Acta Univ. Palacki. Olomuc., Fac. rer. nat., Mathematica **47** (2008) 101–113

Linearization Regions for Confidence Ellipsoids*

Lubomír KUBÁČEK [1], Eva TESAŘÍKOVÁ [2]

[1] *Department of Mathematical Analysis and Applications of Mathematics*
Faculty of Science, Palacký University
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kubacekl@inf.upol.cz

[2] *Department of Algebra and Geometry, Faculty of Science, Palacký University*
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: tesariko@inf.upol.cz

(Received February 12, 2008)

Abstract

If an observation vector in a nonlinear regression model is normally distributed, then an algorithm for a determination of the exact $(1-\alpha)$-confidence region for the parameter of the mean value of the observation vector is well known. However its numerical realization is tedious and therefore it is of some interest to find some condition which enables us to construct this region in a simpler way.

Key words: Confidence ellipsoid; nonlinear regression model; linearization region.

2000 Mathematics Subject Classification: 62F10, 62J05

1 Introduction

In a linear statistical model with normally distributed observation vector the construction of the confidence regions is a simple problem. If the statistical model is nonlinear, i.e. the mean value of the observation vector is a nonlinear vector function of the parameters, then the problem can be also solved, however

*Supported by the Council of the Czech Government MSM 6 198 959 214.

it is much more complicated. Therefore it is reasonable to find another, much more simpler procedure, which can be used at least under some conditions.

The aim of the paper is to find such conditions which enables us to use the procedure from the theory of linear statistical models.

More on the problem of linearization of regression models cf. [3], [4], [5], [6], [10], [11].

2 Notation and auxiliary statement

Let $\mathbf{Y}$ be an n-dimensional random vector (observation vector) which is normally distributed. Its mean value is equal to $\mathbf{f}(\boldsymbol{\beta})$, where $\boldsymbol{\beta} \in R^k$ (k-dimensional real linear space) is an unknown vector parameter and $\mathbf{f}(\cdot)\colon R^k \to R^n$ is a vector function. It is assumed that it can be expressed with sufficiently high accuracy as

$$\mathbf{f}(\mathbf{u}) = \mathbf{f}_0 + \mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0) + \frac{1}{2}\boldsymbol{\kappa}(\mathbf{u} - \boldsymbol{\beta}_0), \quad \mathbf{u} \in R^k,$$

where

$$\mathbf{f}_0 = \mathbf{f}(\boldsymbol{\beta}_0), \quad \mathbf{F} = \frac{\partial \mathbf{f}(\mathbf{u})}{\partial \mathbf{u}'}\Big|_{u=\beta_0},$$

$$\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = [\kappa_1(\delta\boldsymbol{\beta}), \dots, \kappa_n(\delta\boldsymbol{\beta})]', \quad \delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$$

$$\kappa_i(\delta\boldsymbol{\beta}) = \delta\boldsymbol{\beta}' \frac{\partial^2 f_i(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}\Big|_{u=\beta_0} \delta\boldsymbol{\beta}, \quad i = 1, \dots, n.$$

The covariance matrix of the vector $\mathbf{Y}$ is $\sigma^2\mathbf{V}$, where $\sigma^2 \in (0, \infty)$ is either known or unknown parameter and the $n \times n$ matrix $\mathbf{V}$ is given.

The notation

$$\mathbf{Y} \sim N_n[\mathbf{f}(\boldsymbol{\beta}), \sigma^2\mathbf{V}], \quad \boldsymbol{\beta} \in R^k, \tag{1}$$

will be used in the following text.

The quadratized version of the model (1) is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\left[\mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}), \sigma^2\mathbf{V}\right], \quad \boldsymbol{\beta} \in R^k, \tag{2}$$

and the linearized version is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\left(\mathbf{F}\delta\boldsymbol{\beta}, \sigma^2\mathbf{V}\right), \quad \boldsymbol{\beta} \in R^k. \tag{3}$$

Assumption The regularity of the model (3) is assumed in the following text, i.e. the rank of the matrix $\mathbf{F}$ is $r(\mathbf{F}) = k < n$ and the matrix $\mathbf{V}$ is positive definite.

Lemma 2.1 *The $(1-\alpha)$-confidence region for the vector $\boldsymbol{\beta}$ in the model (3) is*

$$\mathcal{E} = \left\{\mathbf{u}\colon (\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}}) \leq \sigma^2\chi^2_k(0; 1-\alpha)\right\}, \tag{4}$$

if the parameter σ^2 is known.

If it is estimated, then

$$\mathcal{E} = \left\{\mathbf{u}\colon (\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}}) \leq \widehat{\sigma^2} F_{k,n-k}(0; 1-\alpha)\right\}. \quad (5)$$

Here

$$\begin{aligned}\widehat{\delta\boldsymbol{\beta}} &= (\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0),\\ \widehat{\sigma^2} &= (\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})'\,\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})/(n-k),\end{aligned}$$

$\chi^2_k(0; 1-\alpha)$ is $(1-\alpha)$-quantile of the central chi-squared distribution with k degrees of freedom and $F_{k,n-k}(0; 1-\alpha)$ is $(1-\alpha)$-quantile of the central Fisher–Snedecor distribution with k and $n-k$ degrees of freedom.

Proof Proof is well known (cf. e.g. [2]) and therefore it is omitted. □

Lemma 2.1 is not valid in the model (2). However if $\delta\boldsymbol{\beta} = \boldsymbol{\beta}^* - \boldsymbol{\beta}_0$ is sufficiently small, where $\boldsymbol{\beta}^*$ is the actual value of the vector parameter $\boldsymbol{\beta}$, it can be expected that the region $\mathcal{E}$ from (4) and (5), respectively, covers the actual value $\boldsymbol{\beta}^*$ with a probability larger than $1 - \alpha - \varepsilon$, where $\varepsilon > 0$ is a sufficiently small real number.

3 Linearization region

Consider the quadratized model (2) with the given covariance matrix $\boldsymbol{\Sigma} = \sigma^2\mathbf{V}$ (i.e. σ^2 is known).

Definition 3.1 The Bates and Watts [1] parametric curvature $K^{(par)}$ and the intrinsic curvature $K^{(int)}$ of the model (1) at the point $\boldsymbol{\beta}_0$ are given as

$$K^{(par)} = \sigma \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k\right\} = \sigma K_0^{(par)}$$

and

$$K^{(int)} = \sigma \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k\right\} = \sigma K_0^{(int)}.$$

Here $\mathbf{P}_F^{V^{-1}} = \mathbf{F}(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}$ and $\mathbf{M}_F^{V^{-1}} = \mathbf{I} - \mathbf{P}_F^{V^{-1}}$.

Let an r-dimensional vector function $\mathbf{d}\colon R^k \to R^r$

$$\mathbf{d}(\boldsymbol{\beta}) = \mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta}, \quad \boldsymbol{\beta} \in R^k,$$

where $r(\mathbf{D}_{r,k}) = r \leq k$, be under consideration.

Theorem 3.1 *Let α and ε be sufficiently small positive real numbers and let $\delta_{\max}$ be solution of the equation*

$$P\{\chi_r^2(\delta_{\max}) \leq \chi_r^2(0; 1-\alpha)\} = 1 - \alpha - \varepsilon.$$

If

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{\mathcal{E}} = \left\{\delta\boldsymbol{\beta}\colon \delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq \frac{2\sqrt{\delta_{\max}}}{K^{(par)}(\boldsymbol{\beta}_0)}\right\}, \quad \text{where} \quad \mathbf{C} = \frac{\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}}{\sigma^2},$$

then

$$\mathcal{E}^* = \left\{\mathbf{u}\colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}) \leq \left(\sqrt{\chi_r^2(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2\right\}$$

covers $\mathbf{D}\delta\boldsymbol{\beta}$ with probability at least $(1 - \alpha - \varepsilon)$.

Proof Let

$$\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq \frac{2\sqrt{\delta_{\max}}}{K^{(par)}(\boldsymbol{\beta}_0)}.$$

Then

$$\forall\{\mathbf{u} \in R^r\}|\mathbf{u}'\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}]| \leq \sqrt{\delta_{\max}}\sqrt{\mathbf{u}'\mathbf{D}\mathbf{C}^{-1}\mathbf{D}'\mathbf{u}},$$

what is equivalent, with respect to the Scheffé theorem [9], to

$$[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}]'\mathbf{D}'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}] \leq \delta_{\max}.$$

Let

$$\left\{\delta\boldsymbol{\beta}\colon [E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}]'\mathbf{D}(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}] \leq \delta_{\max}\right\}. \tag{6}$$

Let $(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1} = \sum_{i=1}^r \lambda_i \mathbf{f}_i\mathbf{f}_i'$ be the spectral decomposition.

The $(1-\alpha)$-confidence ellipsoid in the linearized model is

$$\left\{\mathbf{u}\colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}) \leq \chi_r^2(0; 1-\alpha)\right\}$$

and its semiaxes are $a_i = \sqrt{\chi_r^2(0; 1-\alpha)}/\sqrt{\lambda_i}$, $i = 1, \ldots, r$. The semiaxes of the ellipsoid (6) are $\pi_i = \sqrt{\delta_{\max}}/\sqrt{\lambda_i}, i = 1, \ldots, r$.

The semiaxes of the ellipsoid $\mathcal{E}^*$ are $a_i + \pi_i$, $i = 1, \ldots, r$ and it covers all $(1-\alpha)$-ellipsoids in the linearized model with centers

$$\mathbf{D}\widehat{\delta\boldsymbol{\beta}} + \mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}], E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta} \in \mathcal{E}^*.$$

The random variable

$$(\mathbf{D}\delta\boldsymbol{\beta} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{D}\delta\boldsymbol{\beta} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})$$

is chi-squared with r degrees of freedom and with the parameter of noncentrality equal to

$$\delta = \left\{\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}]\right\}'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}] < \delta_{\max}.$$

If $\delta_{\max}$ satisfies the equality

$$P\{\chi_r^2(\delta_{\max}) \leq \chi_r^2(0; 1-\alpha)\} = 1 - \alpha - \varepsilon,$$

then the ellipsoid $\mathcal{E}^*$ covers the vector $\mathbf{D}\delta\boldsymbol{\beta}$ with probability larger or equal to $1 - \alpha - \varepsilon$. □

Corollary 3.1 *Let* $\mathbf{d}(\boldsymbol{\beta}) = \boldsymbol{\beta}$. *If*

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{\mathcal{E}} = \left\{\delta\boldsymbol{\beta}\colon \delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq \frac{2\sqrt{\delta_{\max}}}{K^{(par)}(\boldsymbol{\beta}_0)}\right\},$$

then

$$\mathcal{E}^* = \left\{\mathbf{u}\colon (\mathbf{u} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{C}(\mathbf{u} - \widehat{\delta\boldsymbol{\beta}}) \leq \left(\sqrt{\chi_r^2(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2\right\}$$

covers $\delta\boldsymbol{\beta}$ *with probability at least* $(1 - \alpha - \varepsilon)$.

Let the function $\mathbf{d}(\boldsymbol{\beta})$ be of the quadratic form, i.e.

$$\mathbf{d}(\boldsymbol{\beta}) = \mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\delta}(\delta\boldsymbol{\beta}),$$

where $\boldsymbol{\delta}(\delta\boldsymbol{\beta}) = [\delta_1(\delta\boldsymbol{\beta}), \ldots, \delta_r(\delta\boldsymbol{\beta})]'$, $\delta_i(\delta\boldsymbol{\beta}) = \delta\boldsymbol{\beta}'\mathbf{A}_i\delta\boldsymbol{\beta}$, $\mathbf{A}_i = \mathbf{A}_i'$, $i = 1, \ldots, r$.

Definition 3.2 The measure of nonlinearity for the confidence ellipsoid is

$$C_{d(\cdot),\text{conf}} = \\ = \sup\left\{\frac{\sqrt{(\boldsymbol{\delta} - \mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\boldsymbol{\delta} - \mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa})}}{\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k\right\}.$$

Theorem 3.2 *Let* $\delta_{\max}$ *satisfies the equality*

$$P\{\chi_r^2(\delta_{\max}) \leq \chi_r^2(0; 1-\alpha)\} = 1 - \alpha - \varepsilon,$$

where α *and* ε *are positive sufficiently small real numbers. Let*

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{d(\cdot),\text{conf}} = \left\{\delta\boldsymbol{\beta}\colon \delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq \frac{2\sqrt{\delta_{\max}}}{C_{d(\cdot),\text{conf}}}\right\}.$$

Then the ellipsoid

$$\mathcal{E}_{d(\cdot)} = \Big\{\mathbf{u} \in R^r \colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})' \leq \left(\sqrt{\chi_r^2(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2\Big\} + \mathbf{D}\widehat{\delta\boldsymbol{\beta}}$$

covers the vector

$$\mathbf{D}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\delta}(\delta\boldsymbol{\beta})$$

with probability larger or equal at least to $(1 - \alpha - \varepsilon)$.

Proof The random variable

$$[\mathbf{d}(\boldsymbol{\beta}) - \mathbf{d}(\boldsymbol{\beta}_0) - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}]'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}[\mathbf{d}(\boldsymbol{\beta}) - \mathbf{d}(\boldsymbol{\beta}_0) - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}]$$

is chi-squared distributed with the parameter of noncentrrality

$$\delta = \frac{1}{4}(\boldsymbol{\delta} - \mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\boldsymbol{\delta} - \mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}).$$

From Definition 3.2 we have

$$4\delta \leq (C_{d(\cdot),\mathrm{conf}})^2(\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta})^2.$$

If $\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq 2\sqrt{\delta_{\max}}/C_{d(\cdot),\mathrm{conf}}$, then $\delta \leq \delta_{\max}$ and then the vector

$$E[\delta(\boldsymbol{\beta}_0) + \mathbf{D}\widehat{\delta\boldsymbol{\beta}}] - [\mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\delta}(\delta\boldsymbol{\beta})]$$

is an element of the ellipsoid

$$\Big\{\mathbf{u} \colon \mathbf{u} \in R^r, \mathbf{u}'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{u} \leq \delta_{\max}\Big\}$$

with probability at least $1 - \alpha - \varepsilon$. Now it is obvious how to finish the proof. □

Corollary 3.2 *If the function* $\mathbf{d}(\cdot)$ *is linear, i.e.* $\mathbf{d}(\boldsymbol{\beta}) = \mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta}$, *then*

$$\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq \frac{2\sqrt{\delta_{\max}}}{C_{D\beta}} \Rightarrow P\{\mathbf{d}(\boldsymbol{\beta}) \in \mathcal{E}\} \geq 1 - \alpha - \varepsilon,$$

where

$$C_{D\beta} = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}\mathbf{D}'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}}}{\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta}} \colon \delta\boldsymbol{\beta} \in R^k\right\},$$

$$\mathcal{E} = \Big\{\mathbf{u} + \mathbf{d}(\boldsymbol{\beta}_0) \colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}) \leq \left(\sqrt{\chi_r^2(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2\Big\}$$

and

$$P\{\chi_r^2(\delta_{\max}) \le \chi_r^2(0; 1-\alpha)\} = 1 - \alpha - \varepsilon$$

(cf. Theorem 3.2).

Corollary 3.3 *If the function $d(\cdot)$ is scalar, i.e. $d(\boldsymbol{\beta}) = d(\boldsymbol{\beta}_0) + \mathbf{d}'\delta\boldsymbol{\beta} + \frac{1}{2}\delta(\delta\boldsymbol{\beta})$, then*

$$\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \le \frac{2\sqrt{\delta_{\max}}}{C_{d(\beta)}} \Rightarrow P\{d(\boldsymbol{\beta}) \in \mathcal{E}\} \ge 1 - \alpha - \varepsilon,$$

where

$$C_{d(\beta)} = \sup\left\{ \frac{\sqrt{\boxed{a}'(\mathbf{d}'\mathbf{C}^{-1}\mathbf{d})^{-1}\boxed{a}}}{\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k \right\},$$

$$\mathcal{E} = \left\{ u + d(\boldsymbol{\beta}_0) \colon (u - \mathbf{d}'\widehat{\delta\boldsymbol{\beta}})^2/(\mathbf{d}'\mathbf{C}^{-1}\mathbf{d}) \le \left(\sqrt{\chi_1^2(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2 \right\},$$

and

$$\boxed{a} = [\delta(\delta\boldsymbol{\beta}) - \mathbf{d}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})].$$

Until now the parameter σ^2 is assumed to be known. Let

$$T(\delta\boldsymbol{\beta}) = U\frac{n-k}{k}, \qquad U = \frac{(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}})}{(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})}.$$

Then $T(\delta\boldsymbol{\beta}^*)$ has the Fisher–Snedecor distribution $F_{k,n-k}(\cdot)$ in case of the linearized version (3) of the regression model.

Lemma 3.1 *In the the quadratized model (2) we have*

(i) $(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}}) \sim \sigma^2\chi_k^2(\delta_1)$,

where $\delta_1 = \boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})/(4\sigma^2)$ and

(ii) $(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}}) \sim \sigma^2\chi_{n-k}^2(\delta_2)$,

where $\delta_2 = \boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})/(4\sigma^2)$.

Proof (i) The parameter of noncentrality δ_1 is

$$\delta_1 = E(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}E(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}})/\sigma^2.$$

Since in the quadratized model (2)

$$\begin{aligned} E(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}}) &= \delta\boldsymbol{\beta}^* - (\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}\left[\mathbf{F}\delta\boldsymbol{\beta}^* + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*)\right] \\ &= -\frac{1}{2}(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*), \end{aligned}$$

the statement (i) is valid.

(ii) Analogously

$$E(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}}) = \mathbf{F}\delta\boldsymbol{\beta}^* + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*) - \mathbf{F}(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1} \times \left[\mathbf{F}\delta\boldsymbol{\beta}^* + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*)\right] = \frac{1}{2}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*)$$

and therefore also the statement (ii) is valid. □

The probability density of the random variable $\chi_f^2(\delta)$ is [7]

$$g_{f,\delta}(y) = \begin{cases} \exp[-(y+\delta)/2] \sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{\delta}{2}\right)^r \frac{y^{r+(f/2)-1}}{2^{r+f/2}\Gamma(r+f/2)}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

Thus the density of the random variable U is

$$g(u;\delta_1,\delta_2) = \int_0^\infty g_{k,\delta_1}(uv) g_{n-k,\delta_2}(v) v dv.$$

Let the set $C_{\delta_1^*,\delta_2^*}$ be defined as follows.

$$C_{\delta_1^*,\delta_2^*} = \left\{(\delta_1^*,\delta_2^*)\colon \int_0^{[(n-k)/k]F_{k,n-k}(0;1-\alpha)} g(u;\delta_1^*,\delta_2^*)du = 1-\alpha-\varepsilon\right\}.$$

Theorem 3.3 *The linearization region for the confidence ellipsoid in the case of the estimated σ^2 is*

$$\mathcal{L}_{\delta_1,\delta_2} = \left\{\delta\boldsymbol{\beta}\colon \delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} < \sigma\frac{2\sqrt{\delta_1^*}}{K_0^{(par)}} \quad \& \quad \delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} < \sigma\frac{2\sqrt{\delta_2^*}}{K_0^{(int)}}\right\}$$

i.e.

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{\delta_1,\delta_2} \Rightarrow P\{\delta\boldsymbol{\beta} \in \mathcal{E}\} \ge 1-\alpha-\varepsilon,$$

where $\mathcal{E}$ is given by (5).

Proof With respect to Definition 3.3 it is valid that

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) \le \frac{1}{4\sigma^4}(\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta})^2\left(\sigma K_0^{(par)}\right)^2.$$

Thus the inequality

$$\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} \le \sigma\frac{2\sqrt{\delta_1^*}}{K_0^{(par)}}$$

implies

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \delta_1 \le \delta_1^*.$$

Analogously

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) \leq \frac{1}{4\sigma^4}(\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta})^2\left(\sigma K_0^{(int)}\right)^2.$$

Thus the inequality

$$\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} \leq \sigma\frac{2\sqrt{\delta_2^*}}{K_0^{(int)}}$$

implies

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \delta_2 \leq \delta_2^*.$$

□

In order not to prefer one of the parameter noncentrality for the other one, the condition

$$\delta_1^*/\delta_2^* = (K_0^{(par)})^2/(K_0^{(int)})^2$$

can be used. Thus

$$\frac{2\sqrt{\delta_1^*}}{K_0^{(par)}} = \frac{2\sqrt{\delta_2^*}}{K_0^{(int)}}.$$

In some cases the Bates and Watts intrinsic curvature is zero and thus the random variable $T = [(n-k)/k]U$ has the the noncentral Fisher–Snedecor distribution

$$F_{k,n-k}(\delta_1) = [\chi_k^2(\delta_1)/k]/[\chi_{n-k}^2(0)/(n-k)],$$

since $\delta_2 = 0$. Let $\delta_{1,\max}$ be solution of the equation

$$P\{F_{k,n-k}(\delta_{1,\max}) \geq F_{k,n-k}(0; 1-\alpha)\} = \alpha + \varepsilon.$$

If

$$C^{(ell,D\delta\beta)} = \sigma C_0^{(ell,D\delta\beta)},$$

where

$$C_0^{(ell,D\delta\beta)} = \\ = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'\mathbf{V}^{-1}\mathbf{F}\mathbf{C}_0^{-1}\mathbf{D}'(\mathbf{D}\mathbf{C}_0^{-1}\mathbf{D}')^{-1}\mathbf{D}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\boldsymbol{\kappa}}}{\delta\boldsymbol{\beta}'\mathbf{C}_0\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k\right\},$$

where $\mathbf{C}_0 = \mathbf{F}'\mathbf{V}^{-1}\mathbf{F}$, then the following implication is valid.

$$\begin{aligned}\delta\boldsymbol{\beta}'\mathbf{C}_0\delta\boldsymbol{\beta} \;\leq\; & \sigma\frac{2\sqrt{\delta_{1,\max}}}{C_0^{(ell,D\delta\beta)}} \\ & \Rightarrow P\Big\{(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{D}'(\mathbf{D}\mathbf{C}_0\mathbf{D}')^{-1}\mathbf{D}(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}}) \\ & \leq k\widehat{\sigma^2}F_{k,n-k}(0; 1-\alpha)\Big\} \geq 1-\alpha-\varepsilon.\end{aligned}$$

It is to be remarked that in the case $\mathbf{D} = \mathbf{I}$, i.e. the confidence ellipsoid for the parameter $\delta\boldsymbol{\beta}$ must be determined, then the equality $C^{(ell,\delta\boldsymbol{\beta})} = K^{(par)}$ can be used.

In the case that σ^2 must be estimated, the decision whether linearization of the model with respect to the confidence ellipsoid can be used, is made with some uncertainty. Therefore a comparison of the given procedure with the exact determination, which is in this case known (cf. [8]), is interesting.

Lemma 3.2 *Let in the model* $\mathbf{Y} \sim N_n[\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Sigma}]$, $\boldsymbol{\beta} \in R^k$, *the matrix* $\boldsymbol{\Sigma}$ *be known.*

(i) Then the set

$$\mathcal{C}_\beta = \left\{\boldsymbol{\beta}\colon [\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}]' \left(\mathbf{P}_{F(\beta)}^{\Sigma^{-1}}\right)' \boldsymbol{\Sigma}^{-1}\mathbf{P}_{F(\beta)}^{\Sigma^{-1}}[\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}] \leq \chi_k^2(0; 1-\alpha)\right\}$$

is the exact $(1-\alpha)$*-confidence region for the parameter* $\boldsymbol{\beta}$.

(ii) If the matrix $\boldsymbol{\Sigma}$ *is of the form* $\boldsymbol{\Sigma} = \sigma^2\mathbf{V}$, *where* $\mathbf{V}$ *is a given* $n \times n$ *p.d. matrix and* σ^2 *is unknown parameter, then the exact* $(1-\alpha)$*-confidence set for the parameter* $\boldsymbol{\beta}$ *is*

$$\begin{aligned}\mathcal{D}_\beta = \Bigg\{&\boldsymbol{\beta}\colon [\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}]' \left(\mathbf{P}_{F(\beta)}^{V^{-1}}\right)' \mathbf{V}^{-1}\mathbf{P}_{F(\beta)}^{V^{-1}}[\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}] \\ &\leq \frac{k}{n-k}[\mathbf{Y} - \mathbf{f}(\boldsymbol{\beta})]' \left(\mathbf{M}_{F(\beta)}^{V^{-1}}\right)' \mathbf{V}^{-1}\mathbf{M}_{F(\beta)}^{V^{-1}}[\mathbf{Y} - \mathbf{f}(\boldsymbol{\beta})]F_{k,n-k}(0; 1-\alpha)\Bigg\}.\end{aligned}$$

Numerical determination of the exact confidence regions is tedious and time consuming unlike procedure given by a linearization.

4 Numerical example

Consider the Michaelis–Menten model, i.e.

$$f_i(\beta_1, \beta_2) = \frac{x_i\beta_1}{x_i + \beta_2}, \quad x_i = 1, 2, 3, 4, 6$$

and $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, $\sigma = 0.1$.

If $\beta_1 = 4$ and $\beta_2 = 1$, then (cf. [12])

$$\{\mathbf{F}\}_{i,\cdot} = \left(\frac{x_i}{1+x_i}, -\frac{4x_i}{(1+x_i)^2}\right), \quad i = 1, 2, 3, 4, 5,$$

$$\mathbf{F}_i = \begin{pmatrix} 0, & -\frac{x_i}{(1+x_i)^2} \\ -\frac{x_i}{(1+x_i)^2}, & \frac{8x_i}{(1+x_i)^3} \end{pmatrix}, \quad i = 1, 2, 3, 4, 5,$$

$$K_0^{(int)} = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{M}_F\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{F}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k\right\} = 0.3326,$$

$$K_0^{(par)} = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{P}_F\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{F}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k\right\} = 1.3212.$$

Let $\sigma(= 0.1)$ be known and let $\varepsilon = 0.05$, i.e. $\delta_{\max} = 0.6398$. Then the linearization region for the confidence ellipsoid is

$$\mathcal{L}_{\mathcal{E}} = \left\{ \delta\boldsymbol{\beta}\colon \delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{F}\delta\boldsymbol{\beta} \leq \sigma \frac{2\sqrt{\delta_{\max}}}{K_0^{(par)}(\boldsymbol{\beta}_0)} \right\}$$

and the 0.95-confidence ellipsoid for $\delta\boldsymbol{\beta}$ is

$$\mathcal{E} = \left\{ \mathbf{u}\colon (\mathbf{u} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}}) \leq 0.0599 \right\}$$

cf. Fig. 1.

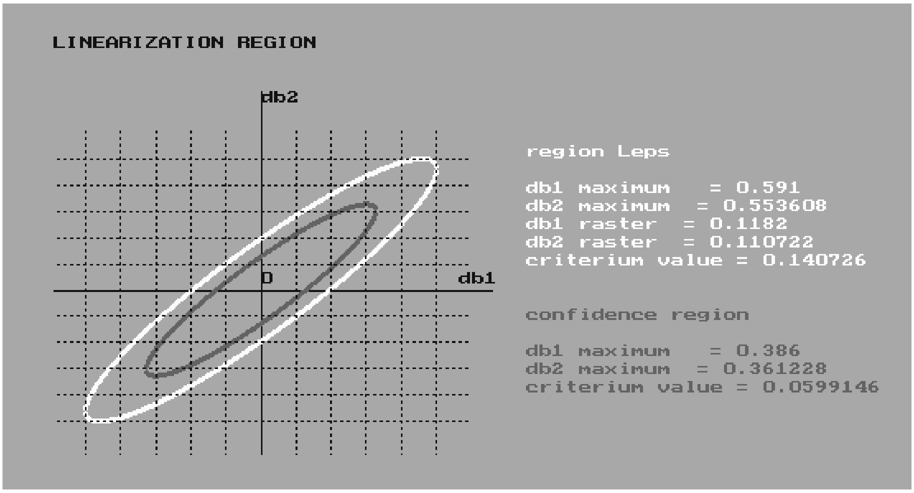

Fig. 1: 0.95-confidence ellipse for $\delta\boldsymbol{\beta}$ and the region $\mathcal{L}_{\mathcal{E}}$

Let set of measured data $\mathbf{y}$ are simulated for $\sigma = 0.1$, i.e.

$$\mathbf{y} = (1.90, 2.57, 3.08, 3.13, 3.58)'.$$

If $\delta_1/\delta_2 = (K^{(par)}/K^{(int)})^2 = 15.779\,478 = t$, then the set $C_{\delta_1^*,\delta_2^*}$ consists of a single point which is a solution of the equations

$$\int_0^{\frac{3}{2}9.552} \left[\int_0^{\infty} g_{2,\delta_1^*}(uv) g_{3,\delta_2^*}(v) v dv \right] du = 0.95 - 0.05, \quad \delta_1^* = t\delta_2^*,$$

where

$$g_{f,\delta}(y) = \begin{cases} \exp[-(y+\delta)/2] \sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{\delta}{2}\right)^r \frac{y^{r+(f/2)-1}}{2^{r+f/2}\Gamma(r+f/2)}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

In this case the linearization region from Theorem 3.3 is given in Fig. 2.

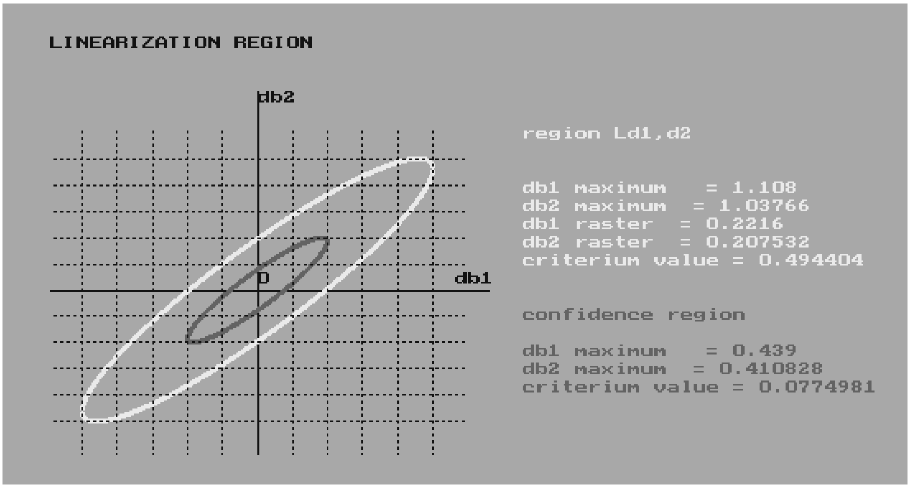

Fig. 2: The region $\mathcal{L}_{\delta_1,\delta_2}$ and 0.95-confidence ellipse (5)

The set $\mathcal{D}_\beta$ from Lemma 3.2 is given for $1-\alpha = 0.90$ at Fig. 3.

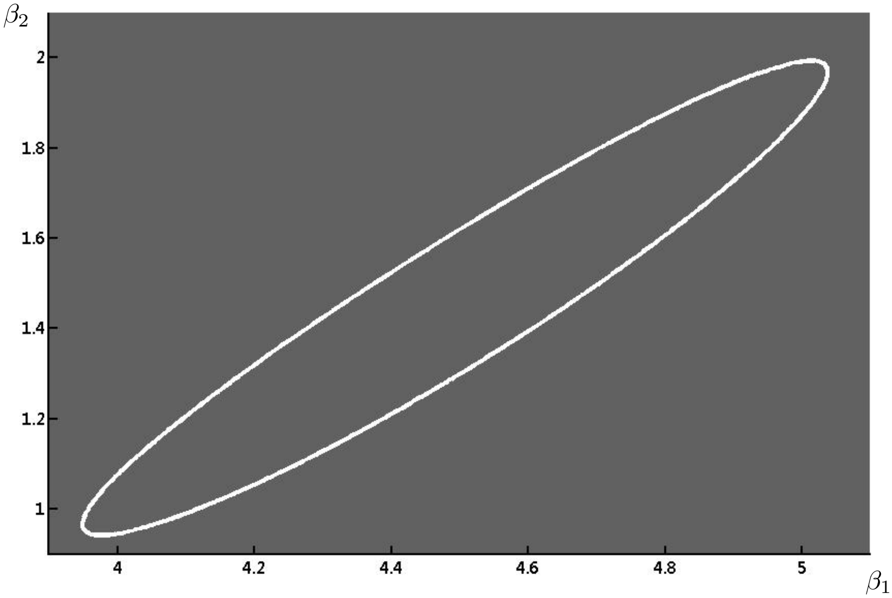

Fig. 3: The set $\mathcal{D}_\beta$ from Lemma 3.2

Acknowledgement Authors would like to express their thanks to Mgr. J. Marek, Ph.D. for his thoroughly check of the manuscript and a help with numerical calculation.

References

[1] Bates, D. M., Watts, D. G.: *Relative curvature measures of nonlinearity.* J. Roy. Stat. Soc. **B 42** (1980), 1–25.

[2] Kubáček, L., Kubáčková, L., Volaufová, J.: Statistical Models with Linear Structures. *Veda (Publishing House of Slovak Academy of Science), Bratislava*, 1995.

[3] Kubáček, L.: *On a linearization of regression models.* Applications of Mathematics **40** (1995), 61–78.

[4] Kubáček, L., Kubáčková, L.: *Regression models with a weak nonlinearity.* Technical report Nr. 1998.1, Universität Stuttgart, 1998, 1–67.

[5] Kubáček, L., Kubáčková, L.: Statistics and Metrology. *Vyd. Univ. Palackého, Olomouc*, 2000 (in Czech).

[6] Kubáček, L., Tesaříková, E.: *Confidence regions in nonlinear models with constraints.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **42**, 2003, 407–426.

[7] Kubáčková, L.: Foundations of Experimental Data Analysis. *CRC-Press, Boca Raton-Ann Arbor–London–Tokyo*, 1992.

[8] Pázman, A.: Nonlinear Statistical Models. *Kluwer Academic Publisher, Dordrecht–Boston–London and Ister Science Press, Bratislava*, 1993.

[9] Scheffé, H.: The Analysis of Variance (fifth printing). J. Wiley, New York–London–Sydney, 1967.

[10] Tesaříková, E., Kubáček, L.: *How to deal with regression models with a weak nonlinearity.* Discuss. Math., Probab. Stat. **21**, 2001, 21–48.

[11] Tesaříková, E., Kubáček, L.: *Estimators of dispersion in models with constraints (demoprogram).* Dept. Algebra and Geometry, Fac. Sci., Palacký Univ., Olomouc, 2003.

[12] Tesaříková, E., Kubáček, L.: *Linearization regions for confidence ellipsoids (demoprogram).* Department of Algebra and Geometry, Faculty of Dept. Algebra and Geometry, Fac. Sci., Palacký Univ., Olomouc, 2007.