Comment.Math.Univ.Carolin. 35,1 (1994)111–125 

# Sensitivity analysis of $M$-estimators of non-linear regression models

Rubio A.M., Quintana F., Víšek J. Án

*Abstract.* An asymptotic formula for the difference of the $M$-estimates of the regression coefficients of the non-linear model for all $n$ observations and for $n-1$ observations is presented under conditions covering the twice absolutely continuous $\varrho$-functions. Then the implications for the $M$-estimation of the regression model are discussed.

*Keywords:* $M$-estimation of non-linear regression models, the influence points

*Classification:* Primary 62F35; Secondary 62F12

## 1. Introduction

In the development of the theory of the linear regression models a considerable attention has been paid to the sensitivity analysis. Let us mention at least Cook and Weisberg(1982), Welsch (1982), Chatterjee and Hadi (1988) or Zvára (1989), among others. One of the important tools of the linear regression analysis (in detail explained below) was the formula describing a change of the coefficient estimates (or the studentized change of the estimates) when excluding one observation from the original data. Such a formula has been used to find out which of the points has the largest influence on the determination of the model. A similar formula is derived here for the non-linear regression scheme considering the $M$-estimation. Let us start with some basic notation to be able to explain the problem in question in detail.

Let $N$ denote the set of all positive integers, $R$ the real line, $R^\ell$ the $\ell$-dimensional Euclidean space ($\ell \in N$) and $(\Omega, \mathcal{A}, P)$ a probability space. Moreover, let for some fixed $p \in N$ and $q \in N$, $\beta^0 = (\beta_1^0, \beta_2^0, \ldots, \beta_p^0)^T$ (where "$T$" denotes the transposition) be the vector of the regression coefficients and $\{X_i\}_{i=1}^\infty, X_i : \Omega \to R^q$, be a sequence of independent and identically distributed random variables (i.i.d.r.v.). Finally, let $\{e_i\}_{i=1}^\infty, e_i : \Omega \to R$, be another sequence of i.i.d.r.v., independent from the sequence $\{X_i\}_{i=1}^\infty$. For a function $g : R^{q+p} \to R$ we shall consider (for all $i \in N$) the regression model

$$Y_i = g(X_i, \beta^0) + e_i. \tag{1}$$

1This paper was written while the author was visiting the Department of Mathematics of The University of Extremadura.

Let us denote by $K(x)$ the distribution function of $X_1$ and by $F(t)$ the distribution function of $e_1$ (by $f(t)$ will be denoted the density of $F(t)$ whenever we shall assume that it exists; moreover let $S_1$ denote the support of $K(x)$). We will be interested in the $M$-estimator of $\beta^0$ given as

$$\hat{\beta}^{(n)} = \mathrm{argmin}_{\beta \in R^p} \{ \sum_{i=1}^{n} \varrho(Y_i - g(X_i, \beta)) \} \tag{2}$$

where $\varrho : R \to R$ is assumed to be differentiable with an absolutely continuous derivative $\psi$. Let us denote the derivative of $\psi$ by $\psi'$ (at the points where it exists).

Specifying for $q = p$ and $g(X, \beta) = X^T \beta$ we obtain the linear regression model

$$Y_i = X_i^T \beta^0 + e_i, \qquad i = 1, 2, \ldots, n.$$

Let us denote by $X^{(n)}$ and $X^{(n-1,\ell)}$ the design matrices $(X_1, X_2, \ldots, X_n)^T$, $X_i \in R^p$ and $(X_1, X_2, \ldots, X_{\ell-1}, X_{\ell+1}, \ldots X_n)^T$, respectively, and the corresponding $LS$-estimators by $\hat{\beta}_{LS}^{(n)}$ and $\hat{\beta}_{LS}^{(n-1,\ell)}$. Comparing the normal equations for $n$ and $n-1$ observations we obtain

$$\hat{\beta}_{LS}^{(n-1,\ell)} - \hat{\beta}_{LS}^{(n)} = -\{[X^{(n-1,\ell)}]^T X^{(n-1,\ell)}\}^{-} X_\ell (Y_\ell - X_\ell^T \hat{\beta}_{LS}^{(n)}) \tag{3}$$

where $\{[X^{(n-1,\ell)}]^T X^{(n-1,\ell)}\}^{-}$ denotes a pseudoinverse to $\{[X^{(n-1,\ell)}]^T X^{(n-1,\ell)}\}$. From it follows that

$$\frac{\|\hat{\beta}_{LS}^{(n-1,\ell)} - \hat{\beta}_{LS}^{(n)}\|}{\sqrt{\mathrm{var}\left(\|\hat{\beta}_{LS}^{(n-1,\ell)} - \hat{\beta}_{LS}^{(n)}\|\right)}} = |Y_\ell - X_\ell^T \hat{\beta}_{LS}^{(n)}|. \tag{4}$$

So to find a point, exclusion of which implies the largest value of the studentized norm of change of estimates of the regression coefficients, we need just to look for the point(s) with the largest absolute value of the residual. Naturally, when we want to take into account also the position of data in space we will prefer to use (3) and the analysis will be a little more complicated. It may be of interest that when we want to analyze the data (and the model) from the point of view of the largest change in the prediction we find the same as above. In fact, for any $\tilde{X} \in R^p$ we obtain

$$\hat{Y}_{LS}^{(n-1,\ell)} - \hat{Y}_{LS}^{(n)} = \tilde{X}^T (\hat{\beta}_{LS}^{(n-1,\ell)} - \hat{\beta}_{LS}^{(n)})$$

and hence

$$\sup_{\|\tilde{X}\|=1} |\hat{Y}_{LS}^{(n-1,\ell)} - \hat{Y}_{LS}^{(n)}| = \|\hat{\beta}_{LS}^{(n-1,\ell)} - \hat{\beta}_{LS}^{(n)}\|.$$

(Similarly as in the case of the change of estimates of the regression coefficients described by (3) and (4) we may want for the prediction to take a position of data in the space also into account. Naturally, the analysis will be again a little more complicated. The present authors, however, believe that we should abandon invariance and prefer the position of data in the factor space only when there are very strong reasons for it.) The purpose of this paper is to establish formulae analogous to (3) and (4) for the $M$-estimators for the non-linear model. Since for the $M$-estimators we usually do not have analytic formulae for their evaluations but only asymptotic representations, our result will be also of the asymptotic type. The $LS$-estimator is under the assumption that $\mathsf{E}_F(e_1) = 0$ unbiased. For the $M$-estimators the situation is somewhat more complicated and hence we will simply assume that $\hat{\beta}^{(n)}$ is consistent, so that our result will be applicable on any consistent $M$-estimator. For the conditions guaranteeing consistency of the $M$-estimators in the non-linear regression see Liese and Vajda (1992) (do not be confused that the authors assume $\varrho$ to be twice continuously differentiable which is slightly stronger than our assumptions; in fact, they need this assumption only for deriving asymptotic normality, so that it is reasonable to consider our **Conditions B** below).

## 2. Asymptotic representation of difference between estimates of regression model

For any finite set $S = \{s_1, s_2, \dots s_k\} \subset R$ and $\alpha > 0$ put $S(\alpha) = \cup_{i=1}^{k} [s_i - \alpha, s_i + \alpha]$. We shall assume:

**Condition A.**

The estimator $\hat{\beta}^{(n)}$ is consistent in the following sense:

$$\forall (\delta > 0 \text{ and } \varepsilon > 0)\ \exists (n_0 \in N)\ \forall (n \in N,\ n \geq n_0 \text{ and } \ell = 1, 2, \dots, n)$$

$$P\left( \|\hat{\beta}^{(n)} - \beta^0\| > \delta \right) < \varepsilon$$

and

$$P\left( \|\hat{\beta}^{(n-1,\ell)} - \beta^0\| > \delta \right) < \varepsilon$$

where

$$\hat{\beta}^{(n-1,\ell)} = \operatorname{argmin}_{\beta \in R^p} \{ \sum_{i=1, i \neq \ell}^{n} \varrho(Y_i - g(X_i, \beta)) \}.$$

**Conditions B.**

(i) The function $\psi(z)$ and the derivative $\psi'(z)$ are uniformly continuous on $R$ and on $R \setminus \mathcal{C}$, respectively, where $\mathcal{C} = \{c_1, c_2, \dots, c_r\}$, $r$ being finite.

(ii) There is $\tau_0$ such that $F(z)$ has a continuous density $f(z)$ on $\mathcal{C}(\tau_0)$.

(iii) There is a finite $L$ such that $\sup_{z \in R} |\psi(z)| < L$ and $\sup_{z \in \mathcal{C}(\tau_0) \setminus \mathcal{C}} |\psi'(z)| < L$.

(iv) The mean value $\mathsf{E}_F \psi'(e_1) > 0$.

**Remark 1.** *Let us observe that due to the continuity of $f(x)$ on $\mathcal{C}(\tau_0)$, $f(x)$ is bounded there, let us say by $M < \infty$.*

**Remark 2.** *Due to the fact that $\psi$ is assumed bounded, the mean value of it exists. Let us assume that it is zero.*

**Remark 3. Conditions B** *essentially coincide with those of Hampel et al.* (1986), 7.2a, *under which a general class of tests of the linear model was studied. The reader who is interested in a heuristic discussion of these conditions may find it at this book.*

**Conditions C.**

(i) The function $g$ is in a neighbourhood of $\beta^0$ twice continuously (and uniformly with respect to $x \in S_1$) differentiable in the coordinates corresponding to the regression coefficients, i.e. there is $\delta_0 > 0$ such that for any $\beta \in R^p$, $\|\beta - \beta^0\| < \delta_0$

$$\frac{\partial}{\partial \beta_j} g(x,\beta) \ (j = 1, 2, \dots, p) \quad \text{and} \quad \frac{\partial^2}{\partial \beta_j \partial \beta_k} g(x,\beta) \ (j,k = 1, 2, \dots, p)$$

exist for any $x \in S_1$ and are uniformly in $x \in S_1$ continuous. Let us denote the corresponding vector and the matrix simply by $g'(x,\beta)$ and $g''(x,\beta)$, respectively, and their coordinates and elements by $g'_j(x,\beta)$ and $g''_{jk}(x,\beta)$.

(ii) There is $J \in (1,\infty)$ such that

$$\max_{1 \le j \le p} \sup_{x \in S_1, \beta \in R^p, \|\beta - \beta^0\| < \delta_0/2} |g'_j(x,\beta)| < J$$

and

$$\max_{1 \le j,k \le p} \sup_{x \in S_1, \beta \in R^p, \|\beta - \beta^0\| < \delta_0/2} |g''_{jk}(x,\beta)| < J.$$

(iii) The matrix $Q = \mathsf{E}_K\{g'(x,\beta^0)[g'(x,\beta^0)]^T\}$ is regular (and hence in our case positive definite).

**Remark 4.** *Observe that under* **Conditions C** *the functions $g$ and $g'$ are absolutely, and uniformly with respect to $x \in S_1$, continuous in $\delta_0$-neighbourhood of $\beta^0$ (let us recall that $S_1$ is the support of $K(x)$).*

**Remark 5.** *From the fact that the sequences $\{e_i\}_{i=1}^{\infty}$ and $\{X_i\}_{i=1}^{\infty}$ are independent and from* **B.iv** *together with* **C.ii** *it follows that $\frac{1}{n}\sum_{i=1}^{n} \psi'(e_i) g'(X_i,\beta^0)[g'(X_i,\beta^0)]^T$ converges in probability to $Q \cdot \mathsf{E}_F \psi'(e_1)$. Similarly, $\frac{1}{n}\sum_{i=1}^{n} \psi(e_i) g''(X_i,\beta^0)$ converges to the zero matrix in probability.*

Due to the assumption of the existence and the continuity of $\psi$ and $g'$ we may look for $\hat\beta^{(n)}$ as

$$\hat\beta^{(n)} = \arg_{\beta \in R^p}\{\sum_{i=1}^{n} \psi(Y_i - g(X_i,\beta)) \ g'(X_i,\beta) = 0\}, \tag{5}$$

as well as for $\hat{\beta}^{(n-1,\ell)}$ as

$$\hat{\beta}^{(n-1,\ell)} = \arg_{\beta \in R^p}\{\sum_{i=1, i\neq \ell}^{n} \psi(Y_i - g(X_i,\beta))\ g'(X_i,\beta) = 0\}. \tag{6}$$

Of course, due to the fact that we have not asked for the monotonicity of the function $\psi(t)$ we have only

$$\operatorname{argmin}_{\beta \in R^p}\{\sum_{i=1}^{n} \varrho(Y_i - g(X_i,\beta))\} \subset \arg_{\beta \in R^p}\{\sum_{i=1}^{n} \psi(Y_i - g(X_i,\beta))\ g'(X_i,\beta) = 0\}.$$

Recalling that $e_i = Y_i - g(X_i,\beta^0)$, let us put for any $\beta \in R^p \quad r_i(\beta) = Y_i - g(X_i,\beta)$ and for any pair $\beta_1, \beta_2 \in R^p$

$$\xi_i(\beta_1,\beta_2) = \min\{r_i(\beta_1), r_i(\beta_2)\} \quad \text{and} \quad \zeta_i(\beta_1,\beta_2) = \max\{r_i(\beta_1), r_i(\beta_2)\}.$$

Finally, for any $\omega \in \Omega$ define

$$\mathcal{H}_{n,1,\ell}(\omega) =$$
$$= \left\{ i \in \{1,2,\ldots,\ell-1,\ell+1,\ldots,n\},\ \left[\xi_i(\hat{\beta}^n, \hat{\beta}^{(n-1,\ell)}), \zeta_i(\hat{\beta}^n, \hat{\beta}^{(n-1,\ell)})\right] \cap \mathcal{C} \neq \emptyset \right\}$$

and

$$\mathcal{H}_{n,2,\ell}(\omega) = \{1,2,\ldots,\ell-1,\ell+1,\ldots,n\} \setminus \mathcal{H}_{n,1,\ell}(\omega).$$

Now using (5) and (6), and employing the mean value theorem we may write

$$\begin{aligned}
&\sum_{i \in \mathcal{H}_{n,1,\ell}} \Big[\psi(Y_i - g(X_i,\hat{\beta}^{(n)}))\ g'(X_i,\hat{\beta}^{(n)}) \\
&\qquad\qquad -\psi(Y_i - g(X_i,\hat{\beta}^{(n-1,\ell)}))\ g'(X_i,\hat{\beta}^{(n-1,\ell)})\ \Big] \\
&\quad + \sum_{i \in \mathcal{H}_{n,2,\ell}} \Big[\psi'(Y_i - g(X_i,\tilde{\beta}))\ g'(X_i,\tilde{\beta})\ \left[g'(X_i,\tilde{\beta})\right]^T \\
&\qquad\qquad -\psi(Y_i - g(X_i,\tilde{\beta}))\ g''(X_i,\tilde{\beta})\Big] \cdot (\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}) \\
&\qquad\qquad = -\psi(Y_\ell - g(X_\ell,\hat{\beta}^{(n)}))\ g'(X_\ell,\hat{\beta}^{(n)}).
\end{aligned} \tag{7}$$

where $\max\{\|\tilde{\beta} - \hat{\beta}^{(n)}\|, \|\tilde{\beta} - \hat{\beta}^{(n-1,\ell)}\|\} \leq \|\hat{\beta}^{(n)} - \hat{\beta}^{(n-1,\ell)}\|$.

**Remark 6.** *It follows from* **B.iii** *and* **C.ii** *that the right hand side of* (7) *is bounded.*

**Lemma 1.** *Let* **Conditions A, B** *and* **C** *hold. Moreover, let us assume that the set* $\mathcal{C} = \emptyset$ *(see* **B.i***). Then*

$$n(\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}) = -Q^{-1}\ \mathsf{E}_F^{-1}\psi'(e_1)\ \psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)}))\ g'(X_\ell, \hat{\beta}^{(n)}) + o_p(1)$$

*uniformly in* $\ell = 1, 2, \ldots n$.

**Remark 7.** *The uniformity claimed in* **Lemma 1** *is of the following type:* $\forall(\delta > 0$ *and* $\varepsilon > 0)$ $\exists(n_0 \in N)$ $\forall(n \in N,\ n \geq n_0$ *and* $\ell = 1, 2, \ldots, n)$

$$P\left(\left\| n\left(\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\right) + Q^{-1}\ \mathsf{E}_F^{-1}\psi(e_1)\ \psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)}))\ g'(X_\ell, \hat{\beta}^{(n)})\right\| > \delta\right) < \varepsilon$$

(*i.e.* $n_0$ *is the same for all* $\ell = 1, 2, \ldots, n$) *but not necessarily*

$$P\left(\max_{\ell=1,2,\ldots,n}\left\| n\left(\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\right) + Q^{-1}\mathsf{E}_F^{-1}\psi(e_1)\ \psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)}))\ g'(X_\ell, \hat{\beta}^{(n)})\right\| > \delta\right) < \varepsilon.$$

Proof of Lemma 1: First of all, we shall prove that for any $u, v = 1, 2, \ldots, p$ we have uniformly in $\ell = 1, 2, \ldots, n$

$$\lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=1}^{n}\left[\psi'(Y_i - g(X_i, \tilde{\beta}))\ g'_u(X_i, \tilde{\beta})g'_v(X_i, \tilde{\beta}) - \psi(Y_i - g(X_i, \tilde{\beta}))\ g''_{uv}(X_i, \tilde{\beta})\right] - q_{uv}\mathsf{E}_F\psi'(e_1)\right| = 0$$

in probability (let us recall that $\tilde{\beta}$ was introduced in (7)). Now let us fix some $\tau > 0$ and find $\nu > 0$ so that for any pair $t_1, t_2 \in R$ such that $|t_1 - t_2| < \nu$ we have $|\psi(t_1) - \psi(t_2)| < \tau J^{-1}$ and also $|\psi'(t_1) - \psi'(t_2)| < \tau J^{-1}$. Moreover, let us find $\kappa \in (0, \delta_0)$ such that for any $\beta^1 \in R^p, \|\beta^1 - \beta^0\| < \kappa$ we have

$$\sup_{x\in S_1}|g(x, \beta^1) - g(x, \beta^0)| < \nu,$$
$$\sup_{x\in S_1}\max_{1\leq u\leq p}|g'_u(x, \beta^1) - g'_u(x, \beta^0)| < L^{-1}\ \tau$$

and

$$\sup_{x\in S_1}\max_{1\leq u,v\leq p}|g''_{uv}(x, \beta^1) - g''_{uv}(x, \beta^0)| < L^{-1}\tau.$$

Now, let us fix some $\varepsilon > 0$ and $\delta > 0$ and making use of the law of large numbers let us find $n_0 \in N$ so that for any $n \in N, n \geq n_0$ we have for the set

$$A_n = \left\{ \omega \in \Omega : \left| \frac{1}{n} \sum_{i=1}^{n} \left[ \psi'(e_i)\ g_u'(X_i, \beta^0)\ g_v'(X_i, \beta^0) \right.\right.\right.$$
$$\left.\left.\left. + \psi(e_i)\ g_{uv}''(X_i, \beta^0) \right] - q_{uv}\ \mathsf{E}_F \psi'(e_1) \right| > \delta \right\}$$

$P(A_n) < \varepsilon$. Moreover, let us find $n_1 \in N, n_1 > n_0$ such that for any $n \in N, n > n_1$ and $\ell = 1, 2, \ldots, n$

$$P\left( \|\hat{\beta}^{(n)} - \beta^0\| > \frac{1}{2}\kappa \right) < \varepsilon$$

and

$$P\left( \|\hat{\beta}^{(n-1,\ell)} - \beta^0\| > \frac{1}{2}\kappa \right) < \varepsilon \qquad \ell = 1, 2, \ldots, n$$

and let us denote by $B_n$ and $B_{n,\ell}$ the sets $\{\omega \in \Omega : \|\hat{\beta}^{(n)} - \beta^0\| > \frac{1}{2}\kappa\}$ and $\{\omega \in \Omega : \|\hat{\beta}^{(n-1,\ell)} - \beta^0\| > \frac{1}{2}\kappa\}$, respectively. Then we have for any $\ell = 1, 2, \ldots, n$

$$P(A_n \cup B_n \cup B_{n,\ell}) < 3\varepsilon$$

and since

$$\|\tilde{\beta} - \beta^0\| \leq \|\hat{\beta}^{(n)} - \beta^0\| + \|\hat{\beta}^{(n-1,\ell)} - \beta^0\|$$

for any $\omega \in [A_n \cup B_n \cup B_{n,\ell}]^c$ we have

$$\max_{1 \leq u,v \leq p} \left| \frac{1}{n} \sum_{i=1}^{n} \left[ \psi'(Y_i - g(X_i, \tilde{\beta})\ g_u'(X_i, \tilde{\beta}))\ g_v'(X_i, \tilde{\beta}) \right.\right.$$
$$\left.\left. + \psi(Y_i - g(X_i, \tilde{\beta}))\ g_{uv}''(X_i, \tilde{\beta}) \right] - q_{uv}\ \mathsf{E}_F \psi'(e_1) \right| < 2\tau^2 + \delta.$$

So we have just proved that the matrices

$$\mathcal{V}^{(n)} = \left\{ \frac{1}{n} \sum_{i=1}^{n} \left\{ \psi'(Y_i - g(X_i, \tilde{\beta}))\ g_u'(X_i, \tilde{\beta})\ g_v'(X_i, \tilde{\beta}) \right.\right.$$
$$\left.\left. + \psi(Y_i - g(X_i, \tilde{\beta}))\ g_{uv}''(X_i, \tilde{\beta}) \right\} \right\}_{u=1,2,\ldots,p}^{v=1,2,\ldots,p} \tag{8}$$

converge in probability to the regular matrix $Q \cdot \mathsf{E}_F \psi(e_1)$. We shall show that it enables us to use **Lemma 2** (see Appendix) to prove that

$$n\left( \hat{\beta}^{(n)} - \hat{\beta}^{(n-1,\ell)} \right) = O_p(1). \tag{9}$$

Let us assume that (9) does not hold. (Let $\ell_0$ be fixed in the rest of the proof.) Then

$$\exists(\varepsilon > 0) \quad \forall(K > 0) \qquad \limsup_{n\to\infty} P(\ n\|\hat\beta^{(n)} - \hat\beta^{(n-1,\ell_0)}\| > K) > \varepsilon.$$

But it means that for $\gamma^{(n)} = n(\hat\beta^{(n)} - \hat\beta^{(n-1,\ell)})$ the conditions of **Lemma 2** are fulfilled. So we have

$$\exists(t \in \{1,2,\ldots,p\} \text{ and } \delta > 0) \qquad \forall(H > 0)$$
$$\limsup_{n\to\infty} P\left(\left|\sum_{j=1}^{p} v_{tj}^{(n)} \cdot n(\hat\beta_j^{(n)} - \hat\beta_j^{(n-1,\ell_0)})\right| > H\right) > \delta.$$

Taking into account that $\mathcal{C} = \emptyset$, and hence $\mathcal{H}_{n,1,\ell} = \emptyset$, we see that the previous inequality yields a contradiction with (7), see **Remark 6**, i.e. (9) holds. The rest of the proof is straightforward. Let us rewrite (7) into the form (keep in mind that $\mathcal{H}_{n,1,\ell} = \emptyset$, and also (8))

$$\left\{\mathcal{V}^{(n)} - Q \cdot \mathsf{E}_F\psi(e_1)\right\} n(\hat\beta^{(n-1,\ell)} - \hat\beta^{(n)})$$
$$+Q \cdot \mathsf{E}_F\psi(e_1) n(\hat\beta^{(n-1,\ell)} - \hat\beta^{(n)}) = -\psi(Y_\ell - g(X_\ell,\hat\beta^{(n)}))g'(X_\ell,\hat\beta^{(n)}) + o_p(1)$$

and the proof follows. □

However, the conditions of **Lemma 1** do not cover the $\psi$-functions frequently used in the $M$-estimation (e. g. they are not fulfilled for Huber's function). Although it is true that by small modifications of these functions we may fulfil the conditions of **Lemma 1** (e.g. imagine Huber's function modified so that it has uniformly continuous derivative), there are at least two reasons why we may try to prove the assertion of **Lemma 1** under more general conditions. At first, these small modifications break the admissibility of the estimators (see Hampel et al. (1986)). (Of course, it is more or less an academic question.) Secondly, the modifications lead to a more complicated evaluation of the $M$-estimators (which is already not very simple). Although an increase of the complexity of evaluation caused by the modifications would not be drastic, if we were able to do without them, it would be preferable. (We have left aside that it is also a theoretical challenge which is interesting to answer.)

So the next step will be to take into account such continuous functions $\psi$, that there are some points at which the derivative of $\psi$ does not exist, i.e. the set $\mathcal{C}$ is not empty (remember again Huber's function).

**Theorem 1.** *Let* **Conditions A, B** *and* **C** *hold. Then*

$$(10) \quad n(\hat\beta^{(n-1,\ell)} - \hat\beta^{(n)})$$
$$= -Q^{-1}\ \mathsf{E}_F^{-1}\psi'(e_1)\ \psi(Y_\ell - g(X_\ell,\hat\beta^{(n)}))\ g'(X_\ell,\hat\beta^{(n)}) + o_p(1)$$

*uniformly in* $\ell = 1, 2, \ldots n$.

PROOF: Let us fix some $\varepsilon > 0$ and let us consider the first term of (7). Due to the uniform (in $x \in S_1$) continuity of the function $g(x, \beta)$ at $\beta^0$ (see **C.i**) we may find $\nu_1 > 0$ such that for any $\beta \in R^p, \|\beta - \beta^0\| < \nu_1$ we have $|g(x, \beta) - g(x, \beta^0)| < \tau_0$. Now, let us find for $B_n = \{\omega \in \Omega, \|\hat{\beta}^{(n)} - \beta^0\| > \frac{1}{2}\nu_1\}$ and $B_{n,\ell} = \{\omega \in \Omega, \|\hat{\beta}^{(n-1,\ell)} - \beta^0\| > \frac{1}{2}\nu_1\}$ such $n_0 \in N$ that for any $n \in N, n \geq n_0$ we have $P(B_n) < \varepsilon$ as well as $P(B_{n-1,\ell}) < \varepsilon$, and consider instead of the first term in (7) the expression

$$\sum_{i \in \mathcal{H}_{n,1,\ell}} \Big[\psi(Y_i - g(X_i, \hat{\beta}^{(n)}))\; g'(X_i, \hat{\beta}^{(n)}) - \psi(Y_i - g(X_i, \hat{\beta}^{(n-1,\ell)}))\; g'(X_i, \hat{\beta}^{(n-1,\ell)})\Big]\, I_{\{B_n \cup B_{n,\ell}\}^c}. \tag{11}$$

where "$c$" denotes the complement. Let us put for any $n \in N, w \in R$ and $k = 1, 2, \ldots, p$

$$\beta^{(n,k)}(w) = (\hat{\beta}_1^{(n)}, \hat{\beta}_2^{(n)}, \ldots, \hat{\beta}_{k-1}^{(n)}, w, \hat{\beta}_{k+1}^{(n-1,\ell)}, \ldots, \hat{\beta}_p^{(n-1,\ell)})^T.$$

Taking into account that for any $i \in \mathcal{H}_{n,1,\ell}$ and any $\omega \in \{B_n \cup B_{n,\ell}\}^c$ we have $r_i(\hat{\beta}^{(n)}) \in \mathcal{C}(\tau_0)$ as well as $r_i(\hat{\beta}^{(n-1,\ell)}) \in \mathcal{C}(\tau_0)$ and making use of the absolute continuity of $\psi$ on $\mathcal{C}(\tau_0)$ and the absolute continuity of $\frac{\partial}{\partial \beta_j} g(x, \beta)$ we may find functions $h_{jk}(X_i, w) : R^p \to R,\ \ j, k = 1, 2, \ldots, p$ such that

$$\Big|\psi(Y_i - g(X_i, \hat{\beta}^{(n)}))\; g'(X_i, \hat{\beta}^{(n)}) - \psi(Y_i - g(X_i, \hat{\beta}^{(n-1,\ell)}))\; g'(X_i, \hat{\beta}^{(n-1,\ell)})\Big| = \left|\sum_{k=1}^{p} \int_{\hat{\beta}_k^{(n-1,\ell)}}^{\hat{\beta}_k^{(n)}} h_{jk}(X_i, \beta^{(n,k)}(w))\, dw\right|$$

and $\max_{1 \leq j,k \leq p} \sup_{\lambda \in [0,1]} |h_{jk}(X_i, \hat{\beta}^{(n-1,\ell)} + \lambda(\hat{\beta}^{(n)} - \hat{\beta}^{(n-1,\ell)}))| \leq L \cdot J^2$ (in fact, the functions $h_{jk}$ are equal a.e. to a sum of products of $\psi'$ and the elements of $g'[g']^T$, and of $\psi$ and the elements of $g''$). It implies that we may find a random $(p \times p)$-matrix, say $\mathcal{K}_i$, such that

$$|\,[\mathcal{K}_i]_{jk}\,| < p^{\frac{1}{2}} \cdot L \cdot J^2 \tag{12}$$

for $j, k = 1, 2, \ldots, p$ and such that

$$|\psi(Y_i - g(X_i, \hat{\beta}^{(n)}))\; g'(X_i, \hat{\beta}^{(n)}) - \psi(Y_i - g(X_i, \hat{\beta}^{(n-1,\ell)}))\; g'(X_i, \hat{\beta}^{(n-1,\ell)})| = \mathcal{K}_i(\hat{\beta}^{(n)} - \hat{\beta}^{(n-1,\ell)}). \tag{13}$$

Now, let us find for (an arbitrary but fixed) $\nu > 0$ such $\kappa$ that for any $\beta \in R^p$ such that $\|\beta - \beta^0\| < \kappa$ we have $|g(x,\beta) - g(x,\beta^0)| < \frac{1}{2}\nu M^{-1}$ (keep in mind the uniform (in $x \in S_1$) continuity of $g(x,\beta)$ at $\beta^0$; for $M$ see **Remark 1**). Now, let us select $n_2 \in N$, $n_2 \geq n_1$ so that for any $n \in N$, $n \geq n_2$ we have for the set $C_n = \{\omega \in \Omega : \|\hat\beta^{(n)} - \beta^0\| > \kappa\}$ and $C_{n,\ell} = \{\omega \in \Omega : \|\hat\beta^{(n-1,\ell)} - \beta^0\| > \kappa\}$ $P(C_n) < \varepsilon$ as well as $P(C_{n,\ell}) < \varepsilon$. Now, we shall consider instead of (11) the expression

$$\sum_{i \in \mathcal{H}_{n,1,\ell}} \Big[\psi(Y_i - g(X_i, \hat\beta^{(n)}))\; g'(X_i, \hat\beta^{(n)})$$

$$-\psi(Y_i - g(X_i, \hat\beta^{(n-1,\ell)}))\; g'(X_i, \hat\beta^{(n-1,\ell)})\Big]\, I_{\{B_n \cup B_{n,\ell} \cup C_n \cup C_{n,\ell}\}^c}.$$

It may be written as $(\hat\beta^{(n)} - \hat\beta^{(n-1,\ell)}) \sum_{i=1}^n \mathcal{K}_i \cdot I_{S_{ni}}$, where $S_{ni} = [B_n \cup B_{n,\ell} \cup C_n \cup C_{n,\ell}]^c \cap \{i \in \mathcal{H}_{n,1,\ell}\}$ (see (13)). Now $I_{S_{ni}} = 1$ implies that there is $j_0 \in \{1, 2, \ldots, r\}$ (see **B.i**) such that $c_{j_0} \in [\xi(\hat\beta^{(n)}, \hat\beta^{(n-1,\ell)}), \zeta(\hat\beta^{(n)}, \hat\beta^{(n-1,\ell)})]$ (see also the definition of $\mathcal{H}_{n,1,\ell}$). Let us consider the case that

$$r_i(\hat\beta^{(n)}) \leq c_{j_0} \leq r_i(\hat\beta^{(n-1,\ell)}).$$

Since $I_{S_{ni}} = 1$ (i.e. we consider a point $\omega \in \{C_n^c \cap C_{n,\ell}^c\}$ )we have $\|\hat\beta^{(n)} - \beta^0\| < \kappa$ as well as $\|\hat\beta^{(n-1,\ell)} - \beta^0\| < \kappa$ and so we have $|e_i - r_i(\hat\beta^{(n)})| = |g(X_i, \hat\beta^{(n)}) - g(X_i, \beta^0)| < \frac{1}{2}\nu M^{-1}$ as well as $|e_i - r_i(\hat\beta^{(n-1,\ell)})| < \frac{1}{2}\nu M^{-1}$. But it means that $I_{S_{ni}} = 1$ implies that $e_i \in [c_{j_0} - \frac{1}{2}\nu M^{-1}, c_{j_0} + \frac{1}{2}\nu M^{-1}]$ and hence taking into account that $S_{ni} \cap [C_n \cup C_{n,\ell}] = \emptyset$ we arrive at

$$P(S_{ni}) \leq 2 \cdot M \cdot \frac{1}{2} \cdot \nu \cdot M^{-1} = \nu.$$

Therefore

$$\mathsf{E}_F I_{S_{ni}} = P(I_{S_{ni}} = 1) \leq \nu$$

and finally for some $\delta > 0$ we obtain (see (12))

$$P(\max_{1 \leq j,k \leq p} \frac{1}{n} |\sum_{i=1}^n [\mathcal{K}_i]_{jk}\, I_{S_{ni}}| > \delta) \leq \frac{\nu \cdot p^{\frac{1}{2}} \cdot L \cdot J^2}{\delta}.$$

Since $\nu$ was arbitrary we conclude that the first term in (7) may be written as (see again (13))

$$\mathcal{K}^{(n)} \cdot n(\hat\beta^{(n)} - \hat\beta^{(n-1,\ell)}) \tag{14}$$

where $\mathcal{K}^{(n)}$ is a $(p \times p)$-matrix with elements of order $o_p(1)$. Now we may along similar lines as in the proof of **Lemma 1** show that (let us recall once again that $\tilde{\beta}$ is given in (7))

$$\frac{1}{n} \sum_{i \notin \mathcal{H}_{n,2,\ell}} \Big[ \psi'(Y_i - g(X_i, \tilde{\beta}))\ g'(X_i, \tilde{\beta})\ \Big[ g'(X_i, \tilde{\beta}) \Big]^T$$
$$(15) \qquad - \psi'(Y_i - g(X_i, \beta^0))\ g'(X_i, \beta^0)\ \Big[ g'(X_i, \beta^0) \Big]^T$$
$$+ \psi(Y_i - g(X_i, \tilde{\beta}))\ g''(X_i, \tilde{\beta}) - \psi(Y_i - g(X_i, \beta^0))\ g''(X_i, \beta^0) \Big] = o_p(1)$$

and finally, carrying out similar steps as in the first part of this proof we show that

$$\frac{1}{n} \sum_{i \in \mathcal{H}_{n,1,\ell}} \Big[ \psi'(Y_i - g(X_i, \beta^0))\ g'(X_i, \beta^0)\ \Big[ g'(X_i, \beta^0) \Big]^T$$
$$(16) \qquad + \psi(Y_i - g(X_i, \beta^0))\ g''(X_i, \beta^0) \Big] = o_p(1).$$

Now taking into account (14),(15) and (16) we may write instead of (7)

$$\left\{ \frac{1}{n} \sum_{i=1}^{n} \Big[ \psi'(e_i)\ g'(X_i, \beta^0)\ \Big[ g'(X_i, \beta^0) \Big]^T + \psi(e_i)\ g''(X_i, \beta^0) \Big] + o_p(1) \right\} \cdot$$
$$\cdot\, n(\hat{\beta}^{(n)} - \hat{\beta}^{(n-1,\ell)}) = -\psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)}) g'(X_\ell, \hat{\beta}^{(n)}).$$

The rest of the proof is the same as the last part of the proof of **Lemma 1** (starting with (8)). □

The conditions of **Theorem 1** cover a majority of the frequently used $\psi$-functions. For instance the most $B$- and $V$-optimal robust estimators, including the bulk of the estimators with the redescending $\psi$-function (e.g. tanh-type estimators) fulfil these conditions – see Hampel et al. (1986), 2.5a. They do not cover the $M$-estimators with the discontinuous $\psi$-functions. On the other hand, it is known that the estimators generated by the $\psi$-function with (at least one) downward jump (i.e. such a $\psi$-function for which at least at one point $d \in R$, $\lim_{z \nearrow d} \psi(z) > \lim_{z \searrow d} \psi(z)$ - under the assumption that such limits exist at all, like the skipped median or the estimator with skipped Huber's function) have the infinite change-of-variance sensitivity. In practical applications we usually avoid such estimators just due to the fact that the infinite change-of-variance sensitivity is an indication of an implausible fluctuation of the estimator (even for small changes of the contamination level).

It is clear that the relation (10) does not allow us to derive for the $M$-estimators a formula analogous to (4). The reason is the presence of $o_p(1)$ in it, causing that we cannot derive from it an approximation to the variance of $n\|\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\|$. But let us look a little closer on the problem. *What does the presence of* $o_p(1)$ *in (10) indicate and what may it cause ?* In fact $o_p(1)$ in (10) may imply that $n(\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)})$ can behave rather "wildly" on a set of (very) small probability. So the behaviour of $\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}$ on the set of small probability may influence (in fact it always increases) the value of $\mathrm{var}(n\|\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\|)$. If this influence is considerable, then the value $\mathrm{var}(n\|\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\|)$ gives a misleading idea about the variability of $n\|\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\|$ because the variability of the typical values of $n\|\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\|$ is in fact smaller, given by $\mathsf{E}^{-2}\psi'(e_1)\ \mathrm{var}(\|Q^{-1}g'(X_\ell, \hat{\beta}^{(n)})\psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)}))\|)$. That is why we would prefer to normalize $n\|\hat{\beta}^{(n-1,\ell)} - \hat{\beta}^{(n)}\|$ by $\mathsf{E}^{-1}\psi'(e_1)\ \mathrm{var}^{\frac{1}{2}}(\|Q^{-1}g'(X_\ell, \hat{\beta}^{(n)})\psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)}))\|)$, compare also Huber (1965). Then the characterization of the changes of the non-linear regression model estimates will be the same as for the $M$-estimators of the linear model (Víšek (1992)), namely

$$\max_{1 \le \ell \le n} \left| \psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)})) \right| . \tag{17}$$

It means that having evaluated the residuals for the given estimate of the non-linear model, we may look for the most influential points just looking for the point with the largest "$\psi$-residual". In the case when the problem of estimating the regression model is not invariant with respect to the position of data in the factor space, i.e. when this position plays an (important) role, we have to use for the sensitivity analysis directly the formula (10) instead of (17) and the computation will be a little more complicated.

From these considerations we may conclude:

**Corollary 1.** *The largest value of the studentized norm of the change of the estimate of regression coefficients is always bounded by* $\sup_{t \in R} |\psi(t)|$.

It is clear that if the $\psi$-function is properly selected (let us say "tuned") then there will be at least one point such that $\psi(Y_\ell - g(X_\ell, \hat{\beta}^{(n)})) \cong \sup_{t \in R} |\psi(t)|$, even for the redescending functions. It is also evident that the change for the $LS$-estimator would be even larger. So the assertion of **Corollary 1** may be interpreted so that using the $M$-estimators we are imposing an upper limit on a possible change of the estimate. On the other hand, it may seem strange that the influence of one point is so "large", where the converted commas indicate that one should keep in mind that we normalize the difference of estimates by the factor $n$, i.e. the change is in fact of order $O_p(n^{-1})$. But even keeping it in mind and considering fixed sample size, it is natural to ask: *Cannot we construct an estimator which would be more stable on subsamples ?*

## Appendix

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

$$\lim_{n\to\infty} v_{ij}^{(n)} = q_{ij} \qquad i,j = 1,2,\ldots,p \qquad \textit{in probability}$$

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

$$\exists\ (\varepsilon > 0) \qquad \forall\ (K > 0) \qquad \limsup_{n\to\infty} P\left(\|\gamma^{(n)}\| > K\right) > \varepsilon. \tag{18}$$

*Then*

$$\begin{gathered} \exists \qquad (k \in \{1,2,\ldots,p\} \quad \textit{and} \quad \delta > 0) \qquad \forall\ (\tau > 0) \\ \limsup_{n\to\infty}\ P\left(\left|\sum_{j=1}^{p} v_{kj}^{(n)}\gamma_j^{(n)}\right| > \tau\right) > \delta\ . \end{gathered} \tag{19}$$

Proof: Let us at first assume that for the sequence $\{\gamma^{(n)}\}_{n=1}^{\infty}$ we have

$$\exists\ (\varepsilon > 0) \qquad \forall\ (K > 0) \qquad \lim_{n\to\infty} P\left(\|\gamma^{(n)}\| > K\right) > \varepsilon. \tag{20}$$

Let us fix a sequence $\{\tilde{K}_r\}_{r=1}^{\infty} \uparrow \infty$, $\tilde{K}_1 = 0$, and construct a sequence $\{K_n\}_{n=1}^{\infty}$ in the following way. For every $r \in N$ find $n_r \in N$ such that for any $n \in N$, $n \geq n_r$

$$P\left(\|\gamma^{(n)}\| > \tilde{K}_r\right) > \frac{\varepsilon}{2}$$

and put for $\ell \in N$, $\ell \in [n_r, n_{r+1})$, $K_\ell = \tilde{K}_r$ (if $n_1 > 1$ put $K_\ell = 0$ for $\ell \leq n_1$). Denote by $B_n = \{\omega \in \Omega :\ \|\gamma^{(n)}\| > K_n\}$, i.e. $P(B_n) > \frac{\varepsilon}{2}$ for all $n \in N$. Let us assume that (19) does not hold, i.e. $\forall\ (k = 1,\ldots,p$ and $\delta > 0)$ $\ \exists\ (\tau_\delta > 0)$ and

$$\limsup_{n\to\infty} P\left(\left|\sum_{j=1}^{p} v_{kj}^{(n)}\gamma_j^{(n)}\right| > \tau_\delta\right) < \delta.$$

Finally it may be written as

$$\begin{gathered} \forall\ (k = 1,\ldots,p \text{ and } \delta > 0) \qquad \exists\ (\tau_\delta > 0 \ \text{ and } \ n_\delta \in N) \qquad \forall\ (n \in N,\ n > n_\delta) \\ P\left(\left|\sum_{j=1}^{p} v_{kj}^{(n)}\gamma_j^{(n)}\right| > \tau_\delta\right) < 2\delta. \end{gathered}$$

Put $\delta = \frac{\varepsilon}{16p}$ and denote by

$$A_n = \left\{ \omega \in \Omega : \max_{k=1,\dots,p} \left| \sum_{j=1}^{p} v_{kj}^{(n)} \gamma_j^{(n)} \right| \leq \tau_\delta \right\}.$$

Then we have for any $n > n_\delta$ $P(A_n^c) \leq \sum_{k=1}^{p} P(\left|\sum_{j=1}^{p} v_{kj}^{(n)} \gamma_j^{(n)}\right| > \tau_\delta) < \frac{2\varepsilon}{16p} \cdot p = \frac{\varepsilon}{8}$. Finally, denote by $\tilde{q}_{ij}$ the elements of $Q^{-1}$ and put $\Gamma = \max_{i,j=1,\dots,p} |\tilde{q}_{ij}|$. Select $\Delta \in (0, \frac{1}{2}p^{-2} \cdot \Gamma^{-1})$ and find $n_\Delta \in N$ such that for any $n \in N,\ n \geq n_\Delta$

$$P\left( \max_{i,j} \left| v_{ij}^{(n)} - q_{ij} \right| \geq \Delta \right) < \frac{\varepsilon}{8p^2}.$$

Denote $C_n = \{\omega \in \Omega :\ \max_{i,j=1,\dots,p} |v_{ij}^{(n)} - q_{ij}| < \Delta\}$. Then we have for any $n > n_\Delta$

$$P\left(C_n^c\right) \leq \sum_{i=1}^{p} \sum_{j=1}^{p} P\left( |v_{ij}^{(n)} - q_{ij}| \geq \Delta \right) < \frac{\varepsilon}{8p^2} p^2 = \frac{\varepsilon}{8}.$$

Since $A_n \cap B_n \cap C_n = (B_n - A_n^c) - C_n^c$ we have for any $n \in N,\ n > n_0 = \max\{n_\delta, n_\Delta\}$,

$$\begin{aligned} P\left(A_n \cap B_n \cap C_n\right) &\geq P\left(B_n - A_n^c\right) - P\left(C_n^c\right) \\ &\geq P\left(B_n\right) - P\left(A_n^c\right) - P\left(C_n^c\right) \geq \frac{\varepsilon}{2} - \frac{\varepsilon}{8} - \frac{\varepsilon}{8} = \frac{\varepsilon}{4}. \end{aligned}$$

Let $\omega \in A_n \cap B_n \cap C_n$. Putting for all $k = 1, \dots, p$

$$\sum_{j=1}^{p} v_{kj}^{(n)} \gamma_j^{(n)} = H_k$$

we have $|H_k| < \tau_\delta$ and we may write

$$\sum_{j=1}^{p} q_{kj} \gamma_j^{(n)} = H_k - \sum_{j=1}^{p} \left( v_{kj}^{(n)} - q_{kj} \right) \gamma_j^{(n)}$$

and also ($\ell = 1, \dots, p$)

$$\gamma_\ell^{(n)} = \sum_{k=1}^{p} \tilde{q}_{\ell k} H_k - \sum_{k=1}^{p} \tilde{q}_{\ell k} \sum_{j=1}^{p} \left( v_{kj}^{(n)} - q_{kj} \right) \gamma_j^{(n)}$$

and finally ($\ell = 1, \dots, p$)

$$\left| \gamma_\ell^{(n)} \right| \leq \sum_{k=1}^{p} |\tilde{q}_{\ell k}| \cdot |H_k| + \Delta \cdot p^2 \cdot \Gamma \cdot \max_{j=1,\dots,p} \left| \gamma_j^{(n)} \right|. \tag{21}$$

Let $\ell_n \in \{1, 2, \ldots, p\}$ be such that $|\gamma_{\ell_n}^{(n)}| = \max_{j=1,\ldots,p} |\gamma_j^{(n)}|$. From (21) we have for any $n \in N$, $n > n_0$ and $\omega \in A_n \cap B_n \cap C_n$

$$\left|\gamma_{\ell_n}^{(n)}\right| \left(1 - \Delta \cdot p^2 \cdot \Gamma\right) \leq p \cdot \Gamma \cdot \tau_\delta,$$

i.e.

$$\left|\gamma_{\ell_n}^{(n)}\right| \leq 2 \cdot p \cdot \Gamma \cdot \tau_\delta.$$

Now it is sufficient to find $n \in N$ so that $K_n > 2 \cdot p^2 \cdot \Gamma \cdot \tau_\delta$ and we obtain

$$2 \cdot p^2 \cdot \Gamma \cdot \tau_\delta < \left\|\gamma^{(n)}\right\| \leq p \left|\gamma_{\ell_n}^{(n)}\right| \leq 2 \cdot p^2 \cdot \Gamma \cdot \tau_\delta,$$

which is a contradiction. To prove the lemma with (18) instead of (20) it is sufficient to assume again that it does not hold and to select a subsequence $\{\gamma^{(n_k)}\}_{k=1}^{\infty}$ for which (20) holds and we get again a contradiction.

## References

Chatterjee S., Hadi A.S. (1988), *Sensitivity Analysis in Linear Regression*, J. Wiley & Sons, New York.

Cook R.D., Weisberg S. (1982), *Residuals and Influence in Regression*, Chapman and Hall, New York.

Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. (1986), *Robust Statistics – The Approach Based on Influence Functions*, J. Wiley & Sons, New York.

Huber P.J. (1964), *A robust version of the probability ratio test*, Ann. Math. Statist. **36,**, 1753–1758.

Víšek J.Á. (1992), *Stability of regression model estimates with respect to subsamples*, Computational Statistics **7**, 183–203.

Welsch R.E. (1982), *Influence function and regression diagnostics*, In: Modern Data Analysis, R.L. Launer and A.F. Siegel, eds., Academic Press, New York, 149–169.

Zvára K. (1989), *Regression analysis* (in Czech), Academia, Prague.

Rubio A.M., Quintana F.
Departamento de Matemáticas, Universidad de Extremadura, 10071 Cáceres, Spain

Víšek J.Á.
Department of Stochastic Informatics, Institute of Information Theory and Automation, Academy of Sciences, Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic

(*Received* April 5, 1993)