Discussiones Mathematicae
Probability and Statistics 25 (2005) 115–133

WEAKLY NONLINEAR REGRESSION MODEL WITH CONSTRAINTS I: NONLINEAR HYPOTHESIS

Lubomír Kubáček *

Department of Mathematical Analysis and Applied Mathematics,
Faculty of Science, Palack Univesity,
Tomkova 40, CZ–779 00 Olomouc

and

Eva Tesaříková

Department of Algebra and Geometry,
Faculty of Science, Palack Univesity
Tomkova 40, CZ–779 00 Olomouc

Abstract

The problem considered is under which conditions in weakly nonlinear regression model with constraints I a weakly nonlinear hypothesis can be tested by linear methods. The aim of the paper is to find a region around the approximate value of the regression parameter with the following property. If we are certain that the actual value of the regression parameter is in this region, then the linear method of testing can be used without any significant deterioration of the inference.

Keywords: regression model with constraints, nonlinear hypothesis, linearization.

2000 Mathematics Subject Classification: 62F03, 62J05.

---

*Supported by the Council of Czech Government J14/98:153 1000 11.

Introduction

If the nonlinearity of the hypothesis on model parameters is not significant and the nonlinearity of the model is weak, then the hypothesis can be tested by linear methods. In the following text the statement will be elaborated stated more precisely.

It is necessary to find some measures of nonlinearity of the test problem and on the basis of it to state some conditions for the linearization. This condition is given in the form of the inclusion $\mathcal{E} \subset \mathcal{L}_{T,I}$ which must occur with a sufficiently high probability. Here $\mathcal{E}$ is the $(1-\alpha)$-confidence region of the model parameters (for sufficiently small α) and $\mathcal{L}_{T,I}$ is a special set in parameter space. The aim of the paper is to determine the set $\mathcal{L}_{T,I}$ (linearization region).

1. Notation

Let $\mathbf{Y} \sim N_n[\mathbf{f}(\boldsymbol{\beta}), \sigma^2\mathbf{V}]$ be the regression model under consideration. Here $\mathbf{Y}$ is the n-dimensional normally distributed observation vector, $\mathbf{f}(\boldsymbol{\beta})$ is the mean value of the vector $\mathbf{Y}$, $\boldsymbol{\beta}$ is an unknown k-dimensional parameter, $\sigma^2\mathbf{V}$ is the covariance matrix of the vector $\mathbf{Y}$, σ^2 is aknown/unknown parameter and $\mathbf{V}$ is a given $n \times n$ positive definite matrix. The constraints I are given by the equality $\mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}$, where $\mathbf{g}(\cdot)$ is a q-dimensional vector function of the parameter $\boldsymbol{\beta}$. (The notation "constraints I" is used in order to dinstinguish the "constraints II" where exept the parameter $\boldsymbol{\beta}$ also another parameter, e.g., $\boldsymbol{\gamma}$ occurs.) The null hypothesis H_0 is given in the form $\mathbf{t}(\boldsymbol{\beta}) = \mathbf{0}$, where $\mathbf{t}(\cdot)$ is a t-dimensional vector function and the alternative is $H_a : \mathbf{t}(\boldsymbol{\beta}) \neq \mathbf{0}$.

The functions $\mathbf{f}(\cdot)$, $\mathbf{g}(\cdot)$ and $\mathbf{t}(\cdot)$ can be given in the form

$$\mathbf{f}(\boldsymbol{\beta}) = \mathbf{f}\left(\boldsymbol{\beta}^{(0)}\right) + \mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}), \quad \mathbf{g}(\boldsymbol{\beta}) = \mathbf{g}\left(\boldsymbol{\beta}^{(0)}\right) + \mathbf{G}\delta\boldsymbol{\beta} + \frac{1}{2}\gamma(\delta\boldsymbol{\beta}),$$

$$\mathbf{t}(\boldsymbol{\beta}) = \mathbf{t}\left(\boldsymbol{\beta}^{(0)}\right) + \mathbf{T}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\tau}(\delta\boldsymbol{\beta}),$$

where $\delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}^{(0)}$, $\boldsymbol{\beta}^{(0)}$ is an approximate value of the parameter $\boldsymbol{\beta}$,

$$\mathbf{F} = \frac{\partial \mathbf{f}(\mathbf{u})}{\partial \mathbf{u}'}\bigg|_{u=\beta^{(0)}}, \quad \mathbf{G} = \frac{\partial \mathbf{g}(\mathbf{u})}{\partial \mathbf{u}'}\bigg|_{u=\beta^{(0)}}, \quad \mathbf{T} = \frac{\partial \mathbf{t}(\mathbf{u})}{\partial \mathbf{u}'}\bigg|_{u=\beta^{(0)}},$$

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

$$\kappa_i(\delta\boldsymbol{\beta}) = (\delta\boldsymbol{\beta})' \frac{\partial^2 f_i(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}\bigg|_{u=\beta^{(0)}} \delta\boldsymbol{\beta}, \quad i = 1, \dots, n,$$

$$\boldsymbol{\gamma}(\delta\boldsymbol{\beta}) = [\gamma_1(\delta\boldsymbol{\beta}), \dots, \gamma_n(\delta\boldsymbol{\beta})]',$$

$$\gamma_i(\delta\boldsymbol{\beta}) = (\delta\boldsymbol{\beta})' \frac{\partial^2 g_i(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}\bigg|_{u=\beta^{(0)}} \delta\boldsymbol{\beta}, \quad i = 1, \dots, q,$$

$$\boldsymbol{\tau}(\delta\boldsymbol{\beta}) = [\tau_1(\delta\boldsymbol{\beta}), \dots, \tau_q(\delta\boldsymbol{\beta})]',$$

$$\tau_i(\delta\boldsymbol{\beta}) = (\delta\boldsymbol{\beta})' \frac{\partial^2 t_i(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}\bigg|_{u=\beta^{(0)}} \delta\boldsymbol{\beta}, \quad i = 1, \dots, t.$$

Let the rank of the $n \times k$ matrix $\mathbf{F}$ be $r(\mathbf{F}) = k < n$, the rank of the $q \times k$ matrix $\mathbf{G}$ be $r(\mathbf{G}) = q < k$ and the rank of the $t \times k$ matrix $\mathbf{T}$ be $r(\mathbf{T}) = t < k$. Let further

$$(1) \qquad r\begin{pmatrix} \mathbf{G} \\ \mathbf{T} \end{pmatrix} = q + t.$$

In what follows text it is assumed that the approximate value $\boldsymbol{\beta}^{(0)}$ is chosen in such a way that $\mathbf{g}(\boldsymbol{\beta}^{(0)}) = \mathbf{0}$ and $\mathbf{t}(\boldsymbol{\beta}^{(0)}) = \mathbf{0}$.

2. Determination of the region $\mathcal{L}_{T,I}$

The linearized form of the model and the hypothesis is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n(\mathbf{F}\delta\boldsymbol{\beta}, \sigma^2\mathbf{V}), \quad \mathbf{G}\delta\boldsymbol{\beta} = \mathbf{0}, \quad \mathbf{T}\delta\boldsymbol{\beta} = \mathbf{0}. \tag{2}$$

The quadratized form of the model and the hypothesis is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\left(\mathbf{F}\delta\boldsymbol{\beta} + \tfrac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}), \sigma^2\mathbf{V}\right),$$

$$\mathbf{G}\delta\boldsymbol{\beta} + \tfrac{1}{2}\boldsymbol{\gamma}(\delta\boldsymbol{\beta}) = \mathbf{0}, \quad \mathbf{T}\delta\boldsymbol{\beta} + \tfrac{1}{2}\boldsymbol{\tau}(\delta\boldsymbol{\beta}) = \mathbf{0}. \tag{3}$$

Lemma 2.1. *If the model* (**??**) *is valid, the test of the hypothesis is*

$$T(\mathbf{Y}) = \left(\widehat{\widehat{\delta\boldsymbol{\beta}}}\right)' \mathbf{T}' \left\{ \mathbf{T} \left[\mathrm{Var}\left(\hat{\hat{\delta\boldsymbol{\beta}}}\right) \right] \mathbf{T}' \right\}^{-1} \mathbf{T}\widehat{\widehat{\delta\boldsymbol{\beta}}} \sim \begin{cases} \chi_t^2(0) & if \;\; H_0 \;\; is\ true, \\ \chi_t^2(\delta) & if \;\; H_0 \;\; is\ not\ true. \end{cases}$$

Here $\chi_t^2(0)$ is a random variable with a central chi-square distribution and with degrees of freedom equal to t. Analogously, $\chi_t^2(\delta)$ is a random variable with a noncentral chi-square distribution, with the parameter of noncentrality equal to δ and with degrees of freedom equal to t,

$$\widehat{\widehat{\delta\boldsymbol{\beta}}} = \left[\mathbf{I} - \mathbf{C}_0^{-1}\mathbf{G}'\left(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}'\right)^{-1}\mathbf{G}\right]\widehat{\delta\boldsymbol{\beta}},$$

$$\widehat{\delta\boldsymbol{\beta}} = \mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\left[\mathbf{Y} - \mathbf{f}\left(\boldsymbol{\beta}^{(0)}\right)\right], \quad \mathbf{C}_0 = \mathbf{F}'\mathbf{V}^{-1}\mathbf{F},$$

$$\mathrm{Var}\left(\hat{\hat{\delta\boldsymbol{\beta}}}\right) = \sigma^2\left[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'\left(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}'\right)^{-1}\mathbf{G}'\mathbf{C}_0^{-1}\right] = \sigma^2\left(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'}\right)^{+}$$

(the notation $^+$ means the Moore–Penrose generalized inverse and $\mathbf{M}_{G'} = \mathbf{I} - \mathbf{P}_{G'}, \mathbf{P}_{G'} = \mathbf{G}'(\mathbf{G}\mathbf{G}')^-\mathbf{G}$, where $^-$ means a generalized inverse; in more detail cf. [?]) and the parameter of noncentrality δ is

$$\delta = \left[E\left(\widehat{\widehat{\delta\boldsymbol{\beta}}}\right)\right]'\mathbf{T}'\left\{\mathbf{T}\left[\mathrm{Var}\left(\widehat{\widehat{\delta\boldsymbol{\beta}}}\right)\right]\mathbf{T}'\right\}^{-1}\mathbf{T}E\left(\widehat{\widehat{\delta\boldsymbol{\beta}}}\right).$$

Proof. Cf. [?], Chapter 4. It is to be remarked that even if the matrix $\mathrm{Var}(\hat{\hat{\boldsymbol{\beta}}})$ is singular, the matrix $\mathbf{T}'[\mathrm{Var}(\hat{\hat{\boldsymbol{\beta}}})]\mathbf{T}$ is regular, what is implied by the assumption (**??**). ■

Lemma 2.2. *If the model* (**??**) *is true, then under the null hypothesis* $H_0 : \mathbf{T}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\tau}(\delta\boldsymbol{\beta}) = \mathbf{0}$, *the following is valid*

$$\frac{1}{\sigma^2}\left(\widehat{\widehat{\delta\boldsymbol{\beta}}}\right)'\mathbf{T}'\left\{\mathbf{T}\left[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'\left(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}'\right)^{-1}\mathbf{G}\mathbf{C}_0^{-1}\right]\mathbf{T}'\right\}^{-1}\mathbf{T}\widehat{\widehat{\delta\boldsymbol{\beta}}} \sim \chi_t^2(\Delta).$$

Here

$$\Delta = \frac{1}{\sigma^2}E\left[\left(\widehat{\widehat{\delta\boldsymbol{\beta}}}\right)'\right]\mathbf{T}'\left\{\mathbf{T}\left[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'\left(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}'\right)^{-1}\mathbf{G}\mathbf{C}_0^{-1}\right]\mathbf{T}\right\}^{-1}\mathbf{T}E\left[\widehat{\widehat{\delta\boldsymbol{\beta}}}\right]$$

and

$$\mathbf{T}E\left(\widehat{\widehat{\delta\boldsymbol{\beta}}}\right) = -\frac{1}{2}\boldsymbol{\tau}(\delta\boldsymbol{\beta})\,\frac{1}{2}\mathbf{T}\mathbf{C}_0^{-1}\mathbf{G}'\left(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}'\right)^{-1}\boldsymbol{\gamma}(\delta\boldsymbol{\beta})$$
$$+\frac{1}{2}\mathbf{T}\mathbf{M}^{C_0}_{C_0^{-1}G'}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \frac{1}{2}\boxed{\mathbf{t}}_I(\delta\boldsymbol{\beta}).$$

Here

$$\mathbf{M}^{C_0}_{C_0^{-1}G'} = \mathbf{I} - \mathbf{P}^{C_0}_{C_0^{-1}G'},$$

$$\mathbf{P}^{C_0}_{C_0^{-1}G'} = \mathbf{C}_0^{-1}\mathbf{G}'\left(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{C}_0\mathbf{C}_0^{-1}\mathbf{G}'\right)^{-1}\mathbf{G}\mathbf{C}_0^{-1}\mathbf{C}_0.$$

Proof.

$$\mathbf{T}E(\widehat{\widehat{\delta\boldsymbol{\beta}}}) = \mathbf{T}\mathbf{M}^{C_0}_{C_0^{-1}G'}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\left(\mathbf{F}\delta\boldsymbol{\beta} + \tfrac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})\right)$$

$$= \mathbf{T}\mathbf{M}^{C_0}_{C_0^{-1}G'}\left(\delta\boldsymbol{\beta} + \mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})\right)$$

$$= \mathbf{T}\left(\mathbf{I} - \mathbf{P}^{C_0}_{C_0^{-1}G'}\right)\delta\boldsymbol{\beta} + \mathbf{T}\mathbf{M}^{C_0}_{C_0^{-1}G'}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})$$

$$= \mathbf{T}\delta\boldsymbol{\beta} - \mathbf{T}\mathbf{P}^{C_0}_{C_0^{-1}G'}\delta\boldsymbol{\beta} + \mathbf{T}\mathbf{M}^{C_0}_{C_0^{-1}G'}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})$$

Under the null hypothesis $\mathbf{T}\delta\boldsymbol{\beta} = -\frac{1}{2}\boldsymbol{\tau}(\delta\boldsymbol{\beta})$ and regarding the constraints we have

$$-\mathbf{T}\mathbf{P}^{C_0}_{C_0^{-1}G'}\delta\boldsymbol{\beta} = -\mathbf{T}\mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\delta\boldsymbol{\beta}$$

$$= \mathbf{T}\mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\frac{1}{2}\boldsymbol{\gamma}(\delta\boldsymbol{\beta}).$$

Thus

$$\mathbf{T}E(\widehat{\widehat{\delta\boldsymbol{\beta}}}) = -\frac{1}{2}\boldsymbol{\tau}(\delta\boldsymbol{\beta}) + \mathbf{T}\mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\frac{1}{2}\boldsymbol{\gamma}(\delta\boldsymbol{\beta})$$

$$+ \mathbf{T}\mathbf{M}^{C_0}_{C_0^{-1}G'}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \frac{1}{2}\boxed{\mathbf{t}}_I(\delta\boldsymbol{\beta}).$$

∎

Definition 2.3. Let

$$K_I^{(test)}(\boldsymbol{\beta}_0) = \sup\left\{\frac{1}{\sigma^3}\frac{\sqrt{\left(\boxed{\mathbf{t}}_I(\mathbf{K}_{(G',T')'}\delta\mathbf{s})\right)'[\mathbf{T}(\mathbf{M}_G'\mathbf{C}_0\mathbf{M}_{G'})^+\mathbf{T}']^{-1}\boxed{\mathbf{t}}_I(\mathbf{K}_{(G',T')'}\delta\mathbf{s})}}{\delta\mathbf{s}'\mathbf{K}'_{(G',T')'}(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'})^+\mathbf{K}_{(G',T')'}\delta\mathbf{s}} : \delta\mathbf{s}\in R^{k-(q+t)}\right\} = \frac{1}{\sigma^3}K_{I,0}^{(test)}(\boldsymbol{\beta}_0)$$

be a measure of nonlinearity for the test in a regression model with constraints I. Here $\mathbf{K}_{(G',T')'}$ is a $k\times(k-q-t)$ matrix with the property

$$\mathcal{K}er\begin{pmatrix}\mathbf{G}\\ \mathbf{T}\end{pmatrix} = \mathcal{M}(\mathbf{K}_{(G',T')'})$$

and

$$K_{I,0}^{(test)}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\left(\boxed{\mathbf{t}}_I(\mathbf{K}_{(G',T')'}\delta\mathbf{s})\right)'[\mathbf{T}(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'})^+\mathbf{T}']^{-1}\boxed{\mathbf{t}}_I(\mathbf{K}_{(G',T')'}\delta\mathbf{s})}}{\delta\mathbf{s}'\mathbf{K}'_{(G',T')'}(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'})^+\mathbf{K}_{(G',T')'}\delta\mathbf{s}} : \delta\mathbf{s}\in R^{k-(q+t)}\right\}.$$

Theorem 2.4. *Let δ_{max} be a solution to the equation*

$$P\left\{\chi_t^2(\delta_{max}) \geq \chi_t^2(0;1-\alpha)\right\} = \alpha+\varepsilon$$

and

$$\mathcal{L}_{T,I} =$$

$$\left\{\delta\boldsymbol{\beta} : \delta\boldsymbol{\beta} = \mathbf{K}_{(G',T')'}\delta\mathbf{s}, \delta\mathbf{s}'\mathbf{K}'_{(G',T')'}(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'})^+\mathbf{K}_{(G',T')'}\delta\mathbf{s} \leq \frac{2\sigma\sqrt{\delta_{max}}}{K_{I,0}^{(test)}(\boldsymbol{\beta}_0)}\right\}.$$

Then

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{T,I} \Rightarrow P_{H_0}\{T(\mathbf{Y}) \geq \chi_t^2(0; 1-\alpha)\} \leq \alpha + \varepsilon.$$

Here $\chi_t^2(0;1-\alpha)$ is the $(1-\alpha)$-quantile of the central chi-square distribution with t degrees of freedom.

Proof. With respect to Definition 2.3 we have

$$\left(\boxed{\mathbf{t}}_I\left(\mathbf{K}_{(G',T')'}\delta\mathbf{s}\right)\right)'\left[\mathbf{T}(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'})^+\mathbf{T}'\right]^{-1}\boxed{\mathbf{t}}_I\left(\mathbf{K}_{(G',T')'}\delta\mathbf{s}\right)$$

$$\leq \left[\delta\mathbf{s}'\mathbf{K}'_{(G',T')'}\left(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'}\right)^+\mathbf{K}_{(G',T')'}\delta\mathbf{s}\right]^2\left(K_{I,0}^{(test)}(\boldsymbol{\beta}_0)\right)^2.$$

If

$$\left[\delta\mathbf{s}'\mathbf{K}'_{(G',T')'}\left(\mathbf{M}_{G'}\mathbf{C}_0\mathbf{M}_{G'}\right)^+\mathbf{K}_{(G',T')'}\delta\mathbf{s}\right]^2\left(K_{I,0}^{(test)}\left(\boldsymbol{\beta}_0\right)^2 \leq 4\sigma^2\delta_{max},\right.$$

then

$$\Delta = \frac{1}{4\sigma^2}\left(\boxed{\mathbf{t}}_I(\mathbf{K}_{(G',T')'}\delta\mathbf{s})\right)'[\mathbf{T}\left(\mathbf{M}'_G\mathbf{C}_0\mathbf{M}_{G'}\right)^+\mathbf{T}']^{-1}\boxed{\mathbf{t}}_I\left(\mathbf{K}_{(G',T')'}\delta\mathbf{s}\right) \leq \delta_{max}$$

and

$$\alpha \leq P\left\{\chi_t^2(\Delta) \geq \chi_t^2(0;1-\alpha) \leq P\left\{\chi_t^2(\delta_{max}) \geq \chi_t^2(0;1-\alpha)\right.\right. = \alpha + \varepsilon.$$

∎

Lemma 2.5. *The* $(1-\alpha)$*-confidence region in the model* (**??**) *is*

$$\mathcal{E}=\left\{\boldsymbol{\beta}:\boldsymbol{\beta}=\hat{\hat{\boldsymbol{\beta}}}+\mathbf{k}_G,\mathbf{k}_G\in\mathcal{K}er(\mathbf{G}),\mathbf{k}_G'(\mathbf{M}_G\mathbf{C}\mathbf{M}_G)^{+}\mathbf{k}_G\leq\sigma^2\chi^2_{k-q}(0;1-\alpha)\right\}.$$

Proof. Cf. [**?**] Chapter 2 Example 2.3.6. ■

Corollary 2.6. *If*

$$\sigma << \frac{2\sqrt{\delta_{max}}}{\chi^2_{k-q}(0;1-\alpha)K^{(test)}_{0,I}(\boldsymbol{\beta}_0)},$$

then $\mathcal{L}_{T,I}$ *from Theorem* 2.3 *is a linearization region for sufficiently small* α *and* ε, *respectively.*

Proof. It is a direct consequence of Theorem 2.4 and Lemma 2.5. ■

Remark 2.7. It is to be remarked that $\mathcal{E}$ is the $(1-\alpha)$-confidence ellipsoid in the case of a weakly nonlinearity of the model only. It needs some special investigation whether $\mathcal{E}$ can be used as the confidence region with a sufficiently high probability (in more detail cf. [**?**] and [**?**]), however this problem is out of the scope of the paper. Nevertheless, the inequality from Corollary 2.6 provides good information on the nonlinearity of the model with respect to the test considered. If in a given case we are not sure whether the confidence region with a sufficiently high probability is included into $\mathcal{L}_{T,I}$, then the exact formula for $\mathcal{E}$, i.e.

$$\left\{\boldsymbol{\beta}:[\mathbf{Y}-\mathbf{f}(\boldsymbol{\beta})]'\mathbf{V}^{-1}\mathbf{P}^{V^{-1}}_{F(\beta)K_G(\beta)}[\mathbf{Y}-\mathbf{f}(\boldsymbol{\beta})]\leq\sigma^2\chi^2_{k-q}(0;1-\alpha\right\},$$

where

$$\mathbf{F}(\boldsymbol{\beta}) = \partial\mathbf{f}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}', \quad \mathcal{M}(\mathbf{K}_G(\boldsymbol{\beta}))=\mathcal{K}er[\mathbf{G}(\boldsymbol{\beta})],$$

$$\mathbf{G}(\boldsymbol{\beta}) = \partial\mathbf{g}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}',$$

$$\mathbf{P}^{V^{-1}}_{F(\beta)K_G(\beta)} = \mathbf{F}(\boldsymbol{\beta})\left\{\mathbf{C}_0^{-1}(\boldsymbol{\beta}) - \mathbf{C}_0^{-1}(\boldsymbol{\beta})\mathbf{G}'(\boldsymbol{\beta})\left[\mathbf{G}(\boldsymbol{\beta})\mathbf{C}_0^{-1}(\boldsymbol{\beta})\mathbf{G}'(\boldsymbol{\beta})\right]^{-1}\times\right.$$

$$\left.\times\mathbf{G}(\boldsymbol{\beta})\mathbf{C}_0^{-1}(\boldsymbol{\beta})\right\}\mathbf{F}'(\boldsymbol{\beta})\mathbf{V}^{-1} = \mathbf{P}^{V^{-1}}_{F(\beta)}\mathbf{P}^{C_0(\beta)}_{\mathcal{K}er([G(\beta)]}$$

can be used. (For more detail cf. [?], pp. 53–54 and [?] Remark 2.1.)

In some case (cf. following a numerical example) the region $\mathcal{L}_{T,I}$ can be extremely large. Then it is necessary to check whether the nonlinear problem can be linearized with respect to other statistical inference, mainly with respect to the bias of the estimator of $\boldsymbol{\beta}$, and to use a smaller linearization region.

Remark 2.8. The influence of nonlinearity on the power function was investigated in the null hypothesis only. In an alternative hypothesis the behaviour of the power function is characterized by the following expression for the noncentrality parameter δ

$$\delta = \frac{1}{4\sigma^2}\left[\boxed{\mathbf{t}}_I(\delta\boldsymbol{\beta})\right]'\left\{\mathbf{T}\left[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}_0^{-1}\right]\mathbf{T}'\right\}^{-1}\boxed{\mathbf{t}}_I(\delta\boldsymbol{\beta})$$

$$+\frac{1}{\sigma^2}\mathbf{d}'\left\{\mathbf{T}\left[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}_0^{-1}\right]\mathbf{T}'\right\}^{-1}\boxed{\mathbf{t}}_I(\delta\boldsymbol{\beta})$$

$$+\frac{1}{\sigma^2}\mathbf{d}'\left\{\mathbf{T}\left[\mathbf{C}_0^{-1} - \mathbf{C}_0^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}_0^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}_0^{-1}\right]\mathbf{T}'\right\}^{-1}\mathbf{d},$$

where $\mathbf{0} \neq \mathbf{d} = \mathbf{T}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\tau}(\delta\boldsymbol{\beta})$.

Thus a value $\mathbf{d}_{min} = -\frac{1}{2}\boxed{\mathbf{t}}_I(\delta\boldsymbol{\beta})$ of the vector $\mathbf{d}$ gives $\delta_{min} = 0$. Therefore the value of the power function in the corresponding alternative hypothesis is α, i.e., it is smaller than $\alpha + \varepsilon = P\{\chi_t^2(\Delta) \geq \chi_t^2(0; 1-\alpha)\}$. In the opposite direction of the vector $\mathbf{d}$ the value of the power function is larger than it is in the linear case. This case is perferable for testing.

3. NUMERICAL EXAMPLE

Let a part of the circle arc

$$y = \beta_2 + \sqrt{\beta_3^2 - x^2 + 2x\beta_1 - \beta_1^2}, \quad \beta_1 - \beta_3 \leq x \leq \beta_1 + \beta_3,$$

be measured at points $x_1, \dots, x_5$. Let it be known that a tangent of the arc at the point $x = 0$ is c_1. It gives the constraint

$$(g(\boldsymbol{\beta}) =)(1 + c_1^2)\beta_1^2 - c_1^2\beta_3^2 = 0.$$

The null hypothesis H_0 states

$$H_0 : (t(\boldsymbol{\beta}) =)\beta_2 - \sqrt{\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2} = 0$$

(i.e., the intersection of the arc and the axis of coordinates x is at the point $x = c_2$) and the alternative is

$$H_a : \beta_2 - \sqrt{\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2} \neq 0.$$

The numbers c_1 and c_2 are given.

Thus we obtain

$$\{\mathbf{F}(\boldsymbol{\beta})\}_{i,.} = \left(\frac{x_i - \beta_1}{\sqrt{\beta_3^2 - x_i^2 + 2x_i\beta_1 - \beta_1^2}}, 1, \frac{\beta_3}{\sqrt{\beta_3^2 - x_i^2 + 2x_i\beta_1 - \beta_1^2}} \right),$$

$$i = 1, \dots, 5,$$

$$\mathbf{F}_i = \frac{\partial^2 f(x_i, \boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}$$

$$= \begin{pmatrix} \frac{-\beta_3^2}{(\beta_3^2 - x_i^2 + 2x_i\beta_1 - \beta_1^2)^{3/2}}, & 0, & \frac{-(x_i - \beta_1)\beta_3}{(\beta_3^2 - x_i^2 + 2x_i\beta_1 - \beta_1^2)^{3/2}} \\ 0, & 0, & 0 \\ \frac{-(x_i - \beta_1)\beta_3}{(\beta_3^2 - x_i^2 + 2x_i\beta_1 - \beta_1^2)^{3/2}}, & 0, & \frac{-(x_i - \beta_1)^2}{(\beta_3^2 - x_i^2 + 2x_i\beta_1 - \beta_1^2)^{3/2}} \end{pmatrix},$$

$$i = 1, \dots, 5,$$

$$\mathbf{G} = \left[2\beta_1 \left(1 + c_1^2\right), 0, -2c_1^2\beta_3\right],$$

$$\mathbf{G}_1 = \frac{\partial^2 g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} = \begin{pmatrix} 2(1 + c_1^2), & 0, & 0 \\ 0, & 0, & 0 \\ 0, & 0, & -2c_1^2 \end{pmatrix},$$

$$\mathbf{T} = \left(\frac{-(c_2 - \beta_1)}{\sqrt{\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2}}, 1, \frac{-\beta_3}{\sqrt{\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2}} \right),$$

$$\mathbf{T}_1 = \frac{\partial^2 t(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}$$

$$= \begin{pmatrix} \frac{\beta_3^2}{(\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2)^{3/2}}, & 0, & \frac{(c_2 - \beta_1)\beta_3}{(\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2)^{3/2}} \\ 0, & 0, & 0 \\ \frac{(c_2 - \beta_1)\beta_3}{(\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2)^{3/2}}, & 0, & \frac{(c_2 - \beta_1)^2}{(\beta_3^2 - c_2^2 + 2c_2\beta_1 - \beta_1^2)^{3/2}} \end{pmatrix}.$$

Since in our case the matrix $\mathbf{K}_{(G',T')'}$ is a column $\mathbf{k}$,

$$\mathbf{k}' = \left(c_1^2\beta_3, \frac{c_1^2\beta_3(c_2-\beta_1)+\beta_1\beta_3(1+c_1^2)}{\sqrt{\beta_3^2-(c_2-\beta_1)^2}}, \beta_1(1+c_1^2) \right),$$

the quantity $K_{I,0}^{(test)}(\boldsymbol{\beta}_0)$ can be expressed as

$$K_{I,0}^{(test)}(\boldsymbol{\beta}_0) = \frac{|A|\sqrt{[\mathbf{T}(\mathbf{M}_{G'}\mathbf{C}\mathbf{M}_{G'})^+\mathbf{T}']^{-1}}}{\mathbf{k}'(\mathbf{M}_{G'}\mathbf{C}\mathbf{M}_{G'})^+\mathbf{k}},$$

where

$$\mathbf{C} = \mathbf{F}'\mathbf{F}, \quad (\mathbf{M}_{G'}\mathbf{C}\mathbf{M}_{G'})^+ = \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1},$$

$$A = -\mathbf{k}'\mathbf{T}_1\mathbf{k} + \mathbf{T}\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{k}'\mathbf{G}_1\mathbf{k} +$$

$$+\mathbf{T}((\mathbf{M}_{G'}\mathbf{C}\mathbf{M}_{G'})^+\mathbf{F}' \begin{pmatrix} \mathbf{k}'\mathbf{F}_1\mathbf{k} \\ \vdots \\ \mathbf{k}'\mathbf{F}_5\mathbf{k} \end{pmatrix}.$$

Let $\alpha = 0.05$ and $\varepsilon = 0.05$, respectively. Let

x_1	$\beta_1 + (c_2 - \beta_1)0.32$
x_2	$\beta_1 + (c_2 - \beta_1)0.44$
x_3	$\beta_1 + (c_2 - \beta_1)0.56$
x_4	$\beta_1 + (c_2 - \beta_1)0.68$
x_5	$\beta_1 + (c_2 - \beta_1)0.80$

(If the parameters $\beta_1, \beta_2, \beta_3$ are chosen, then $c_1 = \beta_1/\sqrt{\beta_3^2 - \beta_2^2}$ and $c_2 = \beta_1 + \sqrt{\beta_3^2 - \beta_2^2}$.)

Four points in the parameter space are chosen, i.e.,

$$\boldsymbol{\beta}_1 = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}; \boldsymbol{\beta}_2 = \begin{pmatrix} 10 \\ 10 \\ 20 \end{pmatrix}; \boldsymbol{\beta}_3 = \begin{pmatrix} 100 \\ 100 \\ 200 \end{pmatrix}; \boldsymbol{\beta}_4 = \begin{pmatrix} 1000 \\ 1000 \\ 2000 \end{pmatrix}.$$

In this case

$$\mathbf{C}_0(\boldsymbol{\beta}_1) = \begin{pmatrix} 253.71, & 419.18, & -482.68 \\ 419.18, & 710.94, & -812.92 \\ -482.68, & -812.92, & 931.40 \end{pmatrix} \approx$$

$$\approx \mathbf{C}_0(\boldsymbol{\beta}_4) = \begin{pmatrix} 253.83, & 419.39, & -482.93 \\ 419.39, & 711.30, & -813.33 \\ -482.93, & -813.33, & 931.87 \end{pmatrix},$$

$$\frac{1}{\sigma^2}\mathrm{Var}(\widehat{\widehat{\delta\boldsymbol{\beta}}}_1) = \begin{pmatrix} 0.8564, & -2.522, & 1.7128 \\ -2.5223, & 7.6269, & -5.044 \\ 1.7128, & -5.044, & 3.4256 \end{pmatrix} =$$

$$= \frac{1}{\sigma^2}\mathrm{Var}(\widehat{\widehat{\delta\boldsymbol{\beta}}}_4) = \begin{pmatrix} 0.8564, & -2.522, & 1.7128 \\ -2.5223, & 7.6269, & -5.044 \\ 1.7128, & -5.044, & 3.4256 \end{pmatrix}.$$

Further

$$K_{0,I}^{(test)}(\boldsymbol{\beta})$$

$\boldsymbol{\beta}_1$	$\boldsymbol{\beta}_2$	$\boldsymbol{\beta}_3$	$\boldsymbol{\beta}_4$
0.012 3278	0.001 23262	0.000 123 253	0.000 012 324

Let us denote

$$\sigma_{crit} = \frac{2\sqrt{\delta_{max}}}{K_{I,0}^{(test)}(\boldsymbol{\beta}_0)\chi_1^2(0; 0.95)}$$

(cf. Corollary 2.6). Since $\delta_{max} = 0.639767, \chi_1^2(0; 0.95) = 3.84$, we have

$$\sigma_{crit}$$

$\boldsymbol{\beta}_1$	$\boldsymbol{\beta}_2$	$\boldsymbol{\beta}_3$	$\boldsymbol{\beta}_4$
33.793	337.972	3 379.965	33 803.214

These values are extremely large and that means that the models are extremely weak nonlinear with respect to testing the given hypothesis.

In order to demonstrate numerically the behaviour of the test let us shift the values $\boldsymbol{\beta}_i, i = 1, 2, 3, 4$, into $\boldsymbol{\beta}_i + \delta\boldsymbol{\beta}_i, i = 1, 2, 3, 4$. The shift $\delta\boldsymbol{\beta}_i$ must respect the Bates and Watts [?] curvatures of the models (cf. also Remark 2.7), i.e.,

$$K_{I,0}^{(int)}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boxed{1}'\mathbf{V}^{-1}\mathbf{M}_{FK_G}^{V^{-1}}\boxed{1}}}{\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}} : \delta\mathbf{s} \in R^{k-q}\right\} = \frac{1}{\sigma}K_I^{(int)}(\boldsymbol{\beta}_0),$$

$$K_{I,0}^{(par)}(\boldsymbol{\beta}_0) = \sup\left\{\frac{\sqrt{\boxed{1}'\mathbf{V}^{-1}\mathbf{P}_{FK_G}^{V^{-1}}\boxed{1}}}{\delta\mathbf{s}'\mathbf{K}_G'\mathbf{C}_0\mathbf{K}_G\delta\mathbf{s}} : \delta\mathbf{s} \in R^{k-q}\right\} = \frac{1}{\sigma}K_I^{(par)}(\boldsymbol{\beta}_0)$$

(for more detail cf. [**?**]). The last two measures of nonlinearity enable us to determine the maximum of σ until which the linearization of the model with respect to the test of agreement between the linear model and the data ($K_{I,0}^{(int)}$) and with the respect to the bias of the estimator of $\boldsymbol{\beta}$ ($K_{I,0}^{(par)}$), respectively, is possible, i.e.,

$$\sigma_{max}^{(int)} = \frac{2\sqrt{\delta_{crit}^{(int)}}}{K_{I,0}^{(int)}(\boldsymbol{\beta}_0)\chi_{k-q}^2(0; 1-\alpha)}, \quad \sigma_{max}^{(par)} = \frac{2\varepsilon}{K_{I,0}^{(par)}(\boldsymbol{\beta}_0)\chi_{k-q}^2(0; 1-\alpha)},$$

where $\delta_{crit}^{(int)}$ is a solution to the equation

$$P\left\{\chi_{n+q-k}^2\left(\delta_{crit}^{(int)}\right) \geq \chi_{n+q-k}^2(0; 1-\alpha)\right\} = \alpha + \varepsilon.$$

In this way we obtain

	$\boldsymbol{\beta}_1$	$\boldsymbol{\beta}_2$	$\boldsymbol{\beta}_3$	$\boldsymbol{\beta}_4$
$K_{I,0}^{(int)}(\boldsymbol{\beta})$	0. 565 683	0.819 145	0.005 196	0.000 540
$\sigma_{max}^{(int)}$	0.472	4.874	51.385	494.629
$K_{I,0}^{(par)}(\boldsymbol{\beta})$	8.190	0.819	0.081 9	0.008 192
$\sigma_{max}^{(par)}$	0.002	0.020	0.203 7	2.037 4

The values $\sigma_{max}^{(par)}$ are used to demonstrate numerically the possibility of considering the models to be linear, when

$$\delta\boldsymbol{\beta}_i = [\mathbf{M}_{G'(\beta_i+\delta\beta_i)}\mathbf{C}_0(\boldsymbol{\beta}_i + \delta\boldsymbol{\beta}_i)\mathbf{M}_{G'(\beta_i+\delta\beta_i)}]^+ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mathbf{s}_i,$$

where

$$\mathbf{s}_i = \sigma_{max}^{(par)}(\boldsymbol{\beta}_i)\frac{\sqrt{5.99}}{\sqrt{(1,1,1)[\mathbf{M}_{G'(\beta_i+\delta\beta_i)}\mathbf{C}_0(\boldsymbol{\beta}_i + \delta\boldsymbol{\beta}_i)\mathbf{M}_{G'(\beta_i+\delta\beta_i)}]^+ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}}$$

(the point $\boldsymbol{\beta}_i + \delta\boldsymbol{\beta}_i$ is at the boundary of the 0.95 confidence ellipse for the parameter $\boldsymbol{\beta}_i$ with the center $\boldsymbol{\beta}_i$).

Now the values $\mathbf{Y}_i^j, i = 1, 2, 3, 4$ and $j = 1, \dots, 100$ are simulated for the parameter $\boldsymbol{\beta}_i + \delta\boldsymbol{\beta}_i$ and

$$P_{H_0}\left\{T(\mathbf{Y}) \geq \chi_1^2(0; 0.95)\right\}$$

is estimated by the value $\frac{1}{m}\sum_{j=1}^{m}\zeta_j$, where

$$T\left(\mathbf{Y}^{(j)}\right) \geq \chi_1^2(0; 0.95) \quad \Rightarrow \quad \zeta_j = 1$$

$$T\left(\mathbf{Y}^{(j)}\right) < \chi_1^2(0; 0.95) \quad \Rightarrow \quad \zeta_j = 0.$$

Then by using [?] we obtain the following values

$$\frac{1}{m}\sum_{j=1}^{m}\zeta_j$$

$\boldsymbol{\beta}_1$	$\boldsymbol{\beta}_2$	$\boldsymbol{\beta}_3$	$\boldsymbol{\beta}_4$
0.07	0.02	0.06	0.01

It is quite obvious that the nonlinearity is negligible in this case and the differences between $\frac{1}{m}\sum_{j=1}^{m}\zeta_j$ and $\alpha = 0.05$ are probably due to the relatively small number of simulations.

References

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

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

[3] L. Kubáček and L. Kubáčková, *Regression models with a weak nonlinearity*, Technical Report Nr 1998.1 Univerity of Stuttgart, 1998, 1–64.

[4] L. Kubáček and L. Kubáčková, *Statistics and Metrology* (in Czech), Publishing House of Palacký University, Olomouc 2000.

[5] L. Kubáček and L. Kubáčková, *Statistical problems of a determination of isobestic points*, Folia Fa. Sci. Nat. Univ. Masarykianae Brunensis, Mathematica **11** (2002), 139–150.

[6] L. Kubáček, *Linearized model with constraints* I, Application of Mathematics **48** (2003), 81–95.

[7] A. Pázman, *Nonlinear Statistical Models*, Kluwer Academic Publisher, Dordrecht-Boston-London- and Ister Science Press, Bratislava 1988.

[8] C.R. Rao, *Linear Statistical Inference and Its Application*, J. Wiley, New York-London-Sydney 1965.

[9] C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and Its Applications*, New York, J. Wiley 1971.

[10] E. Tesaříková and L. Kubáček, *A test in nonlinear regression models*, (in Czech) Demoprogram, Department of Algebra and Geometry, Faculty of Science, Palacký University, Olomouc 2004.

Received 27 April 2005