, 2, \ldots, n$ and $n \in N$

$$P(r^2_i(\beta^0) = r^2_j(\beta^0)) = 0.$$

Due to the fact that the random variables e_i's are i.i.d., we have for all pairs $i, j = 1, 2, \ldots, n$

$$P(r^2_i(\beta^0) = r^2_{(h_n)}(\beta^0)) = P(r^2_j(\beta^0) = r^2_{(h_n)}(\beta^0))$$

and

$$\sum_{i=1}^n P(r^2_i(\beta^0) = r^2_{(h_n)}(\beta^0)) = 1.$$

That concludes the proof. □

Remark 2. The previous assertion shows that under the assumption that the sequence $\{e_i\}_{i=1}^{\infty}$ is i.i.d., for any $n \in N$ the probability space Ω can be decomposed on n equiprobable sets such that for each of them it holds that on it the hth order statistic among $e_1^2, e_2^2, \ldots, e_n^2$ is represented by the square of one fix random variable, say $e_{i_0}^2$ for some $i_0 \in \{1, 2, \ldots, n\}$. It is clear that $h-1$ of the other $n-1$ random variables have to be smaller than $e_{i_0}^2$ and $n-h$ larger than it. Hence these $n-1$ random variables are not even conditionally independent, if the condition is $e_{(h)}^2 = e_{i_0}^2$. Nevertheless we may prove following:

Lemma 2. Let $\{e_i\}_{i=1}^{\infty}$ ($e_i \in R$) be a sequence of independent and identically distributed random variables with absolutely continuous distribution function $F(z)$. Then for any $n \in N$, any $i_0 \in \{1, 2, \ldots, n\}$ and any $h-1$ tuple selected from the indices $\{1, 2, \ldots, i_0 - 1, i_0 + 1, \ldots, n\}$ the random variables $e_1, e_2, \ldots, e_{i_0-1}$, $e_{i_0+1}, \ldots, e_n$ are conditionally independent on the set on which $e_{i_0}^2 = e_{(h)}^2$ and $e_i^2 < e_{(h)}^2$ for $i \in \{i_1, i_2, \ldots, i_{h-1}\}$ while $e_i^2 > e_{(h)}^2$ for $i \notin \{i_0, i_1, i_2, \ldots, i_{h-1}\}$. Moreover the conditional density of each random variable (except of e_{i_0}) is proportional to $f(z)$ and the rest of the corresponding formula may be bounded by the same constant over the whole space Ω.

P r o o f . As we have already said, due to the previous assertion for any $n \in N$ the probability space Ω can be decomposed on n equiprobable sets, on each of them the hth order statistic among $e_1^2, e_2^2, \ldots, e_n^2$, ($h \in \{1, 2, \ldots, n\}$) is represented by the square of one e_i's, say $e_{i_0}^2$ for some $i_0 \in \{1, 2, \ldots, n\}$. Notice that none of the other $n-1$ e_i's ($i \neq i_0$) has a special position among the others, except of the fact that $h-1$ of them are smaller than $e_{(h)}^2$ while the others are larger. So, let us select $h-1$-tuple of indices, say $i_1, i_2, \ldots, i_{h-1}$ of those random variables, squares of which will be assumed to be smaller than $e_{i_0}^2$, i. e. $e_{i_j}^2 < e_{(h)}^2$ for $j = 1, 2, \ldots, h-1$. By this selection we give also the set of indices, say $i_{h+1}, \ldots, i_n$ for which $e_{(h)}^2 < e_{i_j}^2$. Now, formally the conditional density is the same for all possibilities of selection of $h-1$ tuples of r.v.'s. So, the probability space may be decomposed into the sets so that each of them is characterized by

- $e_{(h)}^2 = e_{i_0}^2$
- for the indices $i_1, i_2, \ldots, i_{h-1}$ $\quad e_{i_j}^2 < e_{(h)}^2$ while for other indices $e_{(h)}^2 < e_{i_j}^2$.

Of course, i_0 is successively $1, 2, \ldots, n$ and the $h-1$-tuple $i_1, i_2, \ldots, i_{h-1}$ runs through all $\binom{n}{h-1}$ possibilities. (It is easy to see that we have $n \cdot \binom{n}{h-1}$ of such sets.) Let us call this partition $\mathcal{S}$. Now, the conditional density of $e_1^2, e_2^2, \ldots, e_{i_0-1}^2, e_{i_0+1}^2, \ldots, e_n^2$, under the condition given by the set S_0 (say) from the partition $\mathcal{S}$, is evidently proportional to

$$\prod_{j=1}^{h-1} f(z_j) \prod_{k=h+1}^{n} f(z_k) \qquad \text{for } \max_{1 \le j \le h-1} z_j < e_{(h)}^2 \quad \text{and} \quad e_{(h)}^2 < \min_{h+1 \le k \le n} z_k \tag{14}$$

and equal to 0 otherwise. Since the integral of the conditional density over the set S_0 is equal 1, we can even find the constant by which we need to multiply (14) to obtain the joint conditional density of $e_1^2, e_2^2, \ldots, e_{i_0-1}^2, e_{i_0+1}^2, \ldots, e_n^2$. Since the situation is fully symmetric in all indices, the conditional density is formally the same on all elements of the partition $\mathcal{S}$. It implies that the conditional density can be bounded by the same constant over the whole Ω. □

Remark 3. The most important result of the previous lemma is that the conditional density can be bounded by the same constant over the whole space Ω. Of course, this constant depends on n. Nevertheless, if we look for a probability that e_j^2 falls into an interval, we can evaluate this probability as conditional one over all sets of division $\mathcal{S}$, except of one on which $e_j^2 = e_{(h)}^2$. The unconditional probability (or its upper bound) is then given as the mean value over all these sets. Since the probability of the event $\{e_j^2 = e_{(h)}^2\}$ is $\frac{1}{n}$, we conclude that the probability in question is proportional to the (upper bound of) density and the length of the respective interval. Moreover, due to the fact that we take mean value over all sets of the division $\mathcal{S}$, in this case the corresponding constant (of proportionality) does not depend on $n \in N$. That is why, at some points of the proofs in the text which follows, we shall consider the conditional probabilities of some events under the condition that:

- $e_{(h)}^2 = e_{i_0}^2$,
- $\max\{e_{i_1}^2, e_{i_2}^2, \ldots, e_{i_{h-1}}^2\} < e_{(h)}^2$

and

- $\min\{e_{i_{h+1}}^2, \ldots, e_{i_n}^2\} > e_{(h)}^2$.

Let us denote this condition $\mathcal{C}(i_0, i_1, i_2, \ldots, i_{h-1})$.

Theorem 1. Let $\alpha \in \left(0, \frac{1}{2}\right)$ and let Assumptions $\mathcal{A}$ or $\mathcal{B}$ and $\mathcal{C}$ hold. Further, let $\mathcal{K}$ be a compact subset of R^p, $\beta^0 \in \mathcal{K}^o$. Then $\hat{\beta}^{(\mathrm{LTS},n,h)}$ is $\sqrt{n}$-consistent, i. e.

$$\sqrt{n}\left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right) = \mathcal{O}_p(1) \quad \text{as } n \to \infty.$$

P r o o f. Let us recall that

$$\hat{\beta}^{(\mathrm{LTS},n,h)} = \arg\min_{\beta \in \mathcal{K}} \rho(\beta)$$

where

$$\rho(\beta) = \sum_{i=1}^{n} \left[(Y_i - x_i^{\mathrm{T}}\beta)^2 \cdot I\left\{ r_i^2(\beta) \le r_{(h)}^2(\beta) \right\} \right].$$

Since we already know that $\hat{\beta}^{(\mathrm{LTS},n,h)}$, independently of $\mathcal{K}$ is consistent, we may restrict ourselves in the rest of proof, say, on $\mathcal{K} = \bar{B}(\beta^0, 1)$ and on a corresponding subset (say O_1) of the space Ω (such that for any $\omega \in O_1$ $\hat{\beta}^{(\mathrm{LTS},n,h)} \in \bar{B}(\beta^0, 1)$; of course, it means simultaneously that we restrict ourselves, without recalling it, on n which are larger than some n_1). We have shown in Part I that

$$\frac{\partial\rho(\beta)}{\partial\beta} = -2\sum_{i=1}^{n}\left[(Y_i - x_i^{\mathrm{T}}\beta)\,x_i \cdot I\left\{r_i^2(\beta) \le r_{(h)}^2(\beta)\right\}\right] \qquad \text{a. e.} \tag{15}$$

(see (19) of [4]) and so we may write[1]

$$\left.\frac{\partial\rho(\beta)}{\partial\beta}\right|_{\beta=\hat\beta^{(\mathrm{LTS},n,h)}} - \left.\frac{\partial\rho(\beta)}{\partial\beta}\right|_{\beta=\beta^0}$$
$$= -2\sum_{i=1}^{n}\left[\left(Y_i - x_i^{\mathrm{T}}\hat\beta^{(\mathrm{LTS},n,h)}\right)\,x_i \cdot I\left\{r_i^2(\hat\beta^{(\mathrm{LTS},n,h)}) \le r_{(h)}^2(\hat\beta^{(\mathrm{LTS},n,h)})\right\}\right.$$
$$\left. -(Y_i - x_i^{\mathrm{T}}\beta^0)\,x_i \cdot I\left\{r_i^2(\beta^0) \le r_{(h)}^2(\beta^0)\right\}\right]. \tag{16}$$

Including into (16)

$$\pm\sum_{i=1}^{n}\left(Y_i - x_i^{\mathrm{T}}\hat\beta^{(\mathrm{LTS},n,h)}\right)\,x_i \cdot I\left\{r_i^2(\beta^0) \le r_{(h)}^2(\beta^0)\right\}$$

and taking into account once again (15) together with the fact that $\left.\frac{\partial\rho(\beta)}{\partial\beta}\right|_{\beta=\hat\beta^{(\mathrm{LTS},n,h)}} =$ $= 0$ we arrive at

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[(Y_i - x_i^{\mathrm{T}}\beta^0)\,x_i \cdot I\left\{r_i^2(\beta^0) \le r_{(h)}^2(\beta^0)\right\}\right]$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}x_i x_i^{\mathrm{T}}\left(\hat\beta^{(\mathrm{LTS},n,h)} - \beta^0\right) I\left\{r_i^2(\beta^0) \le r_{(h)}^2(\beta^0)\right\}$$
$$-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(Y_i - x_i^{\mathrm{T}}\hat\beta^{(\mathrm{LTS},n,h)}\right)\cdot x_i\left[I\left\{r_i^2(\hat\beta^{(\mathrm{LTS},n,h)})\right.\right.$$
$$\left.\left.\le r_{(h)}^2(\hat\beta^{(\mathrm{LTS},n,h)})\right\} - I\left\{r_i^2(\beta^0) \le r_{(h)}^2(\beta^0)\right\}\right]$$
$$= \frac{1}{n}\sum_{i=1}^{n}x_i x_i^{\mathrm{T}}\left[I\{e_i^2 \le u_\alpha^2\}\right]\cdot\sqrt{n}\left(\hat\beta^{(\mathrm{LTS},n,h)} - \beta^0\right)$$

---

[1] A very first idea can be to find the second derivative $\frac{\partial^2\rho(\beta)}{\partial\beta\cdot\partial\beta^{\mathrm{T}}} = -2\sum_{i=1}^{n}x_i x_i^{\mathrm{T}}\cdot I\left\{r_i^2(\beta) \le r_{(h)}^2(\beta)\right\}$ (along the same lines as it was done for the first derivative in Part I) and then to use the Mean Value Theorem, see e. g. Hewitt and Stromberg [2]. Unfortunately, the Assertion 1 of Part I indicates that the sets on which the hth order statistic among the squared disturbances is represented by the square of one given random variable have "radiuses" approximately of order $\frac{1}{n}$. In other words, as follows from the considerations which led to the formula for the derivative in Part I, the discontinuities of the first (as well as the second) derivative have distance of order $\frac{1}{n}$. On the other hand we may expect that (at the best) $\left\|\hat\beta^{(\mathrm{LTS},n,h)} - \beta^0\right\| = O(n^{-\frac{1}{2}})$, so that we have to conclude: *For this purpose the second derivative is not continuous.*

$$+\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{\mathrm{T}} \left[I\Big\{ e_i^2 \le e_{(h)}^2 \Big\} - I\{ e_i^2 \le u_\alpha^2 \} \right] \sqrt{n} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right)$$

$$+\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{\mathrm{T}} \left[I\Big\{ r_i^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \le r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \Big\} \right.$$

$$\left. - I\Big\{ r_i^2(\beta^0) \le r_{(h)}^2(\beta^0) \Big\} \right] \cdot \sqrt{n} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right)$$

$$-\frac{1}{\sqrt{n}}\sum_{i=1}^{n} e_i x_i \left[I\Big\{ r_i^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \le r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \Big\} \right.$$

$$\left. - I\Big\{ r_i^2(\beta^0) \le r_{(h)}^2(\beta^0) \Big\} \right]. \tag{17}$$

We shall study terms of (17) one by one. Let us start with that on the left hand side. It can be written as (remember that $r_i^2(\beta^0) = e_i^2$)

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \left\{ e_i x_i \cdot \left[I\Big\{ e_i^2 \le e_{(h)}^2 \Big\} - I\{ e_i^2 \le u_\alpha^2 \} \right] \right\} \tag{18}$$

$$+\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \left[e_i x_i \cdot I\{ e_i^2 \le u_\alpha^2 \} \right].$$

Evidently

$$I\Big\{ e_i^2 \le e_{(h)}^2 \Big\} - I\{ e_i^2 \le u_\alpha^2 \} = 1 \quad \Leftrightarrow \quad u_\alpha^2 < e_i^2 \le e_{(h)}^2 \tag{19}$$

and

$$I\Big\{ e_i^2 \le e_{(h)}^2 \Big\} - I\{ e_i^2 \le u_\alpha^2 \} = -1 \quad \Leftrightarrow \quad e_{(h)}^2 < e_i^2 \le u_\alpha^2. \tag{20}$$

Prior to continuing in proof, let us denote by $\mathcal{C}(i_0, i_1, i_2, \ldots, i_{h-1}, z)$ the condition under which $\mathcal{C}(i_0, i_1, i_2, \ldots, i_{h-1})$ holds (see Remark 3) and $\sqrt{e_{(h)}^2} = z \in R$ (and for the sake of space and simplicity of notations let us write $\mathcal{C}(z)$ instead of $\mathcal{C}(i_0, i_1, i_2, \ldots, i_{h-1}, z)$). Then Lemma 1 of Part I implies that for any $\nu \in (0,1)$ there is $n_\nu \in N$ and a constant $K^{(1)} < \infty$ so that for all $n > n_\nu$ there is a set A_n so that $P(A_n) > 1 - \nu$ and for any $\omega \in A_n$

$$\left| e_{(h)}^2 - u_\alpha^2 \right| < n^{-\frac{1}{2}} K^{(1)}$$

so that whenever (19) and (20) hold, then for all $n > n_\nu$ and any $\omega \in A_n$ also for some finite $K^{(2)}$

$$||e_i| - u_\alpha| < n^{-\frac{1}{2}} K^{(2)}. \tag{21}$$

Lemma 2 then guarantee that there are $K^{(3)} < \infty$ and $K^{(4)} < \infty$ such that

$$P\left(\left\{ -\sqrt{e_{(h)}^2} \le e_i < -u_\alpha \right\} \cap A_n | \mathcal{C}(z) \right) = P\left(\left\{ u_\alpha < e_i \le \sqrt{e_{(h)}^2} \right\} \cap A_n | \mathcal{C}(z) \right)$$

$$= K^{(3)} \cdot f(u_\alpha)\,(z - u_\alpha) + \zeta_i^{(1)} \tag{22}$$

as well as

$$P\left(\left\{-u_\alpha \le e_i < -\sqrt{e^2_{(h)}}\right\} \cap A_n | \mathcal{C}(z)\right) = P\left(\left\{\sqrt{e^2_{(h)}} < e_i \le u_\alpha\right\} \cap A_n | \mathcal{C}(z)\right)$$

$$= K^{(3)} \cdot f(u_\alpha)\,(u_\alpha - z) + \zeta_i^{(2)} \tag{23}$$

where

$$|\zeta_i^{(j)}| \le n^{-1} K^{(4)}, \quad j = 1, 2. \tag{24}$$

(Let us recall that, as follows from Lemma 2, $K^{(3)}$ as well as $K^{(4)}$ are the same for all $i = 1, 2, \ldots, n$ and $z \in R$.) But (21), (22), (23) and (24) immediately implies that

$$\mathsf{E}\left\{e_i \left[I\left\{e_i^2 \le e^2_{(h)}\right\} - I\{e_i^2 \le u_\alpha^2\}\right] \cdot I\{A_n\}\right\}$$

$$= \mathsf{E}_{\mathcal{C}(z)}\left\{\left[\mathsf{E} e_i \left[I\left\{e_i^2 \le e^2_{(h)}\right\} - I\{e_i^2 \le u_\alpha^2\}\right] \cdot I\{A_n\} | \mathcal{C}(z)\right]\right\}$$

$$\le \mathsf{E}_{\mathcal{C}(z)}\left\{\left[-u_\alpha + n^{-\frac{1}{2}} \cdot K^{(2)}\right] \cdot \left[K^{(3)} \cdot f(u_\alpha)\,(z - u_\alpha) + n^{-1} \cdot K^{(4)}\right]\right.$$

$$+ \left[u_\alpha + n^{-\frac{1}{2}} \cdot K^{(2)}\right] \cdot \left[K^{(3)} \cdot f(u_\alpha)\,(z - u_\alpha) + n^{-1} \cdot K^{(4)}\right]$$

$$- \left[-u_\alpha - n^{-\frac{1}{2}} \cdot K^{(2)}\right] \cdot \left[K^{(3)} \cdot f(u_\alpha)\,(u_\alpha - z) - n^{-1} \cdot K^{(4)}\right]$$

$$\left. - \left[u_\alpha - n^{-\frac{1}{2}} \cdot K^{(2)}\right] \cdot \left[K^{(3)} \cdot f(u_\alpha)\,(u_\alpha - z) - n^{-1} \cdot K^{(4)}\right]\right\}$$

where the subscript $\mathcal{C}(z)$ indicates that the mean value is taken over the condition $\mathcal{C}(z)$. In this case it means that we should take into account all possible values of $z = \sqrt{e^2_{(h)}}$ (see Lemma 2). Of course, due to the presence of $I\{A_n\}$ in the expression the values of z are restricted on $\{-u_\alpha - n^{-\frac{1}{2}} K^{(2)}, -u_\alpha + n^{-\frac{1}{2}} K^{(2)}\} \cup \{u_\alpha - n^{-\frac{1}{2}} K^{(2)}, u_\alpha + n^{-\frac{1}{2}} K^{(2)}\}$. In the same way we can find the lower bound for the mean value in question. Finally taking into account that $|z - u_\alpha| = \mathcal{O}(n^{-\frac{1}{2}})$, we conclude that there is a constant $K^{(5)} < \infty$ so that for all $n > n_\nu$ we have

$$\left|\mathsf{E} e_i \left[I\left\{e_i^2 \le e^2_{(h)}\right\} - I\{e_i^2 \le u_\alpha^2\}\right] \cdot I\{A_n\}\right| \le n^{-1} K^{(5)}. \tag{25}$$

Along similar lines we can find $K^{(6)} < \infty$ so that for all $n > n_\nu$

$$\operatorname{var}\left\{e_i \left[I\left\{e_i^2 \le e^2_{(h)}\right\} - I\{e_i^2 \le u_\alpha^2\}\right] \cdot I\{A_n\}\right\} \le n^{-\frac{1}{2}} K^{(6)}. \tag{26}$$

As a side product of the previous considerations we obtain (for some $K^{(7)} < \infty$)

$$\mathsf{E}\left|\left[I\left\{e_i^2 \le e^2_{(h)}\right\} - I\{e_i^2 \le u_\alpha^2\}\right] \cdot I\{A_n\}\right| < n^{-\frac{1}{2}} K^{(7)} \tag{27}$$

which we shall need later on. As all x_i's are deterministic, we have

$$\left\|\mathsf{E}\left\{e_i x_i \left[I\left\{e_i^2 \le e^2_{(h)}\right\} - I\{e_i^2 \le u_\alpha^2\}\right] \cdot I\{A_n\}\right\}\right\|$$

$$= \|x_i\| \cdot \left| \mathsf{E}\left\{ e_i \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right\} \cdot I\{A_n\} \right|$$

and then (25) and (26) imply that there is some $K^{(8)} < \infty$ so that for all $n > n_\nu$

$$\left\| \mathsf{E}\left\{ e_i x_i \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \cdot I\{A_n\} \right\} \right\| < n^{-1} \|x_i\| K^{(8)} \tag{28}$$

and similarly for any $j = 1, 2, \ldots, p$

$$\text{var}\left\{ e_i x_{ij} \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \cdot I\{A_n\} \right\} < n^{-\frac{1}{2}} \|x_i\|^2 K^{(8)}. \tag{29}$$

Now (18) can be modified into the form

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ e_i x_i \cdot \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right. \right.$$

$$\left. \left. - \mathsf{E}\left\{ e_i x_i \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right\} \right] \right\} \tag{30}$$

$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ \mathsf{E}\left\{ e_i x_i \cdot \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right\} \right\}. \tag{31}$$

Finally, taking into account (29) we conclude that for any $\Delta > 0$

$$P\left(\left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ e_i x_{ij} \cdot \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right. \right. \right.$$

$$\left. \left. \left. - \mathsf{E}\left\{ e_i x_{ij} \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right\} \right\} \right| > \Delta \right)$$

$$\leq \mathsf{E}\left\{ \Delta^{-2} \text{var}\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\{ e_i x_{ij} \cdot \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right. \right. \right.$$

$$\left. \left. \left. - \mathsf{E}\left\{ e_i x_{ij} \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right\} \right\} \Big| \mathcal{C}(z) \right] \right\}$$

$$< \Delta^{-2} n^{-\frac{3}{2}} \sum_{i=1}^{n} \|x_i\|^2 \cdot K^{(8)}. \tag{32}$$

Then (6) implies that (30) is $o_p(1)$. Similarly, employing (28) we find that also

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left\| \mathsf{E}\left\{ e_i x_i \cdot \left[I\left\{ e_i^2 \leq e_{(h)}^2 \right\} - I\{e_i^2 \leq u_\alpha^2\} \right] \right\} \right\| < n^{-\frac{3}{2}} \sum_{i=1}^{n} \|x_i\| \cdot K^{(8)},$$

i. e. (31) is $o(1)$. Combining just derived facts we conclude that the left hand side of (17) is equal to

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[e_i x_i \cdot I\{e_i^2 \leq u_\alpha^2\} \right] + o_p(1) \tag{33}$$

and taking into account once again (6), we can utilize Central Limit Theorem and then conclude that the left hand side of (17) is $O_p(1)$. (It is clear that it was possible to show that the left hand side of (17) is $O_p(1)$ in a simpler way. But we shall need the fact that the left hand side has just the form given in (33) later on.)

Now, let us turn to the terms on the right hand side of (17). The first one can be written as

$$\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{\mathrm{T}} \left[I\{e_i^2 \le u_\alpha^2\} - \mathsf{E}I\{e_i^2 \le u_\alpha^2\}\right] \cdot \sqrt{n}\left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right)$$

$$+\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{\mathrm{T}} (1-\alpha) \cdot \sqrt{n}\left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right).$$

Now taking into account (7) and applying the law of large numbers on the sequences

$$\left\{x_{ij}x_{i\ell}\left[I\{e_i^2 \le u_\alpha^2\} - \mathsf{E}I\{e_i^2 \le u_\alpha^2\}\right] I\{A_n\}\right\}_{i=1}^{\infty},$$

(for $j, \ell = 1, 2, \ldots, p$), we conclude that the first term of the right hand side of (17) is equal to

$$[Q_n (1-\alpha) + o_p(1)] \cdot \sqrt{n}\left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right). \tag{34}$$

Let us consider the second term of the right hand side. Taking into account (27), we obtain for any $\varepsilon > 0$

$$P\left(\left\|\frac{1}{n}\sum_{i=1}^{n} x_i x_i^{\mathrm{T}}\left[I\left\{e_i^2 \le e_{(h)}^2\right\} - I\{e_i^2 \le u_\alpha^2\}\right]\right\| > \varepsilon\right)$$

$$< \frac{1}{n\varepsilon}\sum_{i=1}^{n}\|x_i\|^2 \cdot \mathsf{E}\left|I\left\{e_i^2 \le e_{(h)}^2\right\} - I\{e_i^2 \le u_\alpha^2\}\right| < n^{-\frac{3}{2}}\varepsilon^{-1}\sum_{i=1}^{n}\|x_i\|^2 K^{(7)}, \tag{35}$$

so that the second term is of order $o_p(1) \cdot \sqrt{n}\left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right)$.

Now let ε and η be positive numbers and denote by J the upper bound of the density $f(z)$. Due to (6) there is a finite $K^{(9)}$ and $n^{(1)} \in N$ so that for all $n > n^{(1)}$

$$\frac{1}{n}\sum_{i=1}^{n}\|x_i\|^2 < K^{(9)} \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n}\|x_i\|^3 < K^{(9)}. \tag{36}$$

Let us put $\tau = \frac{1}{16}\varepsilon \cdot \eta \cdot \left[K^{(9)} \cdot J\right]^{-1}$. Employing Lemma 1 of Part I and Lemma 1 of this part of paper, we can find $n_\varepsilon > n^{(1)}$, $K^{(10)} < \infty$ and $\delta_\varepsilon \in (0,1)$ so that for all $n > n_\varepsilon$ and all $\beta \in R^p$, $\left\|\beta - \beta^0\right\| < \delta_\varepsilon$ the set

$$B_n = \left\{\omega \in \Omega : \sup_{\beta \in B(\beta^0, \delta_\varepsilon)}\left|r_{(h_n)}^2(\beta) - u_\alpha^2\right| < \frac{1}{4}\tau \text{ and } \left|e_{(h_n)}^2 - u_\alpha^2\right| < n^{-\frac{1}{2}}K^{(10)} \cdot J^{-1}\right\}$$

has probability

$$P(B_n) > 1 - \frac{1}{2}\varepsilon. \tag{37}$$

Since $\hat{\beta}^{(\mathrm{LTS},n,h)}$ is consistent, there is $n_\delta > n_\varepsilon$ such that for all $n > n_\delta$

$$C_n = \left\{\omega \in \Omega : \left\| \hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right\| < \delta_\varepsilon \right\} \qquad \text{and} \qquad P(C_n) > 1 - \frac{1}{2}\varepsilon. \tag{38}$$

Now, let us restrict ourselves on $n > n_\delta$ and $\omega \in B_n \cap C_n$ and let us make an idea when

$$I\left\{ r_i^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \le r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \right\} - I\left\{ r_i^2(\beta^0) \le r_{(h)}^2(\beta^0) \right\} \neq 0. \tag{39}$$

If (39) holds then either

$$r_i^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \le r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \qquad \text{and} \qquad r_i^2(\beta^0) > r_{(h)}^2(\beta^0) \tag{40}$$

or

$$r_i^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) > r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \qquad \text{and} \qquad r_i^2(\beta^0) \le r_{(h)}^2(\beta^0) \tag{41}$$

Due to $r_i(\hat{\beta}^{(\mathrm{LTS},n,h)}) = e_i - x_i^{\mathrm{T}} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right)$, we immediately find that (40) holds iff either

$$-\sqrt{r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)})} + x_i^{\mathrm{T}} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right) \le e_i < -\sqrt{e_{(h)}^2} \tag{42}$$

or

$$\sqrt{e_{(h)}^2} < e_i \le \sqrt{r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)})} + x_i^{\mathrm{T}} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right). \tag{43}$$

Similarly, (41) holds iff either

$$-\sqrt{e_{(h)}^2} \le e_i < -\sqrt{r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)})} + x_i^{\mathrm{T}} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right) \tag{44}$$

or

$$\sqrt{r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)})} + x_i^{\mathrm{T}} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0 \right) < e_i \le \sqrt{e_{(h)}^2}). \tag{45}$$

Now taking into account that the event in (42) is a subset of

$$-u_\alpha - \frac{1}{4}\tau - n^{-\frac{1}{2}} K^{(10)} - \|x_i\|\tau \le e_i < -u_\alpha + n^{-\frac{1}{2}} K^{(10)} \tag{46}$$

we conclude that it has, for $n > n_{(2)} = \max\{n_\delta, (64K^{(9)} \cdot K^{(10)} \cdot J \cdot (3\varepsilon\eta)^{-1})^2\}$, probability less than $\frac{1}{16}\frac{\varepsilon\eta(1+\|x_i\|)}{K^{(9)}}$. Of course, we may carry out similar considerations for all events in (43), (44) and (45). Then we obtain for any $n > n_{(2)}$

$$\begin{aligned} & P\left(\frac{1}{n}\sum_{i=1}^n \left\| x_i x_i^{\mathrm{T}} \right\| \cdot \left| I\left\{ r_i^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \le r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \right\} \right.\right. \\ & \qquad \left.\left. - I\left\{ r_i^2(\beta^0) \le r_{(h)}^2(\beta^0) \right\} \right| \cdot I\{B_n\} \cdot I\{C_n\} > \eta \right) \\ < \; & \eta^{-1} \cdot \mathsf{E}\left\{ \frac{1}{n}\sum_{i=1}^n \left\| x_i x_i^{\mathrm{T}} \right\| \cdot \left| I\left\{ r_i^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \le r_{(h)}^2(\hat{\beta}^{(\mathrm{LTS},n,h)}) \right\} \right.\right. \\ & \qquad \left.\left. - I\left\{ r_i^2(\beta^0) \le r_{(h)}^2(\beta^0) \right\} \right| \cdot I\{B_n\} \cdot I\{C_n\} \right\} \le \varepsilon. \end{aligned} \tag{47}$$

So we have shown that

$$\frac{1}{n}\sum_{i=1}^{n}\left\|x_i x_i^{\mathrm{T}}\right\|\cdot\left|I\left\{r_i^2(\hat\beta^{(\mathrm{LTS},n,h)})\le r_{(h)}^2(\hat\beta^{(\mathrm{LTS},n,h)})\right\}\right.$$
$$\left.-I\left\{r_i^2(\beta^0)\le r_{(h)}^2(\beta^0)\right\}\right|\cdot I\{B_n\cap C_n\}=o_p(1).$$

Since the last but one term in (17) can be written as

$$\frac{1}{n}\sum_{i=1}^{n}x_i x_i^{\mathrm{T}}\left[I\left\{r_i^2(\hat\beta^{(\mathrm{LTS},n,h)})\le r_{(h)}^2(\hat\beta^{(\mathrm{LTS},n,h)})\right\}-I\left\{r_i^2(\beta^0)\le r_{(h)}^2(\beta^0)\right\}\right]\times$$
$$\times\{I\{B_n\}\cdot I\{C_n\}+I\{B_n^c\}+I\{C_n^c\}\}\cdot\sqrt{n}\left(\hat\beta^{(\mathrm{LTS},n,h)}-\beta^0\right)$$

(47) implies that the last but one term of (17) is equal to

$$\sqrt{n}\left(\hat\beta^{(\mathrm{LTS},n,h)}-\beta^0\right)\cdot o_p(1). \tag{48}$$

It remains to cope with the last term of (17). For the (substantial) sake of space let us write (up to the end of the considerations about this term) $r_{(h)}^2(\hat\beta)$ instead of $r_{(h)}^2(\hat\beta^{(\mathrm{LTS},n,h)})$ and Δ_i instead of $x_i^{\mathrm{T}}\left(\hat\beta^{(\mathrm{LTS},n,h)}-\beta^0\right)$. In what follows we shall carry out the analysis of the last term in (17) in a rough way, in order to show that it is $O_p(1)$. It is due to the fact that at the present moment we are able to estimate $|r_{(h)}^2(\hat\beta^{(\mathrm{LTS},n,h)})-e_{(h)}^2|$ only by means of Lemma 1. When we shall know that $\hat\beta^{(\mathrm{LTS},n,h)}$ is $\sqrt{n}$-consistent, we will analyze this term better (in order to establish an asymptotic representation of $\hat\beta^{(\mathrm{LTS},n,h)}$). Similarly as in previous it is straightforward to find that the difference of indicators

$$I\left\{r_i^2(\hat\beta^{(\mathrm{LTS},n,h)})\le r_{(h)}^2(\hat\beta^{(\mathrm{LTS},n,h)})\right\}-I\left\{r_i^2(\beta^0)\le r_{(h)}^2(\beta^0)\right\} \tag{49}$$

is equal to one iff

$$-\sqrt{r_{(h)}^2(\hat\beta)}+\Delta_i\le e_i<-\sqrt{e_{(h)}^2}\quad\text{or}\quad\sqrt{e_{(h)}^2}<e_i\le\sqrt{r_{(h)}^2(\hat\beta)}+\Delta_i \tag{50}$$

and is equal to minus one iff

$$-\sqrt{e_{(h)}^2}\le e_i<-\sqrt{r_{(h)}^2(\hat\beta)}+\Delta_i\quad\text{or}\quad\sqrt{r_{(h)}^2(\hat\beta)}+\Delta_i<e_i\le\sqrt{e_{(h)}^2}. \tag{51}$$

The indicators of the events given in (50) and (51) can be further written as

$$I\left\{-\sqrt{r_{(h)}^2(\hat\beta)}+\Delta_i\le e_i<-\sqrt{e_{(h)}^2}\right\}=I\left\{-\sqrt{r_{(h)}^2(\hat\beta)}\le e_i<-\sqrt{e_{(h)}^2}\right\}$$
$$-I\left\{\min\{-\sqrt{r_{(h)}^2(\hat\beta)},-\sqrt{r_{(h)}^2(\hat\beta)}+\Delta_i\}\le e_i<\min\{-\sqrt{e_{(h)}^2},-\sqrt{r_{(h)}^2(\hat\beta)}+\Delta_i\}\right\}$$

$$+I\left\{-\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i \le e_i < \min\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{e^2_{(h)}}\}\right\}, \qquad (52)$$

$$I\left\{\sqrt{e^2_{(h)}} < e_i \le \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\right\} = I\left\{\sqrt{e^2_{(h)}} < e_i \le \sqrt{r^2_{(h)}(\hat{\beta})}\right\}$$
$$-I\left\{\max\{\sqrt{e^2_{(h)}}, \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\} < e_i \le \max\{\sqrt{r^2_{(h)}(\hat{\beta})}, \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\}\right\}$$
$$+I\left\{\max\{\sqrt{r^2_{(h)}(\hat{\beta})}, \sqrt{e^2_{(h)}}\} < e_i \le \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\right\}, \qquad (53)$$

$$I\left\{-\sqrt{e^2_{(h)}} \le e_i < -\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\right\} = I\left\{-\sqrt{e^2_{(h)}} \le e_i < \sqrt{r^2_{(h)}(\hat{\beta})}\right\}$$
$$-I\left\{\max\{-\sqrt{e^2_{(h)}}, -\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\} \le e_i < \max\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\}\right\}$$
$$+I\left\{\max\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{e^2_{(h)}}\} \le e_i < -\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\right\}, \qquad (54)$$

and

$$I\left\{\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i < e_i \le \sqrt{e^2_{(h)}}\right\} = I\left\{\sqrt{r^2_{(h)}(\hat{\beta})} < e_i \le \sqrt{e^2_{(h)}}\right\}$$
$$-I\left\{\mathrm{mim}\{\sqrt{r^2_{(h)}(\hat{\beta})}, \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\} < e_i \le \min\{\sqrt{e^2_{(h)}}, \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\}\right\}$$
$$+I\left\{\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i < e_i \le \min\{\sqrt{r^2_{(h)}(\hat{\beta})}, \sqrt{e^2_{(h)}}\}\right\}, \qquad (55)$$

So, taking into account that the difference in (49) attains value 1 iff (50) holds and -1 for (51), the last term of (17) can be written as

$$\tfrac{1}{\sqrt{n}}\sum_{i=1}^n e_i x_i \left[I\left\{-\sqrt{r^2_{(h)}(\hat{\beta})} \le e_i < -\sqrt{e^2_{(h)}}\right\} + I\left\{\sqrt{e^2_{(h)}} < e_i \le \sqrt{r^2_{(h)}(\hat{\beta})}\right\}\right] \qquad (56)$$

$$-\tfrac{1}{\sqrt{n}}\sum_{i=1}^n e_i x_i \left[I\left\{-\sqrt{e^2_{(h)}} \le e_i < -\sqrt{r^2_{(h)}(\hat{\beta})}\right\} + I\left\{\sqrt{r^2_{(h)}(\hat{\beta})} < e_i \le \sqrt{e^2_{(h)}}\right\}\right] \qquad (57)$$

$$+I\left\{-\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i \le e_i < \min\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{e^2_{(h)}}\}\right\}$$
$$-I\left\{\max\{\sqrt{e^2_{(h)}}, \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\} < e_i \le \max\{\sqrt{r^2_{(h)}(\hat{\beta})}, \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\}\right\}$$
$$+I\left\{\max\{\sqrt{r^2_{(h)}(\hat{\beta})}, \sqrt{e^2_{(h)}}\} < e_i \le \sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\right\}$$
$$+I\left\{\max\{-\sqrt{e^2_{(h)}}, -\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\} \le e_i < \max\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{r^2_{(h)}(\hat{\beta})}+\Delta_i\}\right\}$$

$$-I\left\{\max\{-\sqrt{r^2_{(h)}(\hat\beta)}, -\sqrt{e^2_{(h)}}\} \le e_i < -\sqrt{r^2_{(h)}(\hat\beta)+\Delta_i}\right\}$$
$$+I\left\{\mathrm{mim}\{\sqrt{r^2_{(h)}(\hat\beta)}, \sqrt{r^2_{(h)}(\hat\beta)+\Delta_i}\} < e_i \le \min\{\sqrt{e^2_{(h)}}, \sqrt{r^2_{(h)}(\hat\beta)+\Delta_i}\}\right\}$$
$$-I\left\{\sqrt{r^2_{(h)}(\hat\beta)+\Delta_i} < e_i \le \min\{\sqrt{r^2_{(h)}(\hat\beta)}, \sqrt{e^2_{(h)}}\}\right\}. \tag{58}$$

Taking into account the tables in the proof of Assertion A.2, we can observe that for each i there is at most one of all indicators in (58) equal to one. Nevertheless, let us start with the term in (56). Since

$$I\left\{-\sqrt{r^2_{(h)}(\hat\beta)} \le e_i < -\sqrt{e^2_{(h)}}\right\} + I\left\{\sqrt{e^2_{(h)}} < e_i \le \sqrt{r^2_{(h)}(\hat\beta)}\right\} = I\left\{e^2_{(h)} \le e_i^2 < r^2_{(h)}(\hat\beta)\right\},$$

(56) can be written as

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n e_i x_i^{\mathrm{T}} I\left\{e^2_{(h)} \le e_i^2 < r^2_{(h)}(\hat\beta)\right\}. \tag{59}$$

Let us observe that the indicator in (59) depends, what concerns e_i, only on its square. Moreover, let us recall (once again) that

$$\hat\beta^{(\mathrm{LTS},n,h)} = \underset{\beta\in\mathcal{K}}{\arg\min}\ \sum_{i=1}^h r^2_{(i)}(\beta) \tag{60}$$

(see (2)), i. e. $\hat\beta^{(\mathrm{LTS},n,h)}$ depends only on the order statistics of the squared residuals. Now let us consider any (but fix) $\omega_0 \in O_1$ (for O_1 see the remark at the beginning of the proof, the fifth and sixth line of proof). Similar considerations as we have carried out at the start of the paper then show that there is an h-tuple of indices, say $i_{1,\omega_0}, i_{2,\omega_0}, \ldots, i_{h,\omega_0}$ such that

$$\underset{\beta\in B(\beta^0,1)}{\arg\min}\ \sum_{i=1}^h r^2_{(i)}(\beta) = \sum_{j=1}^h r^2_{i_{j,\omega_0}}(\hat\beta^{(\mathrm{LTS},n,h)}(\omega_0)). \tag{61}$$

Let us denote for a moment this value of $\hat\beta^{(\mathrm{LTS},n,h)}$ by $\hat\beta^{(\omega_0)}$. The corresponding squared residuals are (in this notation) equal to

$$r_i^2(\hat\beta^{(\omega_0)}) = \left(Y_i - x_i^{\mathrm{T}}\hat\beta^{(\omega_0)}\right)^2. \tag{62}$$

In other words, if we select any other $\beta \in B(\beta^0, 1)$ and any other h-tuple of indices we obtain the sum of squared residuals equal to or larger than the sum on the right hand side of (61). Now, let us consider instead of values of disturbances at point ω_0, i. e. instead of $e_1(\omega_0), e_2(\omega_0), \ldots, e_n(\omega_0)$, the values $e_1^* = -e_1(\omega_0), e_2^* = -e_2(\omega_0), \ldots, e_n^* = -e_n(\omega_0)$. Then values of the response variable will be

$$Y_i^* = x_i^{\mathrm{T}}\beta^0 + e_i^* = x_i^{\mathrm{T}}\beta^0 - e_i(\omega_0). \tag{63}$$

Consequently, the squared values of residuals for $\beta^{(1)} = \beta^0 - \left(\hat{\beta}^{(\omega_0)} - \beta^0\right)$ will be

$$\begin{aligned} r_i^2(\beta^{(1)}) &= \left(Y_i^* - x_i^{\mathrm{T}}\beta^{(1)}\right)^2 = \left[x_i^{\mathrm{T}}\beta^0 - e_i(\omega_0) - x_i^{\mathrm{T}}\left(\beta^0 - \hat{\beta}^{(\omega_0)} + \beta^0\right)\right]^2 \\ &= \left(Y_i - x_i^{\mathrm{T}}\hat{\beta}^{(\omega_0)}\right)^2. \end{aligned}$$

In other words, for the symmetric values of disturbances, i.e. for values $e_1^* = -e_1(\omega_0), e_2^* = -e_2(\omega_0), \ldots, e_n^* = -e_n(\omega_0)$, we have found that values of squared residuals (for $\beta = \beta^{(1)} = \beta^0 - (\hat{\beta}^{(\omega_0)} - \beta^0)$) are the same as the squared values of the "original" residuals and the "new" $\hat{\beta}^{(\mathrm{LTS},n,h)}$, namely $\beta^{(1)}$, is symmetric, around β^0, to the "original" $\hat{\beta}^{(\mathrm{LTS},n,h)} = \hat{\beta}^{(\omega_0)}$. So the interval given in the indicator in (59) is the same for "original" disturbances as well as for "symmetric" ones. So we conclude that the distributions of random variables $e_i x_i I\{e_{(h)}^2 \le e_i^2 < r_{(h)}^2(\hat{\beta})\}$, $i = 1, 2, \ldots, n$ are symmetric and (6) then implies that their mean values exist and are equal to zero. Now, due to the consistency of $\hat{\beta}^{(\mathrm{LTS},n,h)}$, applying Lemma 1 for any positive η we can find $n_\eta > n_\delta$ (see (38)) so that for all $n > n_\eta$

$$\left\{e_{(h)}^2 \le e_i^2 < r_{(h)}^2(\hat{\beta})\right\} \cap B_n \cap C_n \subset \left\{u_\alpha^2 - n^{-\frac{1}{2}} K^{(10)} \cdot J^{-1} \le e_i^2 < u_\alpha^2 + \eta\right\} \quad (64)$$

(let us recall that by J we have denoted the upper bound of the density $f(z)$; for $K^{(10)}$, B_n and C_n see the part of this proof between (36) and (38)). Since η was arbitrary, (64) implies that for any positive ν, there is $n_\nu > n_\eta$ so that for all $n > n_\nu$

$$\operatorname{var}\left(e_i x_i I\left\{e_{(h)}^2 \le e_i^2 < r_{(h)}^2(\hat{\beta})\right\} \cdot I_{B_n \cap C_n}\right) < \nu \cdot (1 + \eta). \quad (65)$$

Considerations, similar to those which produced (32) (of course, with a condition, say, $\mathcal{C}(z, y)$ which assumes, in addition to $\mathcal{C}(z)$ also $r_{(h)}^2(\hat{\beta}) = y$) allow to apply conditional Tchebyshev inequality on the triangular array of random variables

$$\left\{\left\{e_i x_i^{\mathrm{T}} I\left\{e_{(h)}^2 \le e_i^2 < r_{(h)}^2(\hat{\beta})\right\}\right\}_{i=1}^{n}\right\}_{n=1}^{\infty}, \quad (66)$$

and due to the fact that ν was arbitrary we can conclude that (56) is $O_p(1)$. Along the same lines we find that the same is true for (57). Hence both these terms "can be moved" on the left hand side of (17).

It remains to study (58). Unfortunately, the terms in (58) have not mean values equal to zero. Hence we have to study the terms of the type

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Bigg\{ e_i x_i \Bigg[& I\left\{ -\sqrt{r_{(h)}^2(\hat{\beta})} + \Delta_i \le e_i < \min\{-\sqrt{r_{(h)}^2(\hat{\beta})}, -\sqrt{e_{(h)}^2}\}\right\} \\ & - \mathsf{E} I\left\{ -\sqrt{r_{(h)}^2(\hat{\beta})} + \Delta_i \le e_i < \min\{-\sqrt{r_{(h)}^2(\hat{\beta})}, -\sqrt{e_{(h)}^2}\}\right\}\Bigg]\Bigg\}. \end{aligned} \quad (67)$$

First of all, taking into account the tables from the proof of Assertion A.2 once again, we observe that the length of all intervals inside the indicators in (58) is less or equal to Δ_i. Then, due to Lemma 1, performing considerations about conditional probabilities of the events in question similar as in previous (see the proof of Theorem 1 of [4]) we find that

$$P\left(\{-\sqrt{r^2_{(h)}(\hat{\beta})} + \Delta_i \leq e_i < \min\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{e^2_{(h)}}\right\} \cdot I_{B_n \cap C_n})$$

$$= [f(u_\alpha) + o(1)] \cdot \Delta_i = [f(u_\alpha) + o(1)] \cdot x_i^{\mathrm{T}} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right).$$

Now, employing Lemma 1 once again we conclude that

$$|e_i - u_\alpha| \cdot I\left\{ -\sqrt{r^2_{(h)}(\hat{\beta})} + \Delta_i \leq e_i < \min\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{e^2_{(h)}}\}\right\} \cdot I_{B_n \cap C_n} = o(1)$$

(notice that due to the fact that on B_n $e^2_{h_n}$ as well as $r^2_{h_n}$ are bounded, the previous expression is really $o(1)$ not only $o_p(1)$; we shall need it for the next step). Finally we arrive at

$$\mathsf{E} e_i x_i I\left\{ -\sqrt{r^2_{(h)}(\hat{\beta})} + \Delta_i \leq e_i < \min\{-\sqrt{r^2_{(h)}(\hat{\beta})}, -\sqrt{e^2_{(h)}}\}\right\} \cdot I_{B_n \cap C_n}$$

$$= [2u_\alpha f(u_\alpha) + o(1)]\, x_i x_i^{\mathrm{T}} \left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right). \tag{68}$$

Of course, the analysis of the mean value of other terms from (58) is in fact the same. Then the analysis of the terms of type (67) is very similar to the analysis of (56).

Finally, let us denote the sum of mean values of terms given in (58) by T_n. Now recalling that we have observed that for any $i = 1, 2, \ldots, n$ at most one indicators in (58) is nonzero and taking into account (68), we can conclude that there is R_n so that

$$|R_n| \leq 2u_\alpha f(u_\alpha)$$

for all $n \in N$ and

$$T_n - Q \cdot (R_n + o_p(1)) \cdot \sqrt{n}\left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right) = o_p(1).$$

So, the analysis of (17) is finished. It yields that

$$O_p(1) = \{\, (Q + o_p(1)) \cdot [(1-\alpha) - R_n]\} \cdot \sqrt{n}\left(\hat{\beta}^{(\mathrm{LTS},n,h)} - \beta^0\right).$$

From the assumption that $f(z)$ is strictly decreasing on R^+, we have $1 - \alpha > 2u_\alpha f(u_\alpha)$ (it is immediately clear from the graph of $f(z)$). So, finally utilizing Lemma A.3 we conclude the proof. □

APPENDIX

Lemma A.1. Let for some $p \in N$, $\{\mathcal{V}^{(n)}\}_{n=1}^{\infty}$, $\mathcal{V}^{(n)} = \{v_{ij}^{(n)}\}_{i=1,2,\dots,p}^{j=1,2,\dots,p}$ be a sequence of $(p \times p)$ matrices such that for $i = 1, 2, \dots, p$ and $j = 1, 2, \dots, p$

$$\lim_{n\to\infty} v_{ij}^{(n)} = q_{ij} \qquad \text{in probability} \tag{A.69}$$

where $Q = \{q_{ij}\}_{i=1,2,\dots,p}^{j=1,2,\dots,p}$ is a fixed nonrandom regular matrix. Moreover, let $\{\theta^{(n)}\}_{n=1}^{\infty}$ be a sequence of p–dimensional random vectors such that

$$\exists\ (\varepsilon > 0)\ \forall\ (K > 0)\ \limsup_{n\to\infty} P\left(\|\theta^{(n)}\| > K\right) > \varepsilon. \tag{A.70}$$

Then

$$\exists\ (\ \delta > 0) \quad \forall\ (L > 0)$$

so that

$$\limsup_{n\to\infty}\ P\left(\left\|\mathcal{V}^{(n)}\theta^{(n)}\right\| > L\right) > \delta.$$

For the proof see [3].

Assertion A.1. Let $a, b \in (0, \infty)$, $\Delta \in R$. Then

$$\begin{aligned} I\{-a+\Delta \le e < -b\} &= I\{-a \le e < -b\} \\ &\quad - I\{\min\{-a, -a+\Delta\} \le e < \min\{-b, -a+\Delta\}\} \\ &\quad + I\{-a+\Delta \le e < \min\{-a, -b\}\}, \end{aligned} \tag{A.71}$$

$$(\min\{-a, -a+\Delta\}, \min\{-b, -a+\Delta\}) \subset (-a, -a+\Delta) \tag{A.72}$$

and

$$(-a+\Delta, \min\{-a, -b\}) \subset (-a+\Delta, -a)\,. \tag{A.73}$$

Further we have

$$\begin{aligned} I\{b < e \le a+\Delta\} &= I\{b < e \le a\} \\ &\quad - I\{\max\{b, a+\Delta\} < e \le \max\{a, a+\Delta\}\} \\ &\quad + I\{\max\{a, b\} < e \le a+\Delta\}, \end{aligned} \tag{A.74}$$

$$(\max\{b, a+\Delta\}, \max\{a, a+\Delta\}) \subset (a+\Delta, a) \tag{A.75}$$

and

$$(\max\{a, b\}, a+\Delta) \subset (a, a+\Delta)\,. \tag{A.76}$$

Similarly

$$I\{-b \le e < -a + \Delta\} = I\{-b \le e < a\}$$
$$-I\{\max\{-b, -a+\Delta\} \le e < \max\{-a, -a+\Delta\}\}$$
$$+I\{\max\{-a, -b\} \le e < -a+\Delta\}, \tag{A.77}$$

$$(\max\{-b, -a + \Delta\},\ ax\{-a, -a + \Delta\}) = \emptyset \tag{A.78}$$

and

$$(\max\{-a, -b\}, -a + \Delta) \subset (-a, -a + \Delta). \tag{A.79}$$

Finally

$$I\{a + \Delta < e \le b\} = I\{a < e \le b\}$$
$$-I\{\mathrm{mim}\{a, a + \Delta\} < e \le \min\{b, a + \Delta\}\}$$
$$+I\{a + \Delta < e \le \min\{a, b\}\}, \tag{A.80}$$

$$(\mathrm{mim}\{a, a + \Delta\}, \min\{b, a + \Delta\}) \subset (a, a + \Delta) \tag{A.81}$$

and

$$(a + \Delta, \min\{a, b\}) \subset (a + \Delta, a). \tag{A.82}$$

P r o o f. We shall consider successively all possible cases. Let us start with (A.71) and let us abbreviate the left hand side by

$$T_0 = I\{-a + \Delta \le e < -b\}$$

and the terms of the right hand side by

$$T_1 = I\{-a \le e < -b\}, \quad T_2 = I\{\min\{-a, -a + \Delta\} \le e < \min\{-b, -a + \Delta\}\}$$

and

$$T_3 = I\{-a + \Delta \le e < \min\{-a, -e\}\}.$$

Then

Table A1.

$a \le b$				
	T_0	T_1	T_2	T_3
$-b \le e \le -a \le -a+\Delta$	0	0	0	0
$-b \le -a \le e \le -a+\Delta$	0	0	0	0
$-b \le e \le -a+\Delta \le -a$	0	0	0	0
$-b \le -a+\Delta \le e \le -a$	0	0	0	0
$-a+\Delta \le e \le -b \le -a$	1	0	0	1
$-a+\Delta \le -b \le e \le -a$	0	0	0	0
$b < a$				
$-a \le e \le -b \le -a+\Delta$	0	1	1	0
$-a \le -b \le e \le -a+\Delta$	0	0	0	0
$-a \le e \le -a+\Delta \le -b$	0	1	1	0
$-a \le -a+\Delta \le e \le -b$	1	1	0	0
$-a+\Delta \le e \le -a \le -b$	1	0	0	1
$-a+\Delta \le -a \le e \le -b$	1	1	0	0

Similarly for (A.72) let us denote

$$I_1 = (\min\{-a, -a+\Delta\}, \min\{-b, -a+\Delta\})$$

and for (A.73)

$$I_2 = (-a+\Delta, \min\{-a, -b\}).$$

Then we have

Table A2.

$a \le b$		
	I_1	I_2
$-b \le -a \le -a+\Delta$	$(-a, -b) = \emptyset$	$(-a+\Delta, -b) = \emptyset$
$-b \le -a+\Delta \le -a$	$(-a+\Delta, -b) = \emptyset$	$(-a+\Delta, -b) = \emptyset$
$-a+\Delta \le -b \le -a$	$(-a+\Delta, -a+\Delta) = \emptyset$	$(-a+\Delta, -b) \subset (-a+\Delta, -a)$
$b < a$		
$-a \le -b \le -a+\Delta$	$(-a, -b) \subset (-a, -a+\Delta)$	$(-a+\Delta, -a) = \emptyset$
$-a \le -a+\Delta \le -b$	$(-a, -a+\Delta)$	$(-a+\Delta, -a) = \emptyset$
$-a+\Delta \le -a \le -b$	$(-a+\Delta, -a+\Delta) = \emptyset$	$(-a+\Delta, -a)$

Let us continue with (A.74). Abbreviating the left hand side again by

$$T_0 = I\{b < e \leq a + \Delta\}$$

and the terms of the right hand side by

$$T_1 = I\{b < e <\leq a\}, \quad T_2 = I\{\max\{b, a+\Delta\}, \max\{a, a+\Delta\}\}$$

and

$$T_3 = I\{\max\{a, b\}, a+\Delta\}\}.$$

Then

Table A3.

$a \leq b$				
	T_0	T_1	T_2	T_3
$a \leq e \leq b \leq a+\Delta$	0	0	0	0
$a \leq b \leq e \leq a+\Delta$	0	0	0	1
$a \leq e \leq a+\Delta \leq b$	0	0	0	0
$a \leq a+\Delta \leq e \leq b$	0	0	0	0
$a+\Delta \leq e \leq a \leq b$	0	0	0	0
$a+\Delta \leq a \leq e \leq b$	0	0	0	0
$b < a$				
$b \leq e \leq a \leq a+\Delta$	1	1	0	0
$b \leq a \leq e \leq a+\Delta$	1	0	0	1
$b \leq e \leq a+\Delta \leq a$	1	0	0	1
$b \leq a+\Delta \leq e \leq a$	0	1	1	0
$a+\Delta \leq e \leq b \leq a$	0	0	0	0
$a+\Delta \leq b \leq e \leq a$	0	1	1	0

Similarly for (A.75) let us denote

$$I_1 = (\max\{b, a+\Delta\}, \max\{a, a+\Delta\})$$

and for (A.76)

$$I_2 = (\max\{a, b\}, a+\Delta\}).$$

Then we have

Table A4.

$a \le b$		
	I_1	I_2
$a \le b \le a+\Delta$	$(a+\Delta, a+\Delta) = \emptyset$	$(b, a+\Delta) \subset (a, a+\Delta)$
$a \le a+\Delta \le b$	$(b, a+\Delta) = \emptyset$	$(b, a+\Delta) = \emptyset$
$a+\Delta \le a \le b$	$(b, a) = \emptyset$	$(b, a+\Delta) = \emptyset$
$b < a$		
$b \le a \le a+\Delta$	$(a+\Delta, a+\Delta) = \emptyset$	$(a, a+\Delta)$
$b \le e \le a+\Delta \le a$	$(a+\Delta, a+\Delta) = \emptyset$	$(a, a+\Delta) = \emptyset$
$b \le a+\Delta \le a$	$(a+\Delta, a)$	$(a, a+\Delta) = \emptyset$
$a+\Delta \le b \le a$	$(b, a) \subset (a+\Delta, a)$	$(a, a+\Delta) = \emptyset$

The rest of proof runs along similar lines. □

ACKNOWLEDGEMENT

This research was supported by by the Grant Agency of the Czech Republic under Grant 402/06/0408.

(Received May 7, 2004.)

REFERENCES

[1] P. Čížek: Analýza citlivosti k-krokových M-odhadů (Sensitivity analysis of k-step M-estimators, in Czech). Diploma Thesis, Czech Technical University, Prague 1996.

[2] E. Hewitt and K. Stromberg: Real and Abstract Analysis. Springer–Verlag, Berlin 1965.

[3] J. Á. Víšek: Sensitivity analysis M-estimates. Ann. Inst. Statist. Math. *48* (1996), 469–495.

[4] J. Á. Víšek: The least trimmed squares. Part I. Consistency. Kybernetika *42* (2006), 1–36.

[5] J. Á. Víšek: Kolmogorov–Smirnov statistics in linear regression. In: Proc. ROBUST 2006, submitted.

Jan Ámos Víšek, Department of Macroeconomics and Econometrics, Institute of Economic Studies, Faculty of Social Sciences, Charles University, Smetanovo nábřeží 6, 110 01 Praha 1 and Institute of Information Theory and Automation — Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 09 Praha 8. Czech Republic.

e-mail: visek@mbox.fsv.cuni.cz