ADAPTIVE ESTIMATION IN LINEAR REGRESSION MODEL

Part 2. Asymptotic normality

JAN ÁMOS VÍŠEK

Asymptotic representation of an adaptive estimator based on Beran’s idea of minimizing Hellinger distance is derived. It is shown that the estimator is asymptotically normal but not efficient. From the practical point of view the approach may be useful because it selects a model with distribution of residuals symmetric “as much as possible” (in the sense of Hellinger distance applied on $F(x)$ and $1 - F(x)$). It is not difficult to construct a numerical examples showing that sometimes it is the only way how to find proper model.

1. INTRODUCTION

This paper is the second part of the article “Adaptive estimation in linear regression model”. The reasons and clarifying discussions about the adaptive estimation may be found in the first part (cf. [20]). Also the notation of the present paper is the same as in the first part and the numeration of theorems and lemmas continues.

The proof of consistency of the adaptive estimator included in the first part of this paper has shown that the technique which leads to all results concerning the adaptive estimator is simple application of classical tools. The proof of Theorem 2 is of a similar character but much more longer. Therefore it will be divided into a sequence of steps, assertions and lemmas, proofs of which will be omitted. We shall show only as examples the proofs of Lemmas 3, 6 and 8. The reason for inclusion of the last three mentioned proofs is the fact that they represent the steps which yield a little unusual form of the result formulated in Theorem 2. All details can be found in technical report [17].

2. PRELIMINARIES

In this section we shall prepare tools for proving asymptotic normality of the estimator $\hat{\theta}_0(Y)$. To this end we restrict ourselves on such densities $g$ for which:

i) Fisher information is finite, i.e., the derivative of $g$ exists and $\int \frac{[\frac{d_y}{g(x)dy}]}{\frac{d_y}{g(x)dy}} dy < \infty$,
ii) $\sup_{y \in \mathbb{R}} |y'(y)| < K_5$

iii) $\frac{\Delta \alpha_n}{c_n} \to 0$ for $n \to \infty$

where $K_5$ is a finite constant. We need also an additional assumption on the kernel $w$.

We shall assume that

$$\int t^2 w(t) dt$$

is finite and denote it by $K_6$. (Moreover we shall assume that all assumptions made in the part 1 – see Sections 2, 3 and 5 – hold.)

**Remark 4.** Condition iii) seems at the first glance a little strange. But it is clear that for any $g$ with sufficiently smooth tails (even with arbitrarily heavy tails) we may for given $\{c_n\}_{n=1}^{\infty}$ find $\{c_n\}_{n=1}^{\infty}$ such that iii) holds. It may cause that $\{c_n\}_{n=1}^{\infty}$ will converge to zero rather slowly. Nevertheless, it is not inconsistent with other conditions which we assumed to be fulfilled (see, e.g., conditions for Theorem 1).

Moreover from the assumption $\int t^2 w(t) dt < \infty$ we have

$$\lim_{|t| \to \infty} t^2 w(t) = 0,$$

$$\lim_{|t| \to \infty} t w(t) = 0$$

and also

$$\lim_{|t| \to \infty} t w'(t) = 0.$$

Another consequence is that $\int |t| w(t) dt < \infty$ and hence also $\int |t| w^2(t) dt \leq K_1 \int |t| w(t) dt < \infty$.

Let us start with a simple assertion.

**Assertion 1.** For any $\beta \in \mathbb{R}^p$ and for all $k = 1, 2, \ldots, p$ we have

$$\frac{\partial}{\partial \beta_k} \int h_k(y, Y, \beta) h_n(-y, Y, \beta) dy =$$

$$= \int \left[ \frac{\partial h_k(y, Y, \beta)}{\partial \beta_k} \cdot h_n(-y, Y, \beta) + h_n(y, Y, \beta) \frac{\partial h_k}{\partial \beta_k} \right] dy.$$ 

Similarly it is not difficult to show that

$$\frac{\partial E_{\beta_0}(y, Y, \beta)}{\partial \beta_k} = \frac{1}{nc^k} \sum_{i=1}^n x^k_i \int w' \left( c_n^{-1} (y - z + X_i^T (\beta - \beta_0)) \right) g(z) dz.$$ 

Let us denote $\frac{\partial E_{\beta_0}(y, Y, \beta)}{\partial \beta_k}$ by $\frac{\partial E_{\beta_0}(y, Y, \beta)}{\partial \beta_k}$. 
Remark 5. Since
\[ \int \frac{\partial h_n(y, Y, \beta)}{\partial \beta_k} h_n(-y, Y, \beta) \, dy = \int h_n(y, Y, \beta) \cdot \frac{\partial h_n(-y, Y, \beta)}{\partial \beta_k} \, dy \]
we have \( \frac{\partial}{\partial \beta_k} \int h_n(y, Y, \beta) h_n(-y, Y, \beta) \, dy = 2 \int \frac{\partial h_n(y, Y, \beta)}{\partial \beta_k} h_n(-y, Y, \beta) \, dy \).

Lemma 3. For any \( \beta \in \mathbb{R}^p \) and \( k = 1, 2, \ldots, p \) we have
\[ \int \left[ \frac{\partial}{\partial \beta_k} h_n(y, Y, \beta) - \frac{\partial}{\partial \beta_k} E^1 g_n(y, Y, \beta) h_n(y) \right]^2 = O_p \left( n^{-1} c_n^2 \right). \]

Proof. To prove the assertion of the lemma let us write
\[ \frac{\partial}{\partial \beta_k} h_n(y, Y, \beta) - \frac{\partial}{\partial \beta_k} E^1 g_n(y, Y, \beta) h_n(y) = \]
\[ = \frac{1}{2} E^1 g_n(y, Y, \beta) h_n(y) \left[ \frac{\partial}{\partial \beta_k} g_n(y, Y, \beta) - \frac{\partial}{\partial \beta_k} E g_n(y, Y, \beta) \right] - \]
\[ - \frac{1}{2} g_n^2(y, Y, \beta) \frac{\partial}{\partial \beta_k} g_n(y, Y, \beta) \left\{ E^1 g_n(y, Y, \beta) \left[ g_n^2(y, Y, \beta) - \right] \right\} - \]
\[ - E^1 g_n(y, Y, \beta) - g_n(y, Y, \beta) \right)^2 + g_n^2(y, Y, \beta) - E^1 g_n(y, Y, \beta) \right) h_n(y). \]  

So we have arrived at the following inequality
\[ E \left\{ \frac{\partial}{\partial \beta_k} h_n(y, Y, \beta) - \frac{\partial}{\partial \beta_k} E^1 g_n(y, Y, \beta) h_n(y) \right\}^2 \leq \]
\[ \leq 3E \left[ E^{-1} g_n(y, Y, \beta) \right]^2 \left[ \frac{\partial}{\partial \beta_k} g_n(y, Y, \beta) - \frac{\partial}{\partial \beta_k} E g_n(-y, Y, \beta) \right]^2 \]
\[ + g_n^2(y, Y, \beta) \left[ \frac{\partial}{\partial \beta_k} g_n(y, Y, \beta) \right] \left\{ E^{-1} g_n(y, Y, \beta) \right\} - \]
\[ \left\{ g_n^2(y, Y, \beta) - E^1 g_n(y, Y, \beta) \right) \right]^2 + \left\{ g_n^2(y, Y, \beta) - E^1 g_n(y, Y, \beta) \right) \right]^2 \right\} h_n^2(y). \]

Now
\[ E^{-1} g_n(y, Y, \beta) \cdot E \left[ \frac{\partial}{\partial \beta_k} g_n(y, Y, \beta) - \frac{\partial}{\partial \beta_k} E g_n(y, Y, \beta) \right]^2 = \]
\[ = \frac{1}{n^2 c_n^2} E^{-1} g_n(y, Y, \beta) c_n^2(y) \cdot E \left[ \frac{\partial}{\partial \beta_k} \sum_{i=1}^n \left\{ w \left( c_n^{-1} (y - (Y_i - X_i \beta)) \right) - \right\} \right]^2 = \]
\[ = \frac{1}{n^2 c_n^2} E^{-1} g_n(y, Y, \beta) \cdot E \sum_{i=1}^n \left\{ w \left( c_n^{-1} (y - (Y_i - X_i \beta)) \right) - \right\} = \]}
\[ -\mathbb{E}w'(c_1^{-1}(y - (Y_i - X_i^T \beta))) z_{ik}^2 = \]
\[ \frac{1}{n^2 c_n^2} \mathbb{E}^{-1} g_n(y, Y, \beta) \sum_{i=1}^n \mathbb{E} \left\{ w'(c_1^{-1}(y - (Y_i - X_i^T \beta))) \right\} x_{ik} \]
\[ \leq \frac{1}{n^2 c_n^2} \mathbb{E}^{-1} g_n(y, Y, \beta) \sum_{i=1}^n x_{ik} \mathbb{E} \left\{ w'(c_1^{-1}(y - (Y_i - X_i^T \beta))) \right\}^2 \]
\[ = \frac{1}{n^2 c_n^2} \mathbb{E}^{-1} g_n(y, Y, \beta) \sum_{i=1}^n x_{ik} \mathbb{E} \left\{ \frac{w'(c_1^{-1}(y - (Y_i - X_i^T \beta)))}{w(c_1^{-1}(y - (Y_i - X_i^T \beta)))} \right\} \]
\[ \leq \frac{1}{n c_n^2} \sup_{z \in K} \sup_{w(z)} \left\{ \frac{w(z)^2}{w(z)} \right\} \sum_{i=1}^n x_{ik} \mathbb{E}^{-1} g_n(y, Y, \beta) \mathbb{E} \left\{ \frac{1}{n c_n^2} \sum_{i=1}^n w(c_1^{-1}(y - (Y_i - X_i^T \beta))) \right\} \]
\[ = \frac{1}{n c_n^2} \sup_{z \in K} \sup_{w(z)} \left\{ \frac{w(z)^2}{w(z)} \right\} \cdot \sup_{z \in K} \mathbb{E}^{-1} g_n(y, Y, \beta) \cdot \mathbb{E} g_n(y, Y, \beta). \]

Since \( \sup_{z \in K} \frac{w(z)^2}{w(z)} \leq \sup_{z \in K} \frac{w(z)^2}{w(z)} \sup_{z \in K} w(z) \) the last expression is of order \( O(n^{-1} c_n^2) \).

Similarly
\[ \mathbb{E}^{-1} g_n(y, Y, \beta) \cdot \mathbb{E} \left[ g_2^2(y, Y, \beta) - \mathbb{E} g_n(y, Y, \beta) \right] \leq \]
\[ \leq \mathbb{E}^{-1} g_n(y, Y, \beta) \cdot \mathbb{E} \left\{ \left[ g_2^2(y, Y, \beta) - \mathbb{E} g_n(y, Y, \beta) \right]^2 \right\} \cdot \left[ g_1^2(y, Y, \beta) + \mathbb{E}^1 g_n(y, Y, \beta) \right]^2 \]
\[ = \mathbb{E}^{-1} g_n(y, Y, \beta) \cdot \mathbb{E} [g_n(y, Y, \beta) - \mathbb{E} g_n(y, Y, \beta)]^2 \]
and using similar steps as in the proof of Lemma 1 we shall show that this expression is small.

Finally,
\[ g_n^{-1}(y, Y, \beta) \cdot \left. \frac{\partial}{\partial \beta} g_n(y, Y, \beta) \right| = \]
\[ = \left[ \sum_{i=1}^n w(c_1^{-1}(y - (Y_i - X_i^T \beta))) \right]^{-1} \cdot \sum_{i=1}^n x_{ik} w'(c_1^{-1}(y - (Y_i - X_i^T \beta))) \cdot c_1^{-1} \leq \]
\[ \leq c_n^{-1} \sup_{z \in K} \sup_{z \in K} \frac{w(z)}{w(z)} \leq c_n^{-1} K_1 \cdot K_2 \]  
(2)
where we have used an inequality
\[ \frac{a_1 + a_2}{b_1 + b_2} \leq \max \left\{ \frac{a_1}{b_1}, \frac{a_2}{b_2} \right\} \]
valid for \( b_1, b_2 > 0 \).
Using all derived inequalities together with
\[ E\left[ \frac{1}{t}(y, Y, \beta) - E_1 g_n(y, Y, \beta) \right]^2 \leq (nc_n)^{-1} \sup_{z \in S} w(z), \]
see [20], last line before Lemma 2, the assertion of the lemma follows.

**Remark 6.** Notice that
\[
\frac{\partial}{\partial \beta_k} E_1 g_n(-y, Y, \beta) = \frac{1}{2} \left\{ \sum_{i=1}^n w(c_n(-y - z + X_i^T(\beta - \beta_0))) g(z) \, dz \right\}^\frac{1}{2}
\]
\[
= \frac{1}{2} \sum_{i=1}^n \left\{ \sum_{i=1}^n \int w(c_n(-y - z + X_i^T(\beta - \beta_0))) x_i \, g(z) \, dz \right\}^\frac{1}{2}
\]
\[
= \frac{1}{2} \sum_{i=1}^n \left\{ \sum_{i=1}^n \int w(c_n(-y + z - X_i^T(\beta - \beta_0))) x_i \, g(z) \, dz \right\}^\frac{1}{2}
\]
\[
= \frac{1}{2} \sum_{i=1}^n \left\{ \sum_{i=1}^n \int w(c_n(-y + z - X_i^T(\beta - \beta_0))) x_i \, g(t) \, dt \right\}^\frac{1}{2}
\]

It gives
\[
\frac{\partial}{\partial \beta_k} E_1 g_n(-y, Y, \beta) \Bigg|_{\beta = \beta_0} = \frac{\partial}{\partial \beta_k} E_1 g_n(y, Y, \beta) \Bigg|_{\beta = \beta_0}
\]

In a similar way we can show that
\[ E_1 g_n(-y, Y, \beta_0) = E_1 g_n(y, Y, \beta_0). \]

The last equality has to be used to prove the next lemma. That is why this lemma holds only for \( \beta = \beta_0 \).

**Lemma 4.** For any \( k, \ell = 1, 2, \ldots, p \) we have
\[
\int \left\{ h_n(-y, Y, \beta_0) \cdot \frac{\partial^2}{\partial \beta_k \partial \beta_\ell} h_n(y, Y, \beta) - E_1 g_n(-y, Y, \beta_0) \cdot \frac{\partial}{\partial \beta_k} E_1 g_n(y, Y, \beta_0) \beta_\ell(y) \right\}^2 \, dy = O_p(n^{-1} c_n^2 a_n).
\]

**Lemma 5.** For any \( \beta \in \mathbb{R}^p \) and \( k, \ell = 1, 2, \ldots, p \) we have
\[
\int \left\{ \frac{\partial}{\partial \beta_k} h_n(y, Y, \beta) \cdot \frac{\partial}{\partial \beta_\ell} h_n(-y, Y, \beta) - E_1 g_n(y, Y, \beta) \cdot \frac{\partial}{\partial \beta_k} E_1 g_n(-y, Y, \beta) \beta_\ell(y) \right\}^2 \, dy = O_p(n^{-1} c_n^2 a_n).
Lemma 6. For any $k, \ell = 1, 2, \ldots, p$

\[ \int \frac{\partial^2 E_{g_k}(y, Y, \beta^0)}{\partial y_\ell \partial y_k} h_k(y) dy = o(1). \]

Proof. The absolute value of above given integral is not larger than

\[ \frac{1}{n c_n^2} \sum_{i=1}^{n} |x_i x_i| \left| \int \int w''(t) (g(z) - g(z_0)) h_k(y) dy \right| dt = \]

\[ = \frac{1}{n c_n^2} \sum_{i=1}^{n} |x_i x_i| \left| \int \int w''(t) (g(y - t c_n) - g(y)) dy \right| dt \leq \]

\[ \leq \frac{1}{n c_n^2} \sum_{i=1}^{n} |x_i x_i| \left| \int \int w''(t) (g(y - t c_n) - g(y)) dy \right| dt + \]

\[ + \frac{1}{n c_n^2} \sum_{i=1}^{n} |x_i x_i| \left| \int \int w''(t) (g(y - t c_n) - g(y)) dy \right| dt. \]  \hspace{1cm} (3)

Let us consider the first integral. It is equal to

\[ \frac{1}{n c_n^2} \sum_{i=1}^{n} |x_i x_i| \left| \int w''(t) \left( G(a_n - t c_n) - G(-a_n - t c_n) \right) dt \right| = \]

\[ = \frac{1}{n c_n^2} \sum_{i=1}^{n} |x_i x_i| \left| \int w''(t) \left( g(a_n) - g(-a_n) \right) - \left[ g(a_n) - g(-a_n) \right] t c_n + \right| + \]

\[ + \left[ t c_n - g'(\zeta_n) - g'(-\zeta_n) \right] c_n^2 \] \hspace{1cm} (4)

where $\zeta_n \in (\min \{-a_n - t c_n, a_n\}, \max \{-a_n - t c_n, -a_n\})$ and $\zeta_n \in (\min \{a_n - t c_n, a_n\}, \max \{a_n - t c_n, a_n\})$. Since \(\int w''(t) dt = \int [w''(t)]^2 dt = 0\) we have

\[ \int w''(t) \left[ G(a_n) - G(-a_n) \right] dt = 0. \]

Similarly due to \(g(a_n) = g(-a_n)\) we have

\[ \int w''(t) \left[ g(a_n) - g(-a_n) \right] dt = 0. \]

Remember that \(n^{-1} \sum_{i=1}^{n} |x_i x_i| < K_2^2\). So, to finish the proof, we need to show that

\[ c_n \int \left[ g'(\zeta_n) - g'(-\zeta_n) \right] w''(t) c_n^2 dt \]

is small. It may be done as follows. Let us fix some $\varepsilon > 0$ and find $K$ so that

\[ \int_{|t|>K} |w''(t)|^2 dt < \frac{\varepsilon}{4K^2}. \]
(It is possible, because \( \int |w''(t)|^2 dt < \int |w'(t)|^2 dt < K \int |w(t)|^2 dt < \infty. \))

Then we have

\[
\left| \int_{|t| < K} [g'(\xi) - g'(\zeta)] w''(t) \, dt \right| \leq 2 \cdot K \cdot \frac{\varepsilon}{4K^3} = \frac{\varepsilon}{2}.
\]

Now we shall estimate that part of integral which is over \( \{t : |t| < K\} \). Due to \( I(g) = \int \frac{w''(t)}{g(t)} dt \) being finite we have \( \lim_{|t| \to \infty} \frac{w''(t)}{g(t)} = 0 \) and due to fact that \( \lim_{|t| \to \infty} g(t) \) is also zero we have

\[
\lim_{|t| \to \infty} |g'(t)| = 0.
\]

Denote by \( Q \) the integral \( \int |w''(t)|^2 dt \). Due to (4) we may find \( L > 0 \) so that for any \( |y| > L \) we have \( |g'(y)| < \frac{\varepsilon}{4Q} \). Finally find \( n_0 \in \mathcal{N} \) so that for any \( n \geq n_0 \) we have \( c_n > 2L \) and \( c_n \cdot K < L \). Then for any such \( n \) we have \( |g'(\xi_n)| < \frac{\varepsilon}{4Q} \) as well as \( |g'(\zeta_n)| < \frac{\varepsilon}{4Q} \). Hence

\[
\left| \int_{|t| < K} [g'(\xi_n) - g'(\zeta_n)] w''(t) \, dt \right| \leq 2 \cdot \frac{\varepsilon}{4Q} \cdot Q = \frac{\varepsilon}{2}.
\]

The proof for the second member in (3) is based on the Cauchy-Schwarz inequality and the fact that \( \int_{|t| > L} |g''(t)|^2 dt \) is finite.

\[\square\]

**Lemma 7.** For any \( k = 1, 2, \ldots, p \) we have

\[
\int \left[ \frac{\partial h_n(y, Y, \beta^0)}{\partial \beta_k} - \frac{\partial \mathbb{E} h_n(y, Y, \beta^0)}{\partial \beta_k} b_n(y) \right] \times

\times \left[ h_n(-y, Y, \beta^0) - \mathbb{E} h_n(-y, Y, \beta^0) b_n(y) \right] \, dy = O_p(n^{-1} c_n k_n).
\]

**Assertion 2.**

\[
\int \frac{\partial \mathbb{E} h_n(y, Y, \beta^0)}{\partial \beta_k} b_n(y) \, dy = 0.
\]

**Lemma 8.** Let \( n^{-1} c_n k_n \to 0 \). Then for any \( k = 1, 2, \ldots, p \) we have

\[
n^k \int \left\{ \frac{\partial h_n(y, Y, \beta^0)}{\partial \beta_k} - \frac{\partial \mathbb{E} h_n(y, Y, \beta^0)}{\partial \beta_k} b_n(y) \right\} \mathbb{E} h_n(y, Y, \beta^0) b_n(y) \, dy = O_p(1).
\]

**Proof.** Using equality (1) we obtain

\[
\frac{\partial h_n(y, Y, \beta^0)}{\partial \beta_k} - \frac{\partial \mathbb{E} h_n(y, Y, \beta^0)}{\partial \beta_k} b_n(y) = - \frac{\partial \mathbb{E} h_n(y, Y, \beta^0)}{\partial \beta_k} b_n(y) \mathbb{E} h_n(y, Y, \beta^0) + \frac{\partial b_n(y)}{\partial \beta_k} \mathbb{E} h_n(y, Y, \beta^0) = \frac{\partial b_n(y)}{\partial \beta_k} \left[ \frac{1}{2} \left( \frac{\partial h_n(y, Y, \beta^0)}{\partial \beta_k} - \frac{\partial \mathbb{E} h_n(y, Y, \beta^0)}{\partial \beta_k} \right) \right].
\]
Adaptive Estimation in Linear Regression Model, Part 2

\[
- \frac{1}{2} \frac{\partial g_n(y, Y, \beta_0)}{\partial \beta_k} g_n^{-1}(y, Y, \beta_0) \left[ g_n^{-1}(y, Y, \beta_0) - \mathbb{E} g_n(y, Y, \beta_0) \right]^2
- \frac{1}{2} \frac{\partial g_n(y, Y, \beta_0)}{\partial \beta_k} g_n^{-1}(y, Y, \beta_0) \mathbb{E} g_n(y, Y, \beta_0)
\times \left[ g_n^{-1}(y, Y, \beta_0) - \mathbb{E} g_n(y, Y, \beta_0) \right] \right) = \sum_{j=1}^{3} R_j.
\]

Let us consider at first \( R_1 \). We have

\[
\left| \int \left[ \frac{\partial g_n(y, Y, \beta_0)}{\partial \beta_k} - \frac{\partial \mathbb{E} g_n(y, Y, \beta_0)}{\partial \beta_k} \right] \tilde{b}_n^2(y) dy \right| \\
\leq \int_{a_n}^{\infty} \left[ \frac{\partial g_n(y, Y, \beta_0)}{\partial \beta_k} - \frac{\partial \mathbb{E} g_n(y, Y, \beta_0)}{\partial \beta_k} \right] dy + \\
\int_{a_n < b < a_n + 4} \left[ \frac{\partial g_n(y, Y, \beta_0)}{\partial \beta_k} - \frac{\partial \mathbb{E} g_n(y, Y, \beta_0)}{\partial \beta_k} \right] dy.
\]  

(5)

Let us study at first the first integral of the right-hand-side of the last inequality. Let us fix an \( \varepsilon > 0 \) and \( \delta > 0 \). Then a straightforward computation gives (notice the factor \( n \))

\[
P \left\{ n \int_{-a_n}^{a_n} \left[ \frac{\partial g_n(y, Y, \beta_0)}{\partial \beta_k} - \frac{\partial \mathbb{E} g_n(y, Y, \beta_0)}{\partial \beta_k} \right] dy \right| \geq \varepsilon \right) \leq \]

\[
\leq P \left\{ n \int_{-a_n}^{a_n} \left[ \frac{\partial g_n(y, Y, \beta_0)}{\partial \beta_k} - \frac{\partial \mathbb{E} g_n(y, Y, \beta_0)}{\partial \beta_k} \right] dy \right| \geq \varepsilon \right) \\
- \sum_{i=1}^{n} x_i \left[ w(c_n^{-1}(a_n - Y_i + X_i^T \beta_0)) - \int w(c_n^{-1}(a_n - z)) g(z) dz \right] \right| \geq \varepsilon \right) \\
< P \left\{ n^{-1} c_n^{-1} \sum_{i=1}^{n} x_i \left[ w(c_n^{-1}(a_n - Y_i + X_i^T \beta_0)) - \int w(c_n^{-1}(a_n - z)) g(z) dz \right] \right| \geq \varepsilon \right) \\
+ P \left\{ n^{-1} c_n^{-1} \sum_{i=1}^{n} x_i \left[ w(c_n^{-1}(a_n - Y_i + X_i^T \beta_0)) - \int w(c_n^{-1}(a_n - z)) g(z) dz \right] \right| \geq \frac{\varepsilon}{2} \right) \\
+ P \left\{ n^{-1} c_n^{-1} \sum_{i=1}^{n} x_i \left[ w(c_n^{-1}(a_n - Y_i + X_i^T \beta_0)) - \int w(c_n^{-1}(a_n - z)) g(z) dz \right] \right| \geq \frac{\varepsilon}{2} \right) \right).
\]  

(6)

Let us write \( e_i \) instead of \( Y_i - X_i^T \beta_0 \). Then the first probability is bounded by

\[
\frac{4}{\varepsilon^2 c_n} \mathbb{E} \left\{ \sum_{i=1}^{n} x_i \left[ w(c_n^{-1}(a_n - e_i)) - \int w(c_n^{-1}(a_n - z)) g(z) dz \right] \right)^2 \leq \]

\[
\leq \frac{4 \cdot K^2}{\varepsilon^2 c_n} \sum_{i=1}^{n} \mathbb{E} \left\{ w(c_n^{-1}(a_n - e_i)) - \int w(c_n^{-1}(a_n - z)) g(z) dz \right\}^2 \leq \]

\[
\leq \frac{4 \cdot K^2}{\varepsilon^2 c_n} \int w^2(c_n^{-1}(a_n - z)) g(z) dz = \frac{4 \cdot K^2}{\varepsilon^2 c_n} \int w^2(t) g(a_n - e_i) dt = \]

\[
= \frac{4 \cdot K^2}{\varepsilon^2 c_n} \int w^2(t) \left| g(a_n) - g'(a_n + c_n(a_n, t)) \right| dt.
\]
where $|\xi_n(a_n, t)| \leq c_n |t|$. Since $\frac{g(\xi_n)}{c_n} \to 0$ it follows
\[
\frac{4 \cdot K^2}{\varepsilon^2 C_n} g(a_n) \int w^2(t) dt \to 0
\]
(7)

\[
\left( \int w^2(t) dt \leq K_1 \int w(t) dt = K_1 \right).
\]
(8)

The integral
\[
\frac{4}{\varepsilon^2} K^2 \int g'(a_n + \xi_n(a_n, t)) w^2(t) dt
\]
may be bounded using the fact that $|\xi_n(a_n, t)| \leq c_n |t|$. Indeed, for any $L \in \mathcal{R}$
\[
\left| \int g'(a_n + \xi_n(a_n, t)) w^2(t) dt \right| = 
\left| \left\{ \int_{|t| > L} + \int_{|t| \leq L} \right\} g'(a_n + \xi_n(a_n, t)) w^2(t) dt \right|
\]
(9)

At first fix $M > 0$ so that for any $|y| > M$ we have
\[
|g'(y)| < \frac{d \varepsilon^2}{16 \cdot K_1 K_6}.
\]

Then select $n_0 \in \mathcal{N}$ and $L > 0$ so that for any $n \in \mathcal{N}, n \geq n_0$ it holds:

a) $\int_{|t| > L} |w^2(t)| dt < \frac{d \varepsilon^2}{16 \cdot K_1 K_6}$,

b) $a_n - c_n \cdot L > M$

c) $\frac{g(\xi_n)}{c_n} \leq \frac{d \varepsilon}{8 \cdot K_1 K_6}$

(see assumption iii) at the beginning of this part of paper). Now taking into account (7), (8) and (9) we see for $n \geq n_0$ that the first probability in (6) is bounded by $\frac{1}{\varepsilon}$. The second probability in (6) may be treated along the similar lines. Let us consider now the second member or right-hand-side of (5) (again notice $n^\frac{1}{2}$). Probability that this member is larger than $\varepsilon$ may be treated as follows.

\[
P\left( n^{-\frac{3}{2}} c_n^{-3} \sum_{i=1}^{n} \int_{a_n < |y| < c_n + L} x_{ik} \left[ w'(c_n^{-1}(y - e_i)) - E w'(c_n^{-1}(y - e_i)) \right] dy > \varepsilon \right) \leq \frac{1}{\varepsilon^2 n c_n^3} \mathbb{E} \left\{ \int_{a_n < |y| < c_n + L} \sum_{i=1}^{n} x_{ik} \left[ w'(c_n^{-1}(y - e_i)) - E w'(c_n^{-1}(y - e_i)) \right] dy \right\}^2 \leq \frac{1}{\varepsilon^2 n c_n^3} \mathbb{E} \left\{ \int_{a_n < |y| < c_n + L} \left( \sum_{i=1}^{n} x_{ik} \left[ w'(c_n^{-1}(y - e_i)) - E w'(c_n^{-1}(y - e_i)) \right] \right) dy \right\}^2 \]
which converges to zero for \( n \to \infty \). Hence \( n^1 R_1 = o_p(1) \). The same result one obtains for \( n^2 R_2 \) using inequality (2) together with idea which the proof of Lemma 1 was based on. Really, one has

\[
\left| \frac{1}{2} \frac{\partial g_n(y, Y, \beta)}{\partial \beta} g_n^{-1}(y, Y, \beta) \left[ g_n^\dagger (y, Y, \beta) - E_g n(y, Y, \beta) \right] \right|^2 \leq
\]

\[
\frac{1}{2} c_n^{-1} K_2 \cdot K_3 \cdot E^{-1} g_n(y, Y, \beta) \left[ g_n^\dagger (y, Y, \beta) - E_g n(y, Y, \beta) \right]^2 = O_p(n^{-1} c_n^2)
\]

(see proof of Lemma 1). It implies (under assumption of present lemma) that \( n^1 R_2 = o_p(1) \).

For the \( R_3 \) we may write (let us use a little abbreviated form because there cannot be any confusion)

\[
\frac{\partial g_n}{\partial \beta} E^{-1} g_n \left[ g_n^\dagger - E g_n \right] = \left\{ \frac{\partial g_n}{\partial \beta} g_n^{-1} E^{-1} g_n \left[ g_n - E g_n \right] + \right\} +
\]

\[
+ \left( \frac{\partial g_n}{\partial \beta} - \frac{\partial g_n}{\partial \beta} E^{-1} g_n \right) E_g n \left[ g_n^\dagger - E g_n \right] +
\]

\[
+ \frac{\partial E g_n}{\partial \beta} E^{-1} g_n \left[ g_n^\dagger - E g_n \right] = \sum_{j=1}^3 S_j.
\]

Let us start with the first right-hand-side member (obtained after carrying out appropriate multiplication). We shall use again (2). Hence to show that \( P(n^{1/2} S_j > \varepsilon) \to 0 \) for \( n \to \infty \) it is (more than) sufficient for \( T_n = n^{1/2} E^{-1/2} g_n \left[ g_n - E g_n \right] \) and \( V_n = n^{1/2} \left[ g_n^\dagger - E g_n \right] \) to prove that both converge to zero in probability. The Chebyshev inequality helps in both cases.

\[
P \left( |T_n| > \varepsilon \right) \leq \frac{n^{1/2}}{\varepsilon} E^{-1/2} g_n \left[ g_n - E g_n \right]^2 = O \left( n^{-1/2} c_n^{-1} \right)
\]

and

\[
P \left( |V_n| > \varepsilon \right) \leq \frac{n^{1/2}}{\varepsilon} E \left( \left[ g_n^\dagger - E g_n \right]^2 \right) \leq \frac{n^{1/2}}{\varepsilon} E^{-1/2} g_n \left[ g_n - E g_n \right]^2
\]

where we have used Lemma 1 and inequality \((a - b)^2 \leq a^{-2} (a^2 - b^2) \) valid for \( a \geq 0 \) and \( b > 0 \). A similar result may be obtained for \( n^{1/2} S_3 \). The last member, namely \( n^{1/2} S_3 \), stays on the left-hand-side of expression given in present lemma. That concludes the proof. \( \square \)

The following two lemmas have been proved in [1] but were not stated explicitly there.
Lemma 9 (Beran [1]).

\[ \lim_{n \to \infty} c_n^2 \int \frac{\left[ \int w'(c_n^{-1}(y-z))g(z)dz \right]^2}{\int w(c_n^{-1}(y-z))g(z)dz}dy = f(y). \]

Lemma 10 (Beran [1]).

\[ n^{1/2} \int \frac{\partial E^1 g_n(y, Y, \beta^0)}{\partial \beta_k} \left[ h_n(y, Y, \beta^0) - E g_n(y, Y, \beta^0)h_n(y) \right]h_n(y)dy = \]
\[ = \frac{1}{2} n^{1/2} \left\{ \sum_{i=1}^n \frac{x_{1i}}{n} \left[ \sum_{m=1}^n \gamma_i(y_i - X_i, \beta^0)g^{-1}(y_i - X_i, \beta^0) + o_p(1) \right. \right\}. \]

Proof. We shall present nearly literally Beran's proof. We may write

\[ n^{1/2} \int \frac{\partial E^1 g_n(y, Y, \beta^0)}{\partial \beta_k} \left[ h_n(y, Y, \beta^0) - E g_n(y, Y, \beta^0)h_n(y) \right]h_n(y)dy = \]
\[ = \frac{1}{2} n^{1/2} \left\{ \int \frac{\partial E^1 g_n(y, Y, \beta^0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta^0) \times \right\]
\[ \times \left[ g_n(y, Y, \beta^0) - E g_n(y, Y, \beta^0)h_n(y) \right]h_n^2(y)dy \]
\[ - \int \frac{\partial E^1 g_n(y, Y, \beta^0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta^0) \]
\[ \left[ h_n(y, Y, \beta^0) - E g_n(y, Y, \beta^0)h_n(y) \right]h_n^2(y)dy \}

(10)

Since

\[ \left| \frac{\partial E^1 g_n(y, Y, \beta^0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta^0) \right| \leq \]
\[ \leq \frac{c_n}{K_4 \int w(c_n^{-1}(y-z))g(z)dz} \leq \]
\[ \leq \frac{K_4 \int w(c_n^{-1}(y-z))g(z)dz}{w(c_n^{-1}(y-z))} \]
\[ \leq c_n^{-1} \cdot K_3 \cdot K_4 \]

and

\[ E [ h_n(y, Y, \beta^0) - E g_n(y, Y, \beta^0)h_n(y)]^2 \leq n^{-1} c_n^{-1} K_1 \cdot K_3 \]

(see the proof of Lemma 1) we obtain that the second integral of the right-hand-side of (10) is \( O_p(n^{-1} c_n) \) and after multiplication by \( n^{1/2} \) converges to zero in probability. Let us put

\[ s_n(y) = c_n^{-1} \int w(c_n^{-1}(y-z))g(z)dz \]
and $s(y) = g^1(y)$.

Then for the first integral of right-hand-side of (10) we have

$$\begin{align*}
\text{var} \left\{ n^2 \int \frac{\partial \mathbf{E} \mathbf{g}_n(y, Y, \beta_0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta_0) \{ g_n(y, Y, \beta_0) - \mathbf{E} g_n(y, Y, \beta_0) \} \mathbf{k}_y^2(y) dy \right\} \\
= c_n^2 \text{var} \left\{ n^2 \int \frac{\partial \mathbf{E} \mathbf{g}_n(y, Y, \beta_0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta_0) \\
\{ w(c_n^{-1}(y - Y + X^T \beta_0)) - \mathbf{E} w(c_n^{-1}(y - Y + X^T \beta_0)) \} \mathbf{k}_y^2(y) dy \right\} \\
\leq c_n^{-2} E \left\{ n^2 \int \frac{\partial \mathbf{E} \mathbf{g}_n(y, Y, \beta_0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta_0) \{ w(c_n^{-1}(y - Y + X^T \beta_0)) \mathbf{k}_y^2(y) dy \} \right\}^2 \\
= 2c_n^{-2} \left\{ \frac{\sum_{j=1}^n x_{ik}}{n} \right\}^2 \{ \mathbf{E} y_s^2(y) \} \left\{ \mathbf{E} w(c_n^{-1}(y - Y + X^T \beta_0)) \mathbf{k}_y^2(y) dy \} \right\} \\
\leq 2c_n^{-1} \left\{ \frac{\sum_{j=1}^n x_{ik}}{n} \right\}^2 \left\{ \int \left\{ \mathbf{E} y_s^2(y) \right\} \left\{ \frac{\sum_{j=1}^n x_{ik}}{n} \right\} \mathbf{k}_y^2(y) dy \right\} \\
= 2 \left\{ \frac{\sum_{j=1}^n x_{ik}}{n} \right\}^2 \left\{ \int \left\{ \mathbf{E} y_s^2(y) \right\} \mathbf{k}_y^2(y) dy \right\} \\
\leq 2 \left\{ \frac{\sum_{j=1}^n x_{ik}}{n} \right\}^2 \left\{ \mathbf{E} y_s^2(y) \right\} \int \left\{ \mathbf{E} y_s^2(y) \right\} dy. \\
\end{align*}$$

Let us denote by $W_{nk}$ the integral

$$n^2 \int \frac{\partial \mathbf{E} \mathbf{g}_n(y, Y, \beta_0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta_0) \{ g_n(y, Y, \beta_0) - \mathbf{E} g_n(y, Y, \beta_0) \} \mathbf{k}_y^2(y) dy.$$

Further, again following [1], let us denote

$$U_{nk} = n^2 \left\{ \frac{\sum_{j=1}^n x_{ik}}{n} \right\} \sum_{j=1}^n s'(Y_j - X_j^T \beta_0) s^{-1}(Y_j - X_j^T \beta_0).$$

Then we have

$$\text{var} U_{nk} = \left\{ \frac{\sum_{j=1}^n x_{ik}}{n} \right\}^2 \int \left\{ s'(y) \right\}^2 dy$$

and also

$$\text{cov} (W_{nk}, U_{nk}) = E \left\{ n^2 \int \frac{\partial \mathbf{E} \mathbf{g}_n(y, Y, \beta_0)}{\partial \beta_k} E^{-1} g_n(y, Y, \beta_0) \{ g_n(y, Y, \beta_0) - \mathbf{E} g_n(y, Y, \beta_0) \} \mathbf{k}_y^2(y) dy \right\}.$$
\[ \left\{ \frac{\sum_{k=1}^{n} x_k}{n} \right\} \sum_{j=1}^{n} s'(Y_j - X_j^T \beta_0) s^{-1}(Y_j - X_j^T \beta_0) \right\} = \\
= \mathbb{E} \left\{ \frac{1}{nc_n} \sum_{k=1}^{n} \frac{x_k}{n} \int \frac{\partial E_{\mathbf{g}_k}(y,Y_i^0)}{\partial \beta_k} E^{-1} g_{\mathbf{k}}(y,y_i^0) \right\} \\
- \sum_{j=1}^{n} \int [w_{i}(c_{i-1}^{-1}(y - Y_i - X_i^T \beta_0)) - \mathbb{E} w_{i}(c_{i-1}^{-1}(y - Y_i - X_i^T \beta_0))] b_i^2(y) dy \\
+ \sum_{j=1}^{n} s'(Y_j - X_j^T \beta_0) s^{-1}(Y_j - X_j^T \beta_0) \right\} = \\
= c_n^2 \sum_{i=1}^{n} \left\{ \frac{x_i}{n} \right\}^2 \int \left\{ \mathbb{E} w_{i}(c_{i-1}^{-1}(y - Y_i - X_i^T \beta_0)) \cdot E^{-1} w_{i}(c_{i-1}^{-1}(y - Y_i - X_i^T \beta_0)) \right\} \\
- \frac{1}{2} \sum_{i=1}^{n} \left\{ \frac{x_i}{n} \right\}^2 \int \left\{ \int \mathbb{E} w_{i}(c_{i-1}^{-1}(y - Y_i - X_i^T \beta_0)) u(t) \right\} dt \\
\left\{ \int \mathbb{E} w_{i}(c_{i-1}^{-1}(y - Y_i - X_i^T \beta_0)) s'(z) s(z) dz \right\} \\
\text{since } E s'(Y_i - X_i^T \beta_0) s^{-1}(Y_i - X_i^T \beta_0) = 0. \text{ The last expression is equal to} \\
c_n^2 \left\{ \sum_{i=1}^{n} \left\{ \frac{x_i}{n} \right\}^2 \int \left\{ s_i'(y) \cdot s_{i-1}^{-1}(y) \cdot w_{i}(c_{i-1}^{-1}(y - z)) b_i^2(y) dy \cdot s'(z) s(z) dz \right\} \right\} (11) \\
\text{Now let us put} \\
d_{n}(y) = c_{n-1}^{-1} \cdot \int w_{n}(c_{n-1}^{-1}(y - z)) s'(z) s(z) dz. \\
\text{Since } s'(y) \in L_2 \text{ there exists for every } \varepsilon > 0 \text{ a differentiable function } \varphi_{\varepsilon} \in L_2 \text{ such that } \\
\varphi_{\varepsilon}^2 \in L_2 \text{ and } \| s'(y) - \varphi_{\varepsilon} \| < \varepsilon, \text{ where } \| \cdot \| \text{ denotes the } L_2 \text{-norm}. \text{ Then put also} \\
d_{n}(y) = c_{n-1}^{-1} \cdot \int w_{n}(c_{n-1}^{-1}(y - z)) \varphi_{\varepsilon}(z) s(z) dz. \\
\text{By the Cauchy-Schwarz inequality we have} \\
\int d_{n}^2(y) dy \leq \int c_n^2 s_n^2(y) \left\{ \int w_{n}(c_{n-1}^{-1}(y - z)) s'(z) dz \right\} \left\{ \int w_{n}(c_{n-1}^{-1}(y - z)) [s'(z)]^2 dz \right\} dy \\
= c_n^2 \int \left\{ \int w_{n}(c_{n-1}^{-1}(y - z)) [s'(z)]^2 dz \right\} dy \\
= \int \| s'(z) \|^2 dz \leq \| s'(y) \|^2. \\
\text{Since also} \\
\| d_{n}(y) \|^2 = \int c_n^2 s_n^2(y) \left\{ \int w_{n}(c_{n-1}^{-1}(y - z)) s(z) dz \right\}^2 dy \\
= \int c_n^2 s_n^2(y) \left\{ \int w_{n}(c_{n-1}^{-1}(y - z)) s(z) dz \cdot c_n^1 \int w_{n}(c_{n-1}^{-1}(y - z)) [\varphi_{\varepsilon}(z)]^2 dz \right\} dy \\
= c_n^2 \int \left\{ \int w_{n}(c_{n-1}^{-1}(y - z)) [\varphi_{\varepsilon}(z)]^2 dz \right\} dv = \| \varphi_{\varepsilon} \|^2 (12)
and

$$d_{n,\epsilon}(y) = c_n^{-1} \cdot s_n^{-1}(y) \cdot \int w(c_n^{-1}(y - z)) \psi_\epsilon(z) s(z) dz =$$

$$s_n^{-1}(y) \int w(t) \psi_\epsilon(y - c_t) s(y - c_t) dt$$

which implies

$$\lim_{n \to \infty} d_{n,\epsilon}(y) = \psi_\epsilon(y),$$

it follows by Vitali’s theorem that

$$\lim_{n \to \infty} \int \psi_\epsilon(y) d_{n,\epsilon}(y) dy = ||\psi_\epsilon||^2.$$  \hspace{1cm} (13)

Now

$$||d_{n,\epsilon} - d_n||^2 = \int c_n^2 s_n^{-2}(y) \left\{ \int w(c_n^{-1}(y - z)) [\psi_\epsilon(z) - s'(z)] s(z) dz \right\}^2 dy \leq \int \left\{ c_n^2 s_n^{-2}(y) \int w(c_n^{-1}(y - z)) s^2(z) dz \right\}^2 dy \leq \int [\psi_\epsilon(z) - s'(z)]^2 dz = ||\psi_\epsilon - s'||^2 \leq \epsilon.$$

Hence

$$\left| \int \psi_\epsilon(y) d_{n,\epsilon}(y) dy - \int s'(y) d_n(y) dy \right| \leq \left| \int [\psi_\epsilon(y) - s'(y)] d_{n,\epsilon}(y) dy \right| + \left| \int s'(y) [d_{n,\epsilon}(y) - d_n(y)] dy \right| \leq \left\{ \int [\psi_\epsilon(y) - s'(y)]^2 dy \int d_{n,\epsilon}(y) dy \right\}^{\frac{1}{2}} + \left\{ \int [s'(y)]^2 dy \int [d_{n,\epsilon}(y) - d_n(y)]^2 dy \right\}^{\frac{1}{2}} \leq \epsilon \{ ||d_{n,\epsilon}|| + ||s'|| \} \leq \epsilon \{ ||\psi_\epsilon|| + ||s'|| \}.$$

This inequality and (13) imply that

$$\lim_{n \to \infty} \int s'(y) d_{n,\epsilon}(y) dy = ||s'||^2.$$  \hspace{1cm} (14)

(Really we may write

$$s' d_n = s'(d_n - d_{n,\epsilon}) + (s' - \psi_\epsilon) d_{n,\epsilon} + \psi_\epsilon (d_{n,\epsilon} - \psi_\epsilon) + (\psi_\epsilon - s') + [s']^2 + [s']^2$$
and value of the integral of any member of the right-hand-side, except the last one, can be bounded by some constant multiplied by \( \varepsilon \) which was fixed but arbitrary.) Finally using (11), (12) and (14) we obtain

\[
\lim_{n \to \infty} \text{cov} (W_{nk}, U_{nk}) = \left[ \sum_{k=1}^{n} \varepsilon_{nk} \right]^{2} \int [s'(y)]^{2} dy.
\]

\[
\text{cov} (W_{nk}, U_{nk}) = c_{nk}^{-1} \left[ \sum_{k=1}^{n} \varepsilon_{nk} \right]^{2} \int s'(y) s_{nk}(y) w(c_{nk}^{-1}(y-z)) b_{nk}^{2}(y) dy s'(z) s(z) dz.
\]

But,

\[
\left| \int \{ s'(y) d_{nk}(y) - s_{nk}'(y) d_{nk}(y) \} b_{nk}^{2}(y) dy \right| \leq \left\{ \int [s' - s_{nk}']^{2} dy \int d_{nk}^{2}(y) dy \right\}^{\frac{1}{2}} \to 0 \text{ for } n \to \infty
\]

and making use of (14) one obtains

\[
\lim_{n \to \infty} \int s_{nk}'(y) d_{nk}(y) dy = \| s'(y) \|^{2}.
\]

\[\square\]

3. ASYMPTOTIC NORMALITY

In this section we will give the main result of the paper. Let us summarize all assumptions we have made and we will need for the Theorem 2.

We have assumed that “the random errors” in model (1) are i.i.d. according to the d.f. \( G \) which has finite Fisher information \( I(g) \). It mean that the d.f. \( G \) is supposed to be twice differentiable. Denote the first and the second derivative by \( g \) and by \( g' \), respectively. Moreover \( g \) is assumed to be symmetric around zero. Then we have required the existence of constant \( K_{1}, \ldots, K_{6} \) such that for the kernel \( w \), the design matrix and the derivative of density \( g' \) we have

\[
\sup_{y \in \mathbb{R}} w(y) < K_{1}, \quad \sup_{y \in \mathbb{R}} \frac{w'(y)}{w(y)} < K_{2},
\]

\[
\sup_{y \in \mathbb{R}} \frac{w''(y)}{w(y)} < K_{3}, \quad \sup_{y \in \mathbb{R}} |z_{ij}| < K_{4},
\]

\[
\sup_{y \in \mathbb{R}} |g'(y)| < K_{5} \quad \text{and} \quad \int g^{2}(y) dy = K_{6} < \infty.
\]

For the bandwidths \( \{ c_{n} \} \) \( n \to \infty \) and the supports (given by a sequence \( \{ a_{n} \} \to \infty \)) of kernel estimate we need

\[
\lim_{n \to \infty} n c_{n}^{2} a_{n}^{-2p} = \infty \quad \text{and} \quad \lim_{n \to \infty} \theta(c_{n}) = 0.
\]
Basic conditions for identifiability of model were the following: For any $\delta > 0$ there are $\Delta \in (0, 1)$ and $K \in \mathbb{R}$ so that

$$\limsup_{n \to \infty} \, \sup_{\theta \in \Theta(\Delta, K)} \int E^\theta \left( \frac{1}{n} \sum_{i=1}^{n} \left( y_i - \beta \right) \left( \frac{1}{n} \sum_{j=1}^{n} \left( y_j - \beta \right) \right) dy < \Delta \right.$$ 

and

$$\limsup_{n \to \infty} \, \sup_{\theta \in \Theta(\Delta, K)} \int h_n(y, \beta) h_n(-y, \beta) dy < \Delta$$

in probability.

Let us write throughout this section $\hat{\beta}^n$ instead of $\hat{\beta}(Y)$ and for any function $F = F(\beta)$ write $\frac{\partial F(\beta)}{\partial \beta}$ instead of $\frac{\partial F(\beta)}{\partial \beta_n}$.

**Theorem 2.** Under the just summed up conditions we have for $\hat{\beta}^n$ the following asymptotic representation

$$n^{-1/2} f(y) \sum_{t=1}^{p} (\hat{\beta}^n_t - \beta_0) \int_0^1 \left( Y_i - X_i^T \beta \right)^{-1} (Y_i - X_i^T \beta_0) + o_p(1).$$

**Proof.** Since $\hat{\beta}^n$ maximizes

$$\int h_n(y, \beta) h_n(-y, \beta) dy$$

over all $\beta \in \mathbb{R}^p$, it follows (see Assertion 1 and Remark 5) that for every $n \in \mathbb{N}$ and $k = 1, 2, \ldots, p$

$$\int \frac{\partial h_n(y, \beta)}{\partial \beta_k} h_n(-y, \beta) dy = 0.$$  \hspace{1cm} (15)

Now expanding $\frac{\partial h_n(y, \beta)}{\partial \beta_k} h_n(-y, \beta)$ at the point $\beta_0$ to approximate $\frac{\partial h_n(y, \beta)}{\partial \beta_k} h_n(-y, \hat{\beta}^n)$ we obtain

$$\int \frac{\partial h_n(y, \hat{\beta}^n)}{\partial \beta_k} h_n(-y, \hat{\beta}^n) dy = \int \frac{\partial h_n(y, \beta_0)}{\partial \beta_k} h_n(-y, \beta_0) dy +$$

$$+ \sum_{t=1}^{p} \left\{ \int \frac{\partial^2 h_n(y, \beta_0)}{\partial \beta_k \partial \beta_t} h_n(-y, \beta_0) dy + \int \frac{\partial h_n(y, \beta_0)}{\partial \beta_k} \frac{\partial h_n(-y, \beta_0)}{\partial \beta_t} dy \right\}.$$  

where $\sup_{n, \beta_0} |R_{nt}| = O_p(1)$. Now successively using (15), Lemma 4 and 5 and multiplying the whole equality by $n^{1/2}$ we obtain

$$\sqrt{n} \int \frac{\partial h_n(y, \beta_0)}{\partial \beta_k} h_n(-y, \beta_0) dy =$$

$$= - \sum_{t=1}^{p} \left\{ \int \frac{\partial^2 g_n(y, \beta_0)}{\partial \beta_k \partial \beta_t} g_n(-y, \beta_0) + \int \frac{\partial g_n(y, \beta_0)}{\partial \beta_k} \frac{\partial g_n(-y, \beta_0)}{\partial \beta_t} g_n(-y, \beta_0) \right\}.$$
A straightforward computation gives
\[ \frac{\partial^2 \mathcal{E}_1 g_n(y, Y, \beta^0)}{\partial \beta_k \partial \beta_t} \frac{\partial \mathcal{E}_1 g_n(-y, Y, \beta^0)}{\partial \beta_k} = \]
\[ = \frac{1}{2} \frac{\partial \mathcal{E}_1 g_n(y, Y, \beta^0)}{\partial \beta_k} \left\{ \frac{\partial^2 \mathcal{E}_1 g_n(y, Y, \beta^0)}{\partial \beta_k \partial \beta_t} \right\} \mathcal{E}_1 g_n(-y, Y, \beta^0) - \frac{1}{2} \frac{\partial \mathcal{E}_1 g_n(y, Y, \beta^0)}{\partial \beta_k} \frac{\partial^2 \mathcal{E}_1 g_n(y, Y, \beta^0)}{\partial \beta_t} \]
Using Lemma 9 and 10 and denoting for any \( k = 1, 2, \ldots, p \) \( n^{-1} \sum_{r=1}^{n-1} \tilde{x}_r \) by \( \tilde{x}_k \) we arrive at

\[
\sum_{k=1}^{p} \sqrt{n}(\beta_k^* - \beta_k^0)\tilde{x}_k \tilde{x}_r \left\{ I_{(\alpha)} + o(1) + \sum_{j=1}^{p} (\beta_j^* - \beta_j^0) \cdot (\mathbb{R})_{\ell} + O_p(n^{-1/2} a_n) \right\}
\]

\[
= n^{-1} \tilde{x}_k \sum_{i=1}^{n} g' \left( Y_i - X_i^T \beta_0^0 \right) g^{-1} \left( Y_i - X_i^T \beta_0^0 \right) + o_P(1).
\]

From it follows that for any \( \ell = 1, 2, \ldots, p \)

\[
\sqrt{n}(\beta_{\ell}^* - \beta_{\ell}^0) = O_p(1)
\]

and that concludes the proof.

4. NUMERICAL STUDY

A very first idea about numerical performance of adaptive estimator may be built up on the following tables. We have used well known Salinity and Stackloss data sets. Their description and explanation may be found in a lot of papers and books, e.g., [12] or [11].

Let us explain abbreviations in the following tables.

**LS** - denotes Least Squares estimate;

\( \hat{\beta}(.5) \) - regression quantiles for \( \alpha = .5 \);

\( \hat{\beta}_{PQ}(10) \) - the estimator is defined as follows: use a preliminary estimator \( \hat{\beta}_{\text{preliminary}} \) (in our case \( \hat{\beta}_{\text{preliminary}} = \frac{1}{2}(\hat{\beta}(1) + \hat{\beta}(9)) \) was used) and evaluate residuals; after trimming off 10% points having the largest values and 10% points having the smallest values of residuals apply \( LS \) to the rest;

\( \hat{\beta}_{KB}(15) \) - Trimmed Least Squares estimate after trimming off points according to regression quantiles \( \hat{\beta}(15) \) and \( \hat{\beta}(85) \);

Huber - \( M \)-estimate with \( \psi(x) = \text{sign} x \cdot \min(|x|, 1.25) \) and with 1.483 \( \cdot \) \( MAD \) as a scale estimate used for rescaling of residuals;

Andrews - \( M \)-estimate with \( \psi(x) = \sin(x) \cdot I_{(|x|<\pi / 2)} \) (\( MAD \) as a scale estimate was used);

LMS - Least Median of Squares (in fact model in which \( |x| + |x+y| \)-th order statistic of residuals was minimized);

LTS (Rousseeuw) - Least Trimmed Squares (in fact this estimate is \( \hat{\beta}_{PQ}(\alpha) \) where as the preliminary estimator serves \( LMS \));

Adaptive - adaptive estimator from this paper;

TLS (Adaptive) - Trimmed Least Square where trimming was according to Adaptive estimator and in both cases of the data sets four points were trimmed off. More precisely, when calculating results in the last line of the next tables for Salinity data the points 5, 16, 23 and 24 were trimmed off; while for Stackloss data the points 1, 3, 4 and 21 were excluded.
## SALINITY DATA

<table>
<thead>
<tr>
<th>Method</th>
<th>Estimates of coefficients</th>
<th>Intercept</th>
<th>Salting</th>
<th>Trend</th>
<th>H2O Flow</th>
</tr>
</thead>
<tbody>
<tr>
<td>LS</td>
<td></td>
<td>9.59</td>
<td>.777</td>
<td>-.026</td>
<td>-.205</td>
</tr>
<tr>
<td>(\hat{\beta}(.5))</td>
<td></td>
<td>14.21</td>
<td>.740</td>
<td>-.111</td>
<td>-.458</td>
</tr>
<tr>
<td>(\hat{\beta}_{RG}(.10))</td>
<td></td>
<td>14.49</td>
<td>.774</td>
<td>-.160</td>
<td>-.488</td>
</tr>
<tr>
<td>(\hat{\beta}_{KN}(.15))</td>
<td></td>
<td>9.69</td>
<td>.800</td>
<td>-.128</td>
<td>-.290</td>
</tr>
<tr>
<td>Huber</td>
<td></td>
<td>13.36</td>
<td>.756</td>
<td>-.094</td>
<td>-.439</td>
</tr>
<tr>
<td>Andrews</td>
<td></td>
<td>17.22</td>
<td>.733</td>
<td>-.196</td>
<td>-.578</td>
</tr>
<tr>
<td>LMS</td>
<td></td>
<td>36.70</td>
<td>.356</td>
<td>-.073</td>
<td>-1.298</td>
</tr>
<tr>
<td>LTS (Rousseeuw)</td>
<td></td>
<td>35.54</td>
<td>.436</td>
<td>-.061</td>
<td>-1.277</td>
</tr>
<tr>
<td>Adaptive</td>
<td></td>
<td>36.70</td>
<td>.367</td>
<td>-.071</td>
<td>-1.276</td>
</tr>
<tr>
<td>TLS (Adaptive)</td>
<td></td>
<td>30.28</td>
<td>.589</td>
<td>-.259</td>
<td>-1.091</td>
</tr>
</tbody>
</table>

### ACKNOWLEDGEMENT

The author expresses his thanks to Jaromír Antoch for his very valuable comments on the manuscript of paper which considerably improved its comprehensibility.

(Received June 19, 1990.)

### REFERENCES

Adaptive Estimation in Linear Regression Model, Part 2


RNDr. Jan Ámos Víšek, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation – Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8, Czechoslovakia.