Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **49**, 2 (2010) 83–93

Underparametrization of Weakly Nonlinear Regression Models*

Lubomír KUBÁČEK [1], Eva TESAŘÍKOVÁ [2]

[1] *Department of Mathematical Analysis and Applications of Mathematics*
Faculty of Science, Palacký University
17. listopadu 12, 771 46 Olomouc, Czech Republic
e-mail: kubacekl@inf.upol.cz

[2] *Department of Algebra and Geometry, Faculty of Science, Palacký University*
17. listopadu 12, 771 46 Olomouc, Czech Republic
e-mail: tesariko@inf.upol.cz

(Received January 12, 2010)

Abstract

A large number of parameters in regression models can be serious obstacle for processing and interpretation of experimental data. One way how to overcome it is an elimination of some parameters. In some cases it need not deteriorate statistical properties of estimators of useful parameters and can help to interpret them. The problem is to find conditions which enable us to decide whether such favourable situation occurs.

Key words: Weakly nonlinear regression model, underparameterization, MSE, BLUE.

2000 Mathematics Subject Classification: 62J05, 62F10

Introduction

Real events and processes in many professions can be modelled adequately in many situations by help of a large number of parameters. If all of them can be interpreted by a professional language, then it is no reason to neglect them. Their elimination can lead to a misinterpretation of other parameters, since estimators of them are influenced by the underparametrization of the model.

Sometimes not all parameters can be interpreted in professional language and then a problem arises whether noninterpretable parameters can be neglected.

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

1 Prerequisities

Let $\mathbf{Y} \sim_n \left[\mathbf{f}(\boldsymbol{\beta}, \boldsymbol{\gamma}), \boldsymbol{\Sigma}\right]$ means that $\mathbf{Y}$ is an n-dimensional random vector with the mean value $E(\mathbf{Y})$ equal to $\mathbf{f}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ and with the covariance matrix $\mathrm{Var}(\mathbf{Y})$ equal to $\boldsymbol{\Sigma}$, which is assumed to be known and p.d. (positive definite). The function $\mathbf{f}(\cdot, \cdot\cdot)\colon R^{k+l} \to R^n$ (R^s means the s-dimensional linear vector space) can be, with sufficiently high accuracy, expressed as

$$\mathbf{f}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{f}_0 + \mathbf{F}\delta\boldsymbol{\beta} + \mathbf{S}\delta\boldsymbol{\gamma} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma}),$$

where $\mathbf{f}_0 = \mathbf{f}(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)$, $\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0$ are approximate values of the true values of the unknown vectors $\boldsymbol{\beta}, \boldsymbol{\gamma}$, $\delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$, $\delta\boldsymbol{\gamma} = \boldsymbol{\gamma} - \boldsymbol{\gamma}_0$,

$$\begin{aligned}
\mathbf{F} &= \partial\mathbf{f}(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)/\partial\boldsymbol{\beta}', \quad \mathbf{S} = \partial\mathbf{f}(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)/\partial\boldsymbol{\gamma}', \\
\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma}) &= \left[\kappa_1(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma}), \ldots, \kappa_n(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma})\right]', \\
\kappa_i(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma}) &= (\delta\boldsymbol{\beta}', \delta\boldsymbol{\gamma}') \begin{pmatrix} \mathbf{F}_{i,(1,1)}, & \mathbf{F}_{i,(1,2)} \\ \mathbf{F}_{i,(2,1)}, & \mathbf{F}_{i,(2,2)} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\boldsymbol{\gamma} \end{pmatrix}, \\
\mathbf{F}_i &= \begin{pmatrix} \mathbf{F}_{i,(1,1)}, & \mathbf{F}_{i,(1,2)} \\ \mathbf{F}_{i,(2,1)}, & \mathbf{F}_{i,(2,2)} \end{pmatrix}, \\
\mathbf{F}_{i,(1,1)} &= \frac{\partial^2 f_i(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}, \quad \mathbf{F}_{i,(1,2)} = \frac{\partial^2 f_i(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)}{\partial\boldsymbol{\beta}\partial\boldsymbol{\gamma}'}, \\
\mathbf{F}_{i,(2,1)} &= \frac{\partial^2 f_i(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\beta}'}, \quad \mathbf{F}_{i,(2,2)} = \frac{\partial^2 f_i(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}, \\
& i = 1, \ldots, n.
\end{aligned}$$

Let the rank $r(\mathbf{F}, \mathbf{S})$ of the matrix $(\mathbf{F}, \mathbf{S})$ be $r(\mathbf{F}, \mathbf{S}) = k + l$, where k is the dimension of the vector $\boldsymbol{\beta}$ and l is the dimension of the vector $\boldsymbol{\gamma}$.

Lemma 1.1 *In the underparametrized and linearized model*

$$\mathbf{Y} - \mathbf{f}_0 \sim_n (\mathbf{F}\delta\boldsymbol{\beta}, \boldsymbol{\Sigma})$$

the BLUE (best linear unbiased estimator) of $\delta\boldsymbol{\beta}$ *is*

$$\widehat{\delta\boldsymbol{\beta}}_{\text{under}} = (\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0).$$

Its bias in the model

$$\mathbf{Y} - \mathbf{f}_0 \sim_n \left[(\mathbf{F}, \mathbf{S}) \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\boldsymbol{\gamma} \end{pmatrix} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma}), \boldsymbol{\Sigma}\right] \tag{1}$$

is $\mathbf{b}_{\delta\beta} = \mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\delta\boldsymbol{\gamma} + \frac{1}{2}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma})$, *where* $\mathbf{C} = \mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F}$.

Proof is elementary and therefore it is omitted (in more detail cf. [2] and [6]). □

The symbol $\mathbf{I}$ means the identity matrix and $\mathbf{A}^+$ is the Moore–Penrose generalized inverse [7] of the matrix $\mathbf{A}$. The projection matrix on the column space $\mathcal{M}(\mathbf{A}_{m,n}) = \{\mathbf{Au}\colon \mathbf{u} \in R^n\}$ of the matrix $\mathbf{A}$ is denoted as $\mathbf{P}_A$, i.e. $\mathbf{P}_A = \mathbf{AA}^+$ and $\mathbf{M}_A = \mathbf{I} - \mathbf{P}_A$.

In the following text different approaches (cf. also [4] and [5]) are described, which enable us to decide whether the parameter $\delta\boldsymbol{\gamma}$ can or cannot be neglected.

2 The case of a single function

Let a single linear function $\mathbf{h}'\boldsymbol{\beta} = \mathbf{h}'\boldsymbol{\beta}_0 + \mathbf{h}'\delta\boldsymbol{\beta}$, be under consideration in the model (1). Then the following theorem can be stated.

Theorem 2.1 *If* $\begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\boldsymbol{\gamma} \end{pmatrix} \in \mathcal{A}_h$, *then*

$$\mathrm{Var}\left(\mathbf{h}'\widehat{\delta\boldsymbol{\beta}}_{\mathrm{under}}\right) + (\mathbf{h}'\mathbf{b}_{\delta\beta})^2 < \mathrm{Var}\left(\mathbf{h}'\widehat{\delta\boldsymbol{\beta}}_{\mathrm{true}}\right).$$

Here

$$\mathcal{A}_h = \left\{ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} : \mathbf{u} \in R^k, \mathbf{v} \in R^l, \left[\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} + \frac{1}{2}\mathbf{A}_h^+ \begin{pmatrix} \mathbf{0} \\ \mathbf{a}_h \end{pmatrix} \right]' \right.$$

$$\left. \times \mathbf{A}_h \left[\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} + \frac{1}{2}\mathbf{A}_h^+ \begin{pmatrix} \mathbf{0} \\ \mathbf{a}_h \end{pmatrix} \right] + (\mathbf{0}', \mathbf{a}_h')\mathbf{M}_{A_h} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} - \frac{1}{4}(\mathbf{0}', \mathbf{a}_h')\mathbf{A}_h^+ \begin{pmatrix} \mathbf{0} \\ \mathbf{a}_h \end{pmatrix} \le c_h \right\},$$

$$\mathbf{A}_h = \sum_{i=1}^{n} \left\{ \frac{1}{2}\mathbf{h}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1} \right\}_i \mathbf{F}_i, \quad \mathbf{M}_{A_h} = \mathbf{I} - \mathbf{A}_h\mathbf{A}_h^+, \quad \mathbf{a}_h' = \mathbf{h}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S},$$

$$c_h = \sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\left[\mathbf{S}'(\mathbf{M}_F\boldsymbol{\Sigma}\mathbf{M}_F)^+\mathbf{S}\right]^{-1}\mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}\mathbf{h}}$$

and $\widehat{\delta\boldsymbol{\beta}}_{\mathrm{true}}$ *is the BLUE of the vector* $\delta\boldsymbol{\beta}$ *in the model*

$$\mathbf{Y} - \mathbf{f}_0 \sim_n \left[(\mathbf{F}, \mathbf{S}) \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\boldsymbol{\gamma} \end{pmatrix}, \boldsymbol{\Sigma} \right]. \tag{2}$$

Proof Since the BLUE of $\delta\boldsymbol{\beta}$ in the model (2) is

$$\widehat{\delta\boldsymbol{\beta}}_{\mathrm{true}} = \mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0) - \mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\left[\mathbf{S}'(\mathbf{M}_F\boldsymbol{\Sigma}\mathbf{M}_F)^+\mathbf{S}\right]^{-1}$$
$$\times\left[\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0)\right],$$

where $\mathbf{C} = \mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F}$ and the vectors

$$\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0) \quad \text{and} \quad \mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0)$$

are noncorrelated, it is valid that

$$\mathrm{Var}(\widehat{\delta\boldsymbol{\beta}}_{\mathrm{true}}) = \mathbf{C}^{-1} + \mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\left[\mathbf{S}'(\mathbf{M}_F\boldsymbol{\Sigma}\mathbf{M}_F)^+\mathbf{S}\right]^{-1}\mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}.$$

Thus

$$\operatorname{Var}(\mathbf{h}'\widehat{\delta\boldsymbol{\beta}}_{\text{true}}) = \operatorname{Var}(\mathbf{h}'\widehat{\delta\boldsymbol{\beta}}_{\text{under}})$$
$$+\,\mathbf{h}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\big[\mathbf{S}'(\mathbf{M}_F\boldsymbol{\Sigma}\mathbf{M}_F)^{+}\mathbf{S}\big]^{-1}\mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}\mathbf{h}.$$

The bias of the estimator $\mathbf{h}'\widehat{\delta\boldsymbol{\beta}}_{\text{under}}$ in the model (1) is

$$\mathbf{h}'\mathbf{b}_{\delta\beta} = \mathbf{a}_h'\delta\boldsymbol{\gamma} + (\delta\boldsymbol{\beta}', \delta\boldsymbol{\gamma}')\mathbf{A}_h\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}.$$

The linear–quadratic form on the right hand side of the last equality, can be expressed as follows.

In general

$$\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix} \notin \mathcal{M}(\mathbf{A}_h) = \mathcal{M}(\mathbf{A}_h^{+}).$$

Therefore the vector $\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}$ is decomposed, i.e.

$$\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix} = \mathbf{P}_{A_h}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix} + \mathbf{M}_{A_h}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix},$$

where $\mathbf{P}_{A_h}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix} \in \mathcal{M}(\mathbf{A}_h)$. Thus

$$\mathbf{h}'\mathbf{b}_{\delta\beta} = (\mathbf{0}', \mathbf{a}_h')\mathbf{M}_{A_h}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} + (\mathbf{0}', \mathbf{a}_h')\mathbf{P}_{A_h}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} + (\delta\boldsymbol{\beta}', \delta\boldsymbol{\gamma}')\mathbf{A}_h\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}$$

and

$$(\mathbf{0}', \mathbf{a}_h')\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} + (\delta\boldsymbol{\beta}', \delta\boldsymbol{\gamma}')\mathbf{A}_h\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} = \left[\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} + \frac{1}{2}\mathbf{A}_h^{+}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right]'\mathbf{A}_h$$
$$\times\left[\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} + \frac{1}{2}\mathbf{A}_h^{+}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right] + (\mathbf{0}', \mathbf{a}_h')\mathbf{M}_{A_h}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} - \frac{1}{4}(\mathbf{0}', \mathbf{a}_h')\mathbf{A}_h^{+}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix},$$

since

$$\mathbf{A}_h^{+}\mathbf{P}_{A_h}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix} = \mathbf{A}_h^{+}\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix} \quad \text{and} \quad \mathbf{A}_h\mathbf{A}_h^{+} = \mathbf{P}_{A_h}.$$

The assumption $\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} \in \mathcal{A}_h$ implies $|\mathbf{h}'\mathbf{b}_{\delta\beta}| < c_h$ and

$$\operatorname{Var}\left(\mathbf{h}'\widehat{\delta\boldsymbol{\beta}}_{\text{true}}\right) - \operatorname{Var}\left(\mathbf{h}'\widehat{\delta\boldsymbol{\beta}}_{\text{under}}\right) = c_h^2 > (\mathbf{h}'\mathbf{b}_{\delta\beta})^2$$

implies the statement of the theorem. □

Remark 2.2 The semiaxes of the quadratic on the left hand side of the following equality

$$\left[\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}+\frac{1}{2}\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right]'\mathbf{A}_h\left[\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}+\frac{1}{2}\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right]$$
$$= c_h+\frac{1}{4}(\mathbf{0}',\mathbf{a}_h')\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}-(\mathbf{0}',\mathbf{a}_h')\mathbf{M}_{A_h}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}$$

depends on the vector $\mathbf{M}_{A_h}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}$ which is orthogonal to $\mathcal{M}(\mathbf{A}_h)$.

Let

$$\mathbf{u}=\mathbf{P}_{A_h}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix},\ \mathbf{v}=\mathbf{M}_{A_h}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}.$$

Then the boundery of the set $\mathcal{A}_h$ is

$$\left\{\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix}:\ \left[\mathbf{u}+\frac{1}{2}\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right]'\mathbf{A}_h\left[\mathbf{u}+\frac{1}{2}\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right]\right.$$
$$\left.= c_h+\frac{1}{4}(\mathbf{0}',\mathbf{a}_h')\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}-(\mathbf{0}',\mathbf{a}_h')\mathbf{v}\right\},$$

i.e. it is a paraboloid with the section for $\mathbf{v}=\mathbf{0}$ equal to

$$\left\{\mathbf{u}:\ \left[\mathbf{u}+\frac{1}{2}\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right]'\mathbf{A}_h\left[\mathbf{u}+\frac{1}{2}\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right]\right.$$
$$\left.= c_h+\frac{1}{4}(\mathbf{0}',\mathbf{a}_h')\mathbf{A}_h^+\begin{pmatrix}\mathbf{0}\\ \mathbf{a}_h\end{pmatrix}\right\}.$$

Another approach to a problem of a single function is as follows. Since

$$\mathbf{h}'\mathbf{b}_{\delta\beta}=\mathbf{h}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\delta\boldsymbol{\gamma}+(\delta\boldsymbol{\beta}',\delta\boldsymbol{\gamma}')\mathbf{A}_h\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix},$$

two sets, i.e.

$$\mathcal{A}_{1,h}=R^k\times\left\{\delta\boldsymbol{\gamma}\colon |\mathbf{h}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\delta\boldsymbol{\gamma}|\le\frac{\varepsilon}{2}\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}\right\}$$

and

$$\mathcal{A}_{2,h}=\left\{\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}:(\delta\boldsymbol{\beta}',\delta\boldsymbol{\gamma}')\mathbf{A}_h\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}\le\frac{\varepsilon}{2}\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}\right\}$$

can be constructed.

Thus the set

$$\mathcal{C}_h=\mathcal{A}_{1,h}\cap\mathcal{A}_{2,h}$$

is the set with the property

$$\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}\in\mathcal{C}_h\quad\Rightarrow\quad |\mathbf{h}'\mathbf{b}_{\delta\beta}|\le\varepsilon\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}},$$

where $\mathbf{h}$ is given by the function $h(\mathbf{b}) = \mathbf{h}'\mathbf{b}$, $\mathbf{b} \in R^k$, considered. The statement is obvious in the view of the Scheffé inequality [8]

$$\forall\{\mathbf{h} \in R^k\}|\mathbf{h}'\mathbf{b}_{\delta\beta}| \leq \varepsilon\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}} \quad \text{iff} \quad \sqrt{\mathbf{b}'_{\delta\beta}\mathbf{C}\mathbf{b}_{\delta\beta}} \leq \varepsilon.$$

Instead of the matrix $\mathbf{C}^{-1} = \mathrm{Var}(\widehat{\delta\boldsymbol{\beta}}_{\mathrm{under}})$, the matrix

$$\mathbf{C}^{-1} + \mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\left[\mathbf{S}'(\mathbf{M}_F\boldsymbol{\Sigma}\mathbf{M}_F)^{+}\mathbf{S}\right]^{-1}\mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1} = \mathrm{Var}(\widehat{\delta\boldsymbol{\beta}}_{\mathrm{true}})$$

can be used.

3 The case of the whole vector $\delta\boldsymbol{\beta}$

The bias of the estimator $\widehat{\delta\boldsymbol{\beta}}_{\mathrm{under}}$ in the model (1) is composed of the two terms, i.e. $\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\delta\boldsymbol{\gamma}$ which is due to neglecting the parameter $\delta\boldsymbol{\gamma}$ in the model and $\frac{1}{2}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma})$ which is due to the nonlinearity of the model.

The first term is suppressed sufficiently if

$$\delta\boldsymbol{\gamma}'\mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}\mathbf{C}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\delta\boldsymbol{\gamma} \leq \left(\frac{\varepsilon}{2}\right)^2,$$

where $\varepsilon > 0$ is sufficiently small number. In view of the Scheffé inequality it is equivalent to the relation

$$\forall\{\mathbf{h} \in R^k\}|\mathbf{h}'\mathbf{b}_{\delta\beta}| \leq \frac{\varepsilon}{2}\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}.$$

Let

$$\mathcal{A}_1 = R^k \times \left\{\delta\boldsymbol{\gamma}\colon \delta\boldsymbol{\gamma}'\mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}\mathbf{C}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\delta\boldsymbol{\gamma} \leq \left(\frac{\varepsilon}{2}\right)^2\right\}.$$

The second term can be analyzed by a small modification of the Bates and Watts parametric curvature (cf. [1])

$$C^{par}(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0) = \sup\left\{\frac{\sqrt{\frac{1}{4}\boldsymbol{\kappa}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}\mathbf{C}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}}}{(\delta\boldsymbol{\beta}', \delta\boldsymbol{\gamma}')\begin{pmatrix}\mathbf{F}'\\ \mathbf{S}'\end{pmatrix}\boldsymbol{\Sigma}^{-1}(\mathbf{F}, \mathbf{S})\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix}}: \begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} \in R^{k+l}\right\}$$

(a simple algorithm for a numerical determination of $C^{par}(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)$ is given in [1]).

It is valid that (in more detail cf. [3] and [6])

$$\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\boldsymbol{\gamma}\end{pmatrix} \in \mathcal{A}_2$$
$$= \left\{\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix} : (\mathbf{u}', \mathbf{v}')\begin{pmatrix}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F}, & \mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\\ \mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{F}, & \mathbf{S}'\boldsymbol{\Sigma}^{-1}\mathbf{S}\end{pmatrix}\begin{pmatrix}\mathbf{u}\\ \mathbf{v}\end{pmatrix} \leq \frac{\varepsilon/2}{C^{par}(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)}\right\}$$
$$\Rightarrow \quad \forall\{\mathbf{h} \in R^k\}\left|\mathbf{h}'\frac{1}{2}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\boldsymbol{\gamma})\right| \leq \frac{\varepsilon}{2}\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}.$$

Thus the following theorem can be stated.

Theorem 3.1 *The model*

$$\mathbf{Y} - \mathbf{f}_0 \sim_n \left[(\mathbf{F}, \mathbf{S}) \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\boldsymbol{\gamma} \end{pmatrix} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\gamma), \boldsymbol{\Sigma} \right]$$

can be substituted by the model

$$\mathbf{Y} - \mathbf{f}_0 \sim_n (\mathbf{F}\delta\boldsymbol{\beta}, \boldsymbol{\Sigma}) \quad \textit{if} \quad \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\boldsymbol{\gamma} \end{pmatrix} \in \mathcal{A}_1 \cap \mathcal{A}_2.$$

In this case

$$\forall \{\mathbf{h} \in R^k\} |\mathbf{h}'\mathbf{b}_{\delta\beta}| \leq \varepsilon\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}.$$

4 Numerical example

Let

$$y_i = \beta_1 \exp(-\beta_2 x_i) + \gamma x_i + \varepsilon_i,$$

i	1	2	3	4	5	6
x_i	1	2	3	4	5	6

$$\mathbf{Y} = (Y_1, Y_2, \dots, Y_6)', \quad \mathrm{Var}(\mathbf{Y}) = \sigma^2\mathbf{I} = (0.01)^2\mathbf{I}, \quad \begin{pmatrix} \boldsymbol{\beta}_0 \\ \gamma_0 \end{pmatrix} = \begin{pmatrix} 10 \\ 1 \\ 2 \end{pmatrix},$$

$$\mathbf{f}_0 = (5.6788,\ 5.3534,\ 6.4979,\ 8.1832,\ 10.0674,\ 12.0248)',$$

$$\mathbf{F} = \frac{\partial E(\mathbf{Y})}{\partial(\beta_1, \beta_2)} = \begin{pmatrix} \exp(-\beta_2 x_1),\ -x_1\beta_1\exp(-\beta_2 x_1) \\ \dots\dots\dots\dots\dots\dots \\ \exp(-\beta_2 x_6),\ -x_6\beta_1\exp(-\beta_2 x_6) \end{pmatrix}$$

$$= \begin{pmatrix} 0.3679,\ -3.6788 \\ 0.1353,\ -2.7067 \\ 0.0498,\ -1.4936 \\ 0.0183,\ -0.7326 \\ 0.0067,\ -0.3369 \\ 0.0025,\ -0.1487 \end{pmatrix},$$

$$\mathbf{S} = (1, 2, 3, 4, 5, 6)',$$

$$\mathbf{F}_i = \frac{\partial^2 E(\mathbf{Y})}{\partial \begin{pmatrix} \boldsymbol{\beta} \\ \gamma \end{pmatrix} \partial(\boldsymbol{\beta}', \gamma)} = \begin{pmatrix} 0, & -x_i\exp(-\beta_2 x_i), & 0 \\ -x_i\exp(-\beta_2 x_i), & x_i^2\beta_1\exp(-\beta_2 x_i), & 0 \\ 0, & 0, & 0 \end{pmatrix},$$

$$i = 1, 2, \dots, 6,$$

$$\mathbf{F}_1 = \begin{pmatrix} 0.000, & -0.368, & 0.000 \\ -0.368, & 3.679, & 0.000 \\ 0.000, & 0.000, & 0.000 \end{pmatrix}, \ \mathbf{F}_2 = \begin{pmatrix} 0.000, & -0.271, & 0.000 \\ -0.271, & 5.413, & 0.000 \\ 0.000, & 0.000, & 0.000 \end{pmatrix},$$

$$\mathbf{F}_3 = \begin{pmatrix} 0.000, & -0.149, & 0.000 \\ -0.149, & 4.481, & 0.000 \\ 0.000, & 0.000, & 0.000 \end{pmatrix}, \ \mathbf{F}_4 = \begin{pmatrix} 0.000, & -0.073, & 0.000 \\ -0.073, & 2.931, & 0.000 \\ 0.000, & 0.000, & 0.000 \end{pmatrix},$$

$$\mathbf{F}_5 = \begin{pmatrix} 0.000, & -0.034, & 0.000 \\ -0.034, & 1.684, & 0.000 \\ 0.000, & 0.000, & 0.000 \end{pmatrix}, \ \mathbf{F}_6 = \begin{pmatrix} 0.000, & -0.015, & 0.000 \\ -0.015, & -0.892, & 0.000 \\ 0.000, & 0.000, & 0.000 \end{pmatrix},$$

$$\mathbf{C} = \frac{1}{\sigma^2}\mathbf{F}'\mathbf{F} = \frac{1}{0.01^2}\begin{pmatrix} 0.157, & -1.810 \\ -1.810, & 23.763 \end{pmatrix}.$$

(i) The case of a single function $\mathbf{h} = (1,0)'$

$$\mathbf{A}_h = \sum_{i=1}^{6}\left\{\frac{1}{2}(1,0)(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\right\}_i \mathbf{F}_i = \begin{pmatrix} 0.000, & 0.000, & 0.000 \\ 0.000, & -12.584, & 0.000 \\ 0.000, & 0.000, & 0.000 \end{pmatrix},$$

$$a_h = (1,0)(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{S} = -29.169,$$

$$c_h = \sigma\sqrt{(1,0)(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{S}(\mathbf{S}'\mathbf{M}_F\mathbf{S})^{-1}\mathbf{S}'\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}} = 0.037713,$$

$$\frac{1}{2}\mathbf{A}_h^+\begin{pmatrix} \mathbf{0} \\ a_h \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{M}_{A_h} = \begin{pmatrix} 1, & 0, & 0 \\ 0, & 0, & 0 \\ 0, & 0, & 1 \end{pmatrix},$$

$$\frac{1}{4}(\mathbf{0}', a_h)\mathbf{A}_h^+\begin{pmatrix} \mathbf{0} \\ a_h \end{pmatrix} = 0,$$

$$\mathcal{A}_h = \left\{\begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} : \delta\boldsymbol{\beta} \in R^2, \delta\gamma \in R^1, \left[\begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} + \frac{1}{2}\mathbf{A}_h^+\begin{pmatrix} \mathbf{0} \\ a_h \end{pmatrix}\right]'\mathbf{A}_h \right.$$
$$\times\left[\begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} + \frac{1}{2}\mathbf{A}_h^+\begin{pmatrix} \mathbf{0} \\ a_h \end{pmatrix}\right]$$
$$\left. \leq c_h + \frac{1}{4}(\mathbf{0}', a_h)\mathbf{A}_h^+\begin{pmatrix} \mathbf{0} \\ a_h \end{pmatrix} - (\mathbf{0}', a_h)\mathbf{M}_{A_h}\begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix}\right\}$$
$$= \left\{\begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} : (\delta\boldsymbol{\beta}', \delta\gamma)\begin{pmatrix} 0, & 0.000, & 0.000 \\ 0.000, & -12.584, & 0 \\ 0.000, & 0.000, & 0.000 \end{pmatrix}\begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} \leq 0.0377 + 29.2\delta\gamma\right\},$$

cf. also Fig. 1

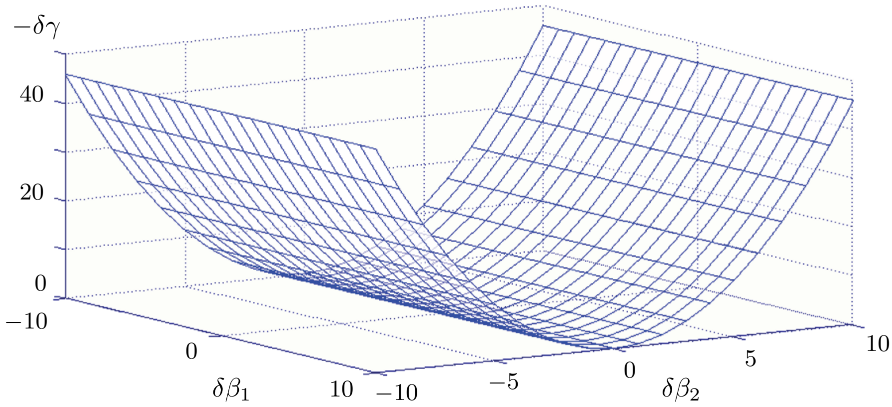

Fig. 1 The set $\mathcal{A}_h$

Another approach is characterized by the sets $\mathcal{A}_{1,h}$ and $\mathcal{A}_{2,h}$ (for $\varepsilon = 0.04$ and $\sigma = 0.01$).

$$\mathcal{A}_{1,h} = R^2 \times \left\{\delta\gamma\colon |(1,0)(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{S}\delta\gamma| \leq \frac{\varepsilon}{2}\sigma\sqrt{\{(\mathbf{F}'\mathbf{F})^{-1}\}_{1,1}}\right\}$$
$$= R^2 \times \{\delta\gamma\colon |\delta\gamma| \leq 5.012 \times 10^{-5}\},$$

$$\mathcal{A}_{2,h} = \left\{\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\gamma\end{pmatrix}\colon (\delta\boldsymbol{\beta}', \delta\gamma)\mathbf{A}_h\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\gamma\end{pmatrix} \leq \frac{\varepsilon}{2}\sigma\sqrt{\{(\mathbf{F}'\mathbf{F})^{-1}\}_{1,1}}\right\}$$

$$= \left\{\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\gamma\end{pmatrix}\colon (\delta\boldsymbol{\beta}', \delta\gamma)\begin{pmatrix}0.000, & 0.000, & 0.000\\ 0.000, & -12.584, & 0.000\\ 0.000, & 0.000, & 0.000\end{pmatrix}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\gamma\end{pmatrix} \leq 0.001446\right\},$$

$$\mathcal{C}_h = \mathcal{A}_{1,h} \cap \mathcal{A}_{2,h}.$$

(ii) The case of the whole vector

$$C^{par}(\boldsymbol{\beta}_0, \gamma_0) =$$

$$\sup\left\{\sigma\frac{\sqrt{\frac{1}{4}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta}, \delta\gamma)\mathbf{P}_F\boldsymbol{\kappa}(\delta\boldsymbol{\beta}, \delta\gamma)}}{(\delta\boldsymbol{\beta}', \delta\gamma)\begin{pmatrix}\mathbf{F}'\mathbf{F}, & \mathbf{F}'\mathbf{S}\\ \mathbf{S}'\mathbf{F}, & \mathbf{S}'\mathbf{S}\end{pmatrix}\begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\gamma\end{pmatrix}}\colon \begin{pmatrix}\delta\boldsymbol{\beta}\\ \delta\gamma\end{pmatrix} \in R^3\right\} = 0.012652,$$

$$\mathcal{A}_1 = R^2 \times \left\{\delta\gamma\colon |\delta\gamma| \leq \frac{\varepsilon\sigma}{\sqrt{4\mathbf{S}'\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{S}}}\right\}$$
$$= R^2 \times \left\{\delta\gamma\colon |\delta\gamma| \leq 3.5818 \times 10^{-5}\right\},$$

$$\mathcal{A}_2 = \left\{ \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} : (\delta\boldsymbol{\beta}', \delta\gamma) \begin{pmatrix} \mathbf{F}'\mathbf{F}, & \mathbf{F}'\mathbf{S} \\ \mathbf{S}'\mathbf{F}, & \mathbf{S}'\mathbf{S} \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} \leq \frac{\sigma^2\varepsilon}{2C^{par}(\boldsymbol{\beta}_0, \gamma_0)} \right\}$$

$$= \left\{ \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} : (\delta\boldsymbol{\beta}', \delta\gamma) \begin{pmatrix} 0.157, & -1.810, & 0.910 \\ -1.810, & 23.763, & -19.080 \\ 0.910, & -19.080, & 91.000 \end{pmatrix} \begin{pmatrix} \delta\boldsymbol{\beta} \\ \delta\gamma \end{pmatrix} \leq 4.16524 \times 10^{-5} \right\}$$

(for $\varepsilon = 0.04$ and $\sigma = 0.01$).

The set $\mathcal{A}_{1,h}$ has the same shape as the set $\mathcal{A}_1$ in Fig. 2b and 2c however it is wider $5.012/3.5818 = 1.399$ times.

The section of the set $\mathcal{A}_1 \cap \mathcal{A}_2$ cf. on Fig. 2.

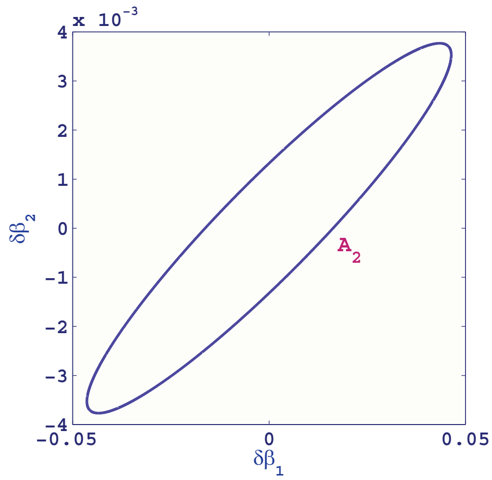

Fig. 2a The section of the set $\mathcal{A}_2$ in coordinates $\delta\beta_1$ and $\delta\beta_2$

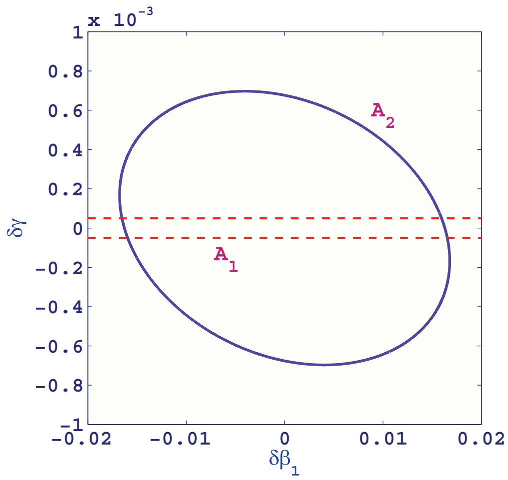

Fig. 2b The section of the set $\mathcal{A}_1 \cap \mathcal{A}_2$ in coordinates $\delta\beta_1$ and $\delta\gamma$

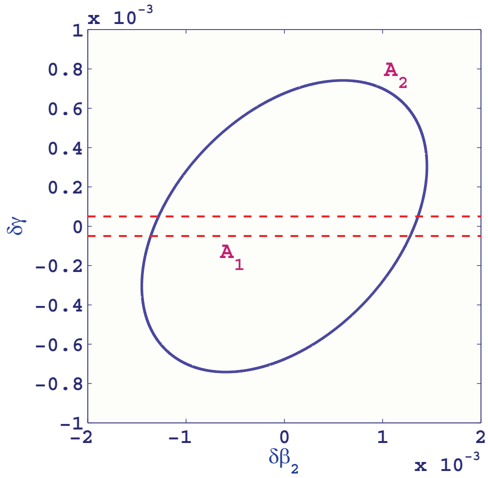

Fig. 2c The section of the set $\mathcal{A}_1 \cap \mathcal{A}_2$ in coordinates $\delta\beta_2$ and $\delta\gamma$

The criterior for neglecting the parameter $\boldsymbol{\gamma}$ is rigorous. If the function $\mathbf{f}(\beta_1, \beta_2, \boldsymbol{\gamma}) = \beta_1$ is estimated, the value $|\beta_2|$ must be smaller than 0.010 and $|\gamma|$ must be smaller than 0.0000501, i.e. in fact they cannot be neglected.

If the whole vector is estimated, the situation is even worse.

The model considered cannot be reduced.

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*, Bratislava, 1995.

[3] Kubáček, L., Kubáčková, L.: Regression Models with a weak Nonlinearity. Technical Report Nr. 1998.1, *Universität Stuttgart*, Stuttgart, 1998, 1–67.

[4] Kubáček, L.: *Underparametrization in a regression model with constraints II.* Math. Slovaca **55** (2005), 579–596.

[5] Kubáček, L.: *Underparametrized regression models.* Tatra Mt. Math. Publ. **14** (2006), 241–254.

[6] Kubáček, L., Tesaříková, E.: Weakly Nonlinear Regression Models. *Vyd. Univerzity Palackého*, Olomouc, 2008.

[7] Rao, C. R., Mitra, S. K.: Generalized Inverse of Matrices and its Applications. *Wiley*, New York–London–Sydney–Toronto, 1971.

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