VERSITA

Mathematica Slovaca

DOI: 10.2478/s12175-009-0146-z
Math. Slovaca **59** (2009), No. 5, 535–544

A MEASURE OF NONLINEARITY IN SINGULAR NONLINEAR REGRESSION MODELS

Lubomír Kubáček

Dedicated to Professor Andrej Pázman on the occasion of his 70th birthday

(*Communicated by Gejza Wimmer*)

ABSTRACT. The method used in nonlinear regression is most of the time the nonlinear squares. If the nonlinearity of the model is not too strong, it is possible to linearize the model and to use the linear statistical theory which is more simple. The question is under which condition this can be done. If the linearized model is regular such criteria were found. To find a solution for singular models is the aim of the paper.

Introduction

The nonlinear least squares method, mostly used in nonlinear regression, is described and analyzed in [7]. It is more sophisticated than the linear least squares method, more complicated numerically and it needs different algorithms for solution of statistical problems. Nevertheless sometimes the situation requires such methods and no other way for a solution is at our disposal.

If a nonlinearity of the model is not too strong, it is possible to linearize the model and to use the more simple linear statistical theory. The question is under what condition this can be done.

If the linearized model is regular, i.e. the design matrix (the matrix of the first derivatives of the mean value of the observation vector with respect to

2000 Mathematics Subject Classification: Primary 62F10, 62J05.
Keywords: weak nonlinearity, singularity, BLUE.
Supported by the Ministry of Education, Youth and Sports of the Czech Republic under research project MSM 6 198 959 214.

the parameters) is of the full column rank and the covariance matrix of the observation vector is positive definite, the criteria were found which enable us to decide whether the nonlinear model can be linearized. The mentioned criteria are inspired by the Bates and Watts curvatures [1] and they are different for different statistical problems (in more detail cf. [4], [5], [6], etc.).

The linearization is admissible if a solution of a statistical problem considered in the linearized model is deteriorated only nonsignificantly by the nonlinearity. It means that in the case of estimation the bias of the estimator caused by nonlinearity is negligible with respect to the standard deviation of the estimator. In the case of confidence region the decrease of the confidence level must be smaller than in advance prescribed small value $\varepsilon > 0$, etc.

A criterion for parameter estimation is based on the Bates and Watts parametric curvature, which in this case is the measure of nonlinearity. This measure of nonlinearity enables us to construct a neighbourhood, called the linearization region (the form of it is an ellipsoid of the same shape as the confidence ellipsoid however of the different size), of the value of the parameter which is used for a linearization of the model. The linearization region is defined by its following property. If the actual value of the parameter is inside the linearization region, then the bias of the best linear estimator of any linear function of parameters caused by nonlinearity is smaller than $100\varepsilon\%$ of the standard deviation of the estimator. Here $\varepsilon > 0$ is chosen in advance by a statistician. The criterion is satisfied if the actual value of the parameter is with sufficiently high probability in the linearization region.

If the linearized model is singular, i.e. the design matrix is not of the full column rank, the covariance matrix need not be positive definite, then the measure of nonlinearity cannot be derived from the Bates and Watts curvature, since it need not exist. However some generalization inspired by the Bates and Watts curvatures can be used for a construction of the linearization region.

The aim of the paper is to find such linearization region in a singular model.

1. Notation and auxiliary statements

Let

$$\mathbf{Y} \sim N_n[\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Sigma}], \qquad \boldsymbol{\beta} \in \mathbb{R}^k,$$

where $\mathbb{R}^k$ is k-dimensional real linear vector space, $\mathbf{Y}$ is an n-dimensional normally distributed random vector (observation vector) with the mean value $E(\mathbf{Y})$

equal to $\mathbf{f}(\boldsymbol{\beta})$ and the known covariance matrix $\mathrm{Var}(\mathbf{Y}) = \boldsymbol{\Sigma}$. The function $\mathbf{f}(\cdot) : \mathbb{R}^k \to \mathbb{R}^n$ is given and it is assumed that it can be expressed in the Taylor series of the second order, i.e.

$$\begin{aligned} \mathbf{f}(\boldsymbol{\beta}) &= \mathbf{f}_0 + \mathbf{F}\boldsymbol{\delta} + \frac{1}{2}\boldsymbol{\kappa}(\boldsymbol{\delta}), \qquad \boldsymbol{\delta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0, \quad \mathbf{f}_0 = \mathbf{f}(\boldsymbol{\beta}_0), \\ \mathbf{F} &= \frac{\partial \mathbf{f}(\mathbf{u})}{\partial \mathbf{u}'}\Big|_{u=\beta_0}, \qquad \boldsymbol{\kappa}(\boldsymbol{\delta}) = \big(\kappa_1(\boldsymbol{\delta}), \dots, \kappa_n(\boldsymbol{\delta})\big)', \\ \kappa_i(\boldsymbol{\delta}) &= \boldsymbol{\delta}' \frac{\partial^2 f_i(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}\Big|_{u=\beta_0} \boldsymbol{\delta}, \qquad i = 1, \dots, n. \end{aligned}$$

In fact it means that quadratic models are considered only. The vector $\boldsymbol{\beta}_0$ is an approximate value of the actual parameter $\boldsymbol{\beta}$. The k-dimensional parameter $\boldsymbol{\beta}$ is unknown and some linear functions of it must be estimated on the basis of a realization of the observation vector $\mathbf{Y}$. The matrix $\mathbf{F}$ need not be of the full column rank and the matrix $\boldsymbol{\Sigma}$ need not be positive definite.

If the model

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n(\mathbf{F}\boldsymbol{\delta}, \boldsymbol{\Sigma}), \qquad \boldsymbol{\delta} \in \mathbb{R}^k, \tag{1}$$

is considered instead of the quadratic model, the following statements are valid. They are given without proofs, since they are well known (for more detail cf. [2]).

Lemma 1. *The function $h(\boldsymbol{\beta}) = \mathbf{h}'(\boldsymbol{\beta}_0 + \boldsymbol{\delta})$, $\boldsymbol{\delta} \in \mathbb{R}^k$, is linearly unbiasedly estimable iff $\mathbf{h} \in \mathscr{M}(\mathbf{F}') = \Big\{\mathbf{F}'\mathbf{u} : \ \mathbf{u} \in \mathbb{R}^n\Big\}$.*

Lemma 2. *The BLUE (best linear unbiased estimator) of the unbiasedly estimable linear function $\mathbf{h}'\boldsymbol{\beta}$, $\boldsymbol{\beta} \in \mathbb{R}^k$, is*

$$\widehat{\mathbf{h}'\boldsymbol{\beta}} = \mathbf{h}'\boldsymbol{\beta}_0 + \mathbf{h}'\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'(\mathbf{Y} - \mathbf{f}_0),$$

where $(\mathbf{F}')^-_{m(\Sigma)}$ is the minimum $\boldsymbol{\Sigma}$-seminorm g-inverse of the matrix $\mathbf{F}'$ (in more detail cf. [9]).

Thus the class of all unbiasedly estimable linear functions can be characterized by the vector $\mathbf{P}_{F'}\boldsymbol{\beta}_0 + \mathbf{P}_{F'}\boldsymbol{\delta}$, where $\mathbf{P}_{F'} = \mathbf{F}'(\mathbf{F}\mathbf{F}')^-\mathbf{F}$ is the projection matrix on the subspace $\mathscr{M}(\mathbf{F}') = \Big\{\mathbf{F}'\mathbf{u} : \ \mathbf{u} \in \mathbb{R}^n\Big\}$ in the Euclidean norm.

Lemma 3. *The BLUE of the vector $\mathbf{P}_{F'}\boldsymbol{\delta}$ is*

$$\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}} = \mathbf{P}_{F'}\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'(\mathbf{Y} - \mathbf{f}_0),$$

and its covariance matrix is

$$\begin{aligned} \mathrm{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}) &= \mathbf{P}_{F'}\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'\boldsymbol{\Sigma}(\mathbf{F}')^-_{m(\Sigma)}\mathbf{P}_{F'} \\ &= \mathbf{P}_{F'}\big[(\mathbf{F}'\mathbf{T}^-\mathbf{F})^- - \mathbf{I}\big]\mathbf{P}_{F'}, \end{aligned}$$

where $\mathbf{T} = \mathbf{\Sigma} + \mathbf{F}\mathbf{F}'$.

2. Linearization region for the bias

If the model

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

is considered instead of of the model (1), then the term $\frac{1}{2}\boldsymbol{\kappa}(\boldsymbol{\delta})$ causes a bias in the BLUE from Lemma 3.

Lemma 4. *The bias* $\mathbf{b}$ *of the estimator* $\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}$ *from* Lemma 3 *is*

$$\mathbf{b} = \frac{1}{2}\mathbf{P}_{F'}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\boldsymbol{\kappa}(\boldsymbol{\delta})$$

in the model (2).

P r o o f. Proof is obvious. □

If $\boldsymbol{\delta}$ is sufficiently small, then $\mathbf{b}$ is small as well. However the size of $\mathbf{b}$ must be compared with a statistical uncertainty of the estimator $\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}$. Here the problem due to the statement of the following lemma occurs.

Lemma 5. *It is valid that*

$$\mathscr{M}\left[\operatorname{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\right] \subset \mathscr{M}(\mathbf{F}'), \qquad r\left[\operatorname{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\right] = r(\mathbf{F}'\mathbf{T}^-\mathbf{\Sigma}).$$

P r o o f. The first statement is obvious. As far as the following statement is concerned, the relationship

$$\operatorname{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}) = \mathbf{F}'(\mathbf{F}\mathbf{F}')^-\mathbf{F}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\mathbf{\Sigma}(\mathbf{F}')^-_{m(\Sigma)}\mathbf{F}'(\mathbf{F}\mathbf{F}')^-\mathbf{F},$$

implies that

$$\begin{aligned} r\left[\operatorname{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\right] &= r\left\{\mathbf{F}'(\mathbf{F}\mathbf{F}')^-\mathbf{F}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\mathbf{\Sigma}(\mathbf{F}')^-_{m(\Sigma)}\mathbf{F}'(\mathbf{F}\mathbf{F}')^-\mathbf{F}\right\} \\ &= r\left\{\mathbf{F}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\mathbf{\Sigma}(\mathbf{F}')^-_{m(\Sigma)}\mathbf{F}'\right\} = r\left\{\mathbf{F}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\mathbf{\Sigma}\right\} \\ &\geq r\left\{\left[\operatorname{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\right]\right\}. \end{aligned}$$

Here the equality

$$\mathbf{F}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\mathbf{\Sigma}(\mathbf{F}')^-_{m(\Sigma)}\mathbf{F}' = \mathbf{F}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\mathbf{\Sigma}$$

was utilized (in more detail cf. [9]). The expression $\mathbf{F}\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'\boldsymbol{\Sigma}$ is invariant with respect to the choice of the g-inverse $(\mathbf{F}')^-_{m(\Sigma)}$ and thus

$$
\begin{aligned}
&r\Big\{\mathbf{F}\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'\boldsymbol{\Sigma}\Big\} = \\
&\quad = r\big[\mathbf{F}(\mathbf{F}'\mathbf{T}^-\mathbf{F})^-\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma}\big] \geq r\big[\mathbf{F}'\mathbf{T}^-\mathbf{F}(\mathbf{F}'\mathbf{T}^-\mathbf{F})^-\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma}\big] \\
&\quad = r(\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma}) \geq r\big[\mathbf{F}(\mathbf{F}'\mathbf{T}^-\mathbf{F})^-\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma}\big] = r\Big\{\mathbf{F}\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'\boldsymbol{\Sigma}\Big\}.
\end{aligned}
$$

□

The bias $\mathbf{b}$ need not be in $\mathscr{M}\big[\mathrm{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\big]$ and thus the expression

$$\mathbf{b}'\big[\mathrm{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\big]^-\mathbf{b},$$

which could be used for a definition of the measure of nonlinearity, may have no meaning. If $\mathbf{b} \in \mathscr{M}\big[\mathrm{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\big]$, then the inequality

$$\mathbf{b}'\,\mathrm{Var}\big[\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\big]\mathbf{b} \leq \varepsilon^2$$

is equivalent with the inequalities

$$\forall\{\mathbf{h} \in \mathbb{R}^k\}|\mathbf{h}'\mathbf{b}| \leq \varepsilon\sqrt{\mathbf{h}'\,\mathrm{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})\mathbf{h}}$$

(regarding the Scheffé theorem). The last inequalities can serve as a good criterion for the admissibility of the bias. In the general case when $\mathbf{b}$ need not be in $\mathscr{M}[\mathrm{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})]$ some modification is necessary.

Let $\mathrm{Var}(\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}) = \mathbf{U}$.

Lemma 6. *The* $(1-\alpha)$*-confidence region* $\mathscr{E}(\mathbf{P}_{F'}\boldsymbol{\delta})$ *for* $\mathbf{P}_{F'}\boldsymbol{\delta}$

$$
\begin{aligned}
\mathscr{E}(\mathbf{P}_{F'}\boldsymbol{\delta}) \quad = \quad \Big\{\mathbf{u}:\ (\mathbf{u}-\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}) \in \mathscr{M}(\mathbf{U}),\ (\mathbf{u}-\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})'\mathbf{U}^+(\mathbf{u}-\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}) \\
\leq \chi^2_{r(F'T^-\Sigma)}(0;1-\alpha)\Big\}
\end{aligned}
$$

is for any $c^2 > 0$ *(how to choose the value* c^2 *is mentioned in* Remark 7*) subset of*

$$
\begin{aligned}
&\overline{\mathscr{E}}(\mathbf{P}_{F'}\boldsymbol{\delta}) = \\
&\quad = \Big\{\mathbf{u}:\ \mathbf{u} \in \mathscr{M}(\mathbf{F}'),\ (\mathbf{u}-\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}})'\Big[\mathbf{U}^+ + \frac{1}{c^2}(\mathbf{P}_{F'}-\mathbf{P}_U)\Big](\mathbf{u}-\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}) \\
&\qquad\qquad \leq \chi^2_{r(F'T^-\Sigma)}(0;1-\alpha)\Big\}.
\end{aligned}
$$

P r o o f. The expression for $\mathscr{E}(\mathbf{P}_{F'}\boldsymbol{\delta})$ is implied by the Pearson lemma (cf. [8]) and Lemma 5, where degrees of freedom are given. The inclusion $\mathscr{E}(\mathbf{P}_{F'}\boldsymbol{\delta}) \subset \overline{\mathscr{E}}(\mathbf{P}_{F'}\boldsymbol{\delta})$ is a consequence of the fact that

$$\mathbf{u} \in \mathscr{M}(\mathbf{U}) \implies \mathbf{u}'\Big[\mathbf{U}^+ + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{u} = \mathbf{u}'\mathbf{U}^+\mathbf{u}.$$

□

Remark 7. The length of semi-axes in the subspace $\mathscr{M}(\mathbf{P}_{F'} - \mathbf{P}_U)$ of the ellipsoid $\overline{\mathscr{E}}(\mathbf{P}_{F'}\boldsymbol{\delta})$ is $\sqrt{\chi^2_{r(F'T^-\Sigma)}(0; 1-\alpha)c^2}$ and it is the same for all $r(\mathbf{F}') - r(\mathbf{U})$ semi-axes. Thus the length of the projected vector $(\mathbf{P}_{F'} - \mathbf{P}_U)\mathbf{b} \perp \mathscr{M}(\mathbf{U})$ can be judged by the inequality

$$\begin{aligned} &\mathbf{b}'(\mathbf{P}_{F'} - \mathbf{P}_U)\Big[\mathbf{U}^+ + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big](\mathbf{P}_{F'} - \mathbf{P}_U)\mathbf{b} \\ &= \frac{1}{c^2}\mathbf{b}'(\mathbf{P}_{F'} - \mathbf{P}_U)\mathbf{b} \le \varepsilon. \end{aligned}$$

If the spectral decomposition of the matrix $\mathbf{U}$ is $\mathbf{U} = \sum_{i=1}^{r(F'T^-\Sigma)} \lambda_i \mathbf{f}_i \mathbf{f}_i'$, then the lengths of the semi-axes in the subspace $\mathscr{M}(\mathbf{U})$ of the ellipsoid $\overline{\mathscr{E}}(\mathbf{P}_{F'}\boldsymbol{\delta})$ are

$$\sqrt{\chi^2_{r(F'T^-\Sigma)}(0; 1-\alpha)\lambda_i}, \qquad i = 1, \ldots, r(\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma}).$$

Now the number c^2 can be chosen, e.g. as

$$c^2 = \max\{\lambda_i : \ i = 1, \ldots, r(\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma})\}.$$

(Other choice regarding the viewpoint of the user can be made as well.) Then, even if $\mathbf{b} \notin \mathscr{M}(\mathbf{U})$, the inequality

$$\mathbf{b}'\Big[\mathbf{U}^+ + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{b} \le \varepsilon$$

can serve as a good starting point for a determination of a measure of nonlinearity and a linearization region.

Definition 8. The measure of nonlinearity for the bias of the estimator $\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}$ is

$$C_{P_F'\delta} = \sup\left\{\frac{\sqrt{R(\boldsymbol{\delta})}}{\boldsymbol{\delta}'\mathbf{P}_{F'}\Big[\mathbf{U}^+ + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{P}_{F'}\boldsymbol{\delta}} : \ \boldsymbol{\delta} \in \mathbb{R}^k\right\},$$

where

$$R(\boldsymbol{\delta}) \ = \ \boldsymbol{\kappa}'(\boldsymbol{\delta})(\mathbf{F}')^-_{m(\Sigma)}\mathbf{P}_{F'}\Big[\mathbf{U}^+ + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{P}_{F'}\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'\boldsymbol{\kappa}(\boldsymbol{\delta}).$$

Theorem 9. *If* $\boldsymbol{\delta} \in \mathscr{L}_{P_{F'}\delta}$ *(the linearization region for the bias of the estimator* $\widehat{\mathbf{P}_{F'}\boldsymbol{\delta}}$*),*

$$\mathscr{L}_{P_{F'}\delta\beta} \quad = \quad \left\{\mathbf{P}_{F'}\mathbf{u}: \ \mathbf{u}'\mathbf{P}_{F'}\Big[\mathbf{U}^{+} + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{P}_{F'}\mathbf{u} \leq \frac{2\varepsilon}{C_{P_{F'}\delta}}\right\},$$

then

$$\forall\{\mathbf{h} \in \mathbb{R}^k\}|\mathbf{h}'\mathbf{b}| \leq \varepsilon\sqrt{\mathbf{h}'\Big[\mathbf{U} + c^2(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{h}}.$$

Here the value $\varepsilon > 0$ *is chosen with respect to the opinion of a statistician.*

Proof. Regarding Definition 8 we have

$$\sqrt{R(\boldsymbol{\delta})} \leq \boldsymbol{\delta}'\mathbf{P}_{F'}\Big[\mathbf{U}^{+} + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{P}_{F'}\boldsymbol{\delta}C_{P'_F\delta}.$$

If

$$\boldsymbol{\delta}'\mathbf{P}_{F'}\Big[\mathbf{U}^{+} + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{P}_{F'}\boldsymbol{\delta} \leq \frac{2\varepsilon}{C_{P'_F\delta}},$$

then

$$2\sqrt{\mathbf{b}'\Big[\mathbf{U}^{+} + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)]\Big]\mathbf{b}} \leq 2\varepsilon,$$

i.e.

$$\mathbf{b}'\Big[\mathbf{U}^{+} + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{b} \leq \varepsilon^2.$$

Now the Scheffé theorem [10]

$$\forall\{\mathbf{h} \in \mathbb{R}^k\}|\mathbf{h}'\mathbf{b}| \leq \varepsilon\sqrt{\mathbf{h}'\Big[\mathbf{U} + c^2(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{h}}$$
$$\Longleftrightarrow \ \mathbf{b}'\Big[\mathbf{U}^{+} + \tfrac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{b} \leq \varepsilon^2$$

can be used. The last equivalence is valid in this case, since

$$\mathbf{b} \in \mathscr{M}\Big[\mathbf{U} + c^2(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]$$

and further the equality

$$\Big[\mathbf{U} + c^2(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]^{+} = \mathbf{U}^{+} + \frac{1}{c^2}(\mathbf{P}_{F'} - \mathbf{P}_U)$$

is valid. □

Remark 10. If $\mathbf{h} \in \mathscr{M}(\mathbf{U})$, then

$$\mathbf{h}'\Big[\mathbf{U} + c^2(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{h} = \mathrm{Var}(\widehat{\mathbf{h}'\mathbf{P}_{F'}\boldsymbol{\delta}}),$$

i.e. in this case $|\mathbf{h}'\mathbf{b}| \leq \varepsilon\sqrt{\mathrm{Var}(\widehat{\mathbf{h}'\mathbf{P}_{F'}\boldsymbol{\delta}})}$.

If $\mathbf{h} \in \mathscr{M}(\mathbf{P}_{F'} - \mathbf{P}_U)$, then $|\mathbf{h}'\mathbf{b}| \leq \varepsilon c\sqrt{\mathbf{h}'\mathbf{h}}$ (in this case $\mathrm{Var}(\widehat{\mathbf{h}'\mathbf{P}_{F'}\delta\boldsymbol{\beta}}) = 0$).

Remark 11. Let's investigate the result of Theorem 9 when the the linearized model is regular, i.e. if $r(\mathbf{F}) = k \leq n$ and $\mathbf{\Sigma}$ is positive definite. In this case the Bates and Watts parametric curvature coincides with the measure of nonlinearity defined in Definition 8. Particulary

$$\begin{aligned}
\mathbf{U} &= \mathbf{P}_{F'}(\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{P}_{F'} = (\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F})^{-1},\\
r(\mathbf{U}) &= k, \quad \mathscr{M}(\mathbf{U}) = \mathscr{M}(\mathbf{F}') = \mathbb{R}^k, \quad \mathbf{P}_{F'} - \mathbf{P}_U = \mathbf{0},\\
C_{P_{F'}\delta} &= \sup\left\{ \frac{\sqrt{R(\boldsymbol{\delta})}}{\boldsymbol{\delta}'\mathbf{F}'(\mathbf{F}\mathbf{F}')^{-}\mathbf{F}\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F}\mathbf{F}'(\mathbf{F}\mathbf{F}')^{-}\mathbf{F}\boldsymbol{\delta}} : \ \boldsymbol{\delta} \in \mathbb{R}^k \right\},
\end{aligned}$$

where

$$\begin{aligned}
R(\boldsymbol{\delta}) &= \boldsymbol{\kappa}'(\boldsymbol{\delta})\mathbf{\Sigma}^{-1}\mathbf{F}(\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'(\mathbf{F}\mathbf{F}')^{-}\mathbf{F}(\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'(\mathbf{F}\mathbf{F}')^{-}\mathbf{F}\\
&\quad \times(\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\boldsymbol{\kappa}(\boldsymbol{\delta})\\
&= \boldsymbol{\kappa}'(\boldsymbol{\delta})\mathbf{\Sigma}^{-1}\mathbf{F}(\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\boldsymbol{\kappa}(\boldsymbol{\delta}) = \boldsymbol{\kappa}'(\boldsymbol{\delta})\mathbf{\Sigma}^{-1}\mathbf{P}_{F'}^{\Sigma^{-1}}\boldsymbol{\kappa}(\boldsymbol{\delta})
\end{aligned}$$

and

$$\boldsymbol{\delta}'\mathbf{F}'(\mathbf{F}\mathbf{F}')^{-}\mathbf{F}\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F}\mathbf{F}'(\mathbf{F}\mathbf{F}')^{-}\mathbf{F}\boldsymbol{\delta} = \boldsymbol{\delta}'\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F}\boldsymbol{\delta}.$$

The Bates and Watts parametric curvature $K^{(par)}(\boldsymbol{\beta}_0)$ of the model (2) at the point $\boldsymbol{\beta}_0$ is defined as [1]

$$K^{(par)}(\boldsymbol{\beta}_0) = \sup\left\{ \frac{\sqrt{\boldsymbol{\kappa}'(\boldsymbol{\delta})\mathbf{\Sigma}^{-1}\mathbf{P}_{F'}^{\Sigma^{-1}}\boldsymbol{\kappa}(\boldsymbol{\delta})}}{\boldsymbol{\delta}'\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F}\boldsymbol{\delta}} : \ \boldsymbol{\delta} \in \mathbb{R}^k \right\}.$$

Thus in a regular model $C_{P_{F'}\delta} = K^{(par)}(\boldsymbol{\beta}_0)$. The linearization region $\mathscr{L}$ constructed in a regular model on the basis $K^{(par)}(\boldsymbol{\beta}_0)$ is

$$\mathscr{L} = \left\{ \boldsymbol{\delta} : \ \boldsymbol{\delta}'\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F}\boldsymbol{\delta} \leq \frac{2\varepsilon}{K^{(par)}(\boldsymbol{\beta}_0)} \right\}$$

(in more detail cf. [4]) and thus

$$\begin{aligned}
\mathscr{L}_{P_{F'}\delta} &= \left\{ \mathbf{P}_{F'}\mathbf{u} : \ \mathbf{u}'\mathbf{P}_{F'}\Big[\mathbf{U}^{+} + c^2(\mathbf{P}_{F'} - \mathbf{P}_U)\Big]\mathbf{P}_{F'}\mathbf{u} \leq \frac{2\varepsilon}{C_{P_{F'}\delta}} \right\}\\
&= \left\{ \mathbf{P}_{F'}\boldsymbol{\delta} = \boldsymbol{\delta} : \ \boldsymbol{\delta}'\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F}\boldsymbol{\delta} \leq \frac{2\varepsilon}{K^{(par)}(\boldsymbol{\beta}_0)} \right\} = \mathscr{L}.
\end{aligned}$$

Theorem 9 is a generalization of a theorem valid for a regular model.

Remark 12. The linearization region $\mathscr{L}_{P_{F'}\delta}$ is of practical use only, when we are sure (or practically sure) that the actual value of the vector $\mathbf{P}_{F'}\boldsymbol{\delta}$ is an element of $\mathscr{L}_{P_{F'}\delta}$. Our information on the actual position of the vector $\mathbf{P}_{F'}\boldsymbol{\delta}$ is given by the relation

$$P\Big\{\mathscr{E}(\mathbf{P}_{F'}\boldsymbol{\delta}) \ni \mathbf{P}_{F'}\boldsymbol{\delta}\Big\} = 1 - \alpha$$

(Lemma 6). Since $\mathscr{E}(\mathbf{P}_{F'}\boldsymbol{\delta}) \subset \overline{\mathscr{E}}$ and $\overline{\mathscr{E}}$ is of the same shape as the linearization region $\mathscr{L}_{P_{F'}\delta}$, the necessary condition for $\mathscr{E}(\mathbf{P}_{F'}\boldsymbol{\delta}) \subset \mathscr{L}_{P_{F'}\delta}$ is

$$\chi^2_{r(F'T^-\Sigma)}(0; 1-\alpha) \ll \frac{2\varepsilon}{C_{P_{F'}\delta}},$$

which is the good first information on a possibility to consider the model (1) instead of (2).

Acknowledgement. Author wuld like to express his thank to reviewers for their thorough reading the manuscript and for their advice how to make the formulation of the text better.

REFERENCES

[1] BATES, D. M.—WATTS, D. G.: *Relative curvature measures of nonlinearity*, J. Roy. Statist. Soc. Ser. B **42** (1980), 1–25.

[2] FIŠEROVÁ, E.—KUBÁČEK, L.—KUNDEROVÁ, P.: *Linear Statistical Models, Regularity and Singularities*, Academia, Praha, 2007.

[3] KUBÁČEK, L.—KUBÁČKOVÁ, L.—VOLAUFOVÁ, J.: *Statistical Models with Linear Structures*, Veda, Bratislava, 1995.

[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*, Publishing House of Palacký University, Olomouc, 2000 (Czech).

[6] KUBÁČEK, L.—TESAŘÍKOVÁ, E.: *Weakly nonlinear regression model with constraints I: nonlinear hypothesis*, Discuss. Math. Probab. Stat. **25**, 2005, 115–133.

[7] PÁZMAN, A.: *Nonlinear Statistical Models*, Kluwer Academic Publisher/Ister Science Press, Dordrecht-Boston-London/Bratislava, 1993.

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

[9] RAO, C. R.—MITRA, S. K.: *Generalized Inverse of Matrices and Its Applications*, John Wiley&Sons, New York–London–Sydney–Toronto, 1971.

[10] SCHEFFÉ, H.: *The Analysis of Variance* (5th ed.), J. Wiley. New York-London-Sydney, 1967.

Received 10. 6. 2008

Department of Mathematical Analysis
and Applied Mathematics
Faculty of Natural Science
Palacký University
CZ–779 00 Tomkova 40
Olomouc
CZECH REPUBLIC
E-mail: kubacekl@inf.upol.cz