Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **48** (2009) 73–82

Linearization Regions for a Confidence Ellipsoid in Singular Nonlinear Regression Models*

Lubomír KUBÁČEK [1], Eva TESAŘÍKOVÁ [2]

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

(Received March 24, 2009)

Abstract

A construction of confidence regions in nonlinear regression models is difficult mainly in the case that the dimension of an estimated vector parameter is large. A singularity is also a problem. Therefore some simple approximation of an exact confidence region is welcome. The aim of the paper is to give a small modification of a confidence ellipsoid constructed in a linearized model which is sufficient under some conditions for an approximation of the exact confidence region.

Key words: Nonlinear regression model, confidence region, singularity.

2000 Mathematics Subject Classification: 62F10, 62J05

1 Introduction

A construction of a confidence region for unbiasedly estimable functions of nonlinear singular regression model parameters can be a difficult numerical problem (for more detail on nonlinear models cf. [6]). Mainly the case of a large dimension of a vector parameter is unwelcome. If a confidence region can be

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

approximated by a confidence ellipsoid (in the case of normally distributed observation vector), then a numerical calculation and an interpretation of results are much more easier and simpler.

Therefore an attempt to find a simple measure of nonlinearity which enable us to decide whether an approximate confidence ellipsoid can be used instead of exact confidence region, is the aim of the paper.

2 Notation and some useful statements

The following notation is used.

$$\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Sigma}) \tag{1}$$

means that $\mathbf{Y}$ is an n-dimensional normally distributed random vector with the mean value $E(\mathbf{Y})$ equal to $\mathbf{f}(\boldsymbol{\beta})$ and with the covariance matrix $\mathrm{Var}(\mathbf{Y}) = \boldsymbol{\Sigma}$. Let the function $\mathbf{f}(\cdot)\colon R^k \to R^n$ (R^n is the n-dimensional real linear vector space) can be expressed by the Taylor series of the second order, i.e.

$$\begin{aligned}
\mathbf{f}(\boldsymbol{\beta}) &= \mathbf{f}_0 + \mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}), \quad \delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0,\\
\mathbf{f}_0 &= \mathbf{f}(\boldsymbol{\beta}_0), \quad \boldsymbol{\beta}_0 \text{ is an approximate value of } \boldsymbol{\beta},\\
\mathbf{F} &= \frac{\partial \mathbf{f}(\mathbf{u})}{\partial \mathbf{u}}\Big|_{u=\beta_0}, \quad \boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \left[\kappa_1(\delta\boldsymbol{\beta}), \dots, \kappa_n(\delta\boldsymbol{\beta})\right]',\\
\kappa_i(\delta\boldsymbol{\beta}) &= \delta\boldsymbol{\beta}' \frac{\partial^2 f_i(\mathbf{u})}{\partial \mathbf{u}\partial \mathbf{u}'}\Big|_{u=\beta_0} \delta\boldsymbol{\beta}, \quad i = 1, \dots, n.
\end{aligned}$$

The matrix $\mathbf{F}$ need not be of the full rank in columns and $\boldsymbol{\Sigma}$ need not be positive definite.

The linearized version of the model (1) is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n(\mathbf{F}\delta\boldsymbol{\beta}, \boldsymbol{\Sigma}) \tag{2}$$

and the quadratized version is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\left(\mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}), \boldsymbol{\Sigma}\right). \tag{3}$$

In the following text the notations

$\mathbf{A}^-$ … g-inverse (generalized inverse) of the matrix $\mathbf{A}$,

$\mathbf{A}^+$ … the Moore–Penrose g-inverse of the matrix $\mathbf{A}$,

$\mathbf{A}^-_{m(W)}$ … minimum $\mathbf{W}$-seminorm g-inverse of the matrix $\mathbf{A}$, ($\mathbf{W}$ is positive semidefinite matrix),

$\mathcal{M}(\mathbf{A}_{m,n}) = \{\mathbf{A}\mathbf{u}\colon \mathbf{u} \in R^n\}$ (column space of the matrix) $\mathbf{A}$,

$\mathbf{I}$ … identity matrix,

$\mathbf{P}_{F'} = \mathbf{F}'(\mathbf{F}\mathbf{F}')^{-}\mathbf{F}$ the projection matrix on the space $\mathcal{M}(\mathbf{F}')$ in the Euclidean norm,

$r(\mathbf{A})$ ... the rank of the matrix $\mathbf{A}$,

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

$\mathbf{T} = \boldsymbol{\Sigma} + \mathbf{F}\mathbf{F}'$,

will be used. More on a g-inverse of a matrix cf. [7].

In the model (2) a representative of all unbiasedly estimable linear functions of the parameter $\boldsymbol{\beta}$ is the vector

$$\boldsymbol{\gamma} = \mathbf{P}_{F'}\boldsymbol{\beta} = \mathbf{P}_{F'}\boldsymbol{\beta}_0 + \mathbf{P}_{F'}\delta\boldsymbol{\beta} = \boldsymbol{\gamma}_0 + \delta\boldsymbol{\gamma}.$$

Lemma 1 *In the model (2) the* $(1-\alpha)$*-confidence ellipsoid of the vector* $\mathbf{P}_{F'}\delta\boldsymbol{\beta}$ *is*

$$\begin{aligned}\mathcal{E}_{P_{F'}\delta\beta} = \Big\{&\mathbf{P}_{F'}\mathbf{u}\colon \mathbf{P}_{F'}\mathbf{u} - \widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}} \in \mathcal{M}\big[\mathrm{Var}(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})\big], (\mathbf{P}_{F'}\mathbf{u} - \widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})' \\ &\times \big[\mathrm{Var}(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})\big]^{-}(\mathbf{P}_{F'}\mathbf{u} - \widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}}) \leq \chi^2_{r[F'(\Sigma+\mathbf{F}\mathbf{F}')^{-}\Sigma]}(0;1-\alpha)\Big\},\end{aligned}$$

where

$$\begin{aligned}\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}} &= \mathbf{P}_{F'}\big[(\mathbf{F}')^{-}_{m(\Sigma)}\big]'(\mathbf{Y} - \mathbf{f}_0), \\ \mathrm{Var}(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}}) &= \mathbf{P}_{F'}\big[(\mathbf{F}'\mathbf{T}^{-}\mathbf{F})^{-} - \mathbf{I}\big]\mathbf{P}_{F'}, \quad \mathbf{T} = \boldsymbol{\Sigma} + \mathbf{F}\mathbf{F}'.\end{aligned}$$

Proof is given in [2]. □

In the following text it is necessary to take into account the fact that even $\boldsymbol{\beta}_0$ can be considered to be known, only $\mathbf{P}_{F'}(\boldsymbol{\beta} - \boldsymbol{\beta}_0) = \mathbf{P}_{F'}\delta\boldsymbol{\beta}$ can be unbiasedly estimated. Let

$$\boldsymbol{\beta}_0 = \boldsymbol{\gamma}_0 + \boldsymbol{\omega}_0, \quad \boldsymbol{\gamma}_0 = \mathbf{P}_{F'}\boldsymbol{\beta}_0, \quad \boldsymbol{\omega}_0 = \mathbf{M}_{F'}\boldsymbol{\beta}_0;$$

the parameter $\delta\boldsymbol{\gamma} = \mathbf{P}_{F'}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)$ is unbiasedly estimable in the model (2), however $\delta\boldsymbol{\omega} = \mathbf{M}_{F'}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)$ is not. Therefore the model

$$\mathbf{Y} \sim N_n\left[\mathbf{f}(\boldsymbol{\beta}_0) + \mathbf{F}\delta\boldsymbol{\gamma} + \frac{1}{2}\boldsymbol{\kappa}_{\omega_0}(\delta\boldsymbol{\gamma}), \boldsymbol{\Sigma}\right] \tag{4}$$

will be considered instead the model (3). Here

$$\begin{aligned}\boldsymbol{\kappa}_{\omega_0} &= (\kappa_{\omega_0,1}, \ldots, \kappa_{\omega_0,n})', \\ \kappa_{\omega_0,i} &= \delta\boldsymbol{\gamma}'\frac{\partial^2 f_i(\boldsymbol{\gamma}_0 + \boldsymbol{\omega}_0)}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}\delta\boldsymbol{\gamma}, \quad i = 1, \ldots, n, \\ \mathbf{F} &= \frac{\partial\mathbf{f}(\boldsymbol{\gamma}_0 + \boldsymbol{\omega}_0)}{\partial\boldsymbol{\gamma}'}.\end{aligned}$$

Lemma 2 *The bias* $\mathbf{b}$ *of the estimator*

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

in the model (4) is

$$\begin{aligned}\mathbf{b} \;&=\; E(\widehat{\delta\boldsymbol{\gamma}}) - \delta\boldsymbol{\gamma} = \frac{1}{2}\mathbf{P}_{F'}\big[(\mathbf{F}')^-_{m(\Sigma)}\big]'\boldsymbol{\kappa}_{\omega_0}(\delta\boldsymbol{\gamma})\\ &=\; \frac{1}{2}\mathbf{P}_{F'}(\mathbf{F}'\mathbf{T}^-\mathbf{F})^-\mathbf{F}'\mathbf{T}^-\boldsymbol{\kappa}_{\omega_0}(\delta\boldsymbol{\gamma}).\end{aligned}$$

Proof is implied by the definition of the bias. □

Lemma 3 *Let* $\mathbf{Y} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Then*

$$\mathbf{Y}'\boldsymbol{\Sigma}^+\mathbf{Y} \sim \chi^2_{r(\Sigma)}(\delta),$$

where $\delta = \boldsymbol{\mu}'\boldsymbol{\Sigma}^+\boldsymbol{\mu} = \boldsymbol{\mu}'\mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma\boldsymbol{\mu}$.

Proof Let $\mathbf{J}$ be a $k \times r(\boldsymbol{\Sigma})$ matrix such that $\mathbf{J}\mathbf{J}' = \boldsymbol{\Sigma}$ and $\mathbf{K}$ such a $k \times r(\boldsymbol{\Sigma})$ matrix that $\mathbf{K}\mathbf{K}' = \boldsymbol{\Sigma}^+$ (i.e. $\mathbf{J}'\mathbf{K} = \mathbf{I}$). Then $\mathbf{K}'\mathbf{Y} = \mathbf{K}'\boldsymbol{\mu} + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim N_{r(\Sigma)}(\mathbf{0}, \mathbf{I})$. Thus

$$\mathbf{Y}'\mathbf{K}\mathbf{K}'\mathbf{Y} \;=\; \mathbf{Y}'\boldsymbol{\Sigma}^+\mathbf{Y} = \boldsymbol{\eta}'\boldsymbol{\eta} + 2\boldsymbol{\eta}'\mathbf{K}'\boldsymbol{\mu} + \boldsymbol{\mu}'\boldsymbol{\Sigma}^+\boldsymbol{\mu} \sim \chi^2_{r(\Sigma)}(\boldsymbol{\mu}'\boldsymbol{\Sigma}^+\boldsymbol{\mu}).$$

However $\boldsymbol{\Sigma}^+ = \mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma$, since

$$\begin{gathered}\boldsymbol{\Sigma}\mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma\boldsymbol{\Sigma} = \boldsymbol{\Sigma}, \quad \mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma\boldsymbol{\Sigma}\mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma = \mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma,\\ \boldsymbol{\Sigma}\mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma = \mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma\boldsymbol{\Sigma} = \mathbf{P}_\Sigma, \quad \mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{P}_\Sigma\boldsymbol{\Sigma}^-\mathbf{P}_\Sigma = \mathbf{P}_\Sigma.\end{gathered}$$

(in more detail cf. [7]). □

3 A linearization region for a confidence ellipsoid

Since $r\big[\operatorname{Var}(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})\big] = r[\mathbf{F}'(\boldsymbol{\Sigma} + \mathbf{F}\mathbf{F}')^-\boldsymbol{\Sigma}]$, it can happen that

$$r\big[\operatorname{Var}(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})\big] = r\big[\operatorname{Var}(\widehat{\delta\boldsymbol{\gamma}})\big] < r(\mathbf{F}').$$

Therefore the vector $\mathbf{b}$ need not be an element of $\mathcal{M}\big[\operatorname{Var}(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})\big]$.

The relation

$$\delta\boldsymbol{\gamma} = \mathbf{P}_{F'}\delta\boldsymbol{\beta} = E(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}}) - \mathbf{b} = E(\widehat{\delta\boldsymbol{\gamma}}) - \mathbf{b},$$

valid in the model (3) and (4), respectively, implies that in general case the vector $\mathbf{P}_{F'}\delta\boldsymbol{\beta}$ need not be an element of $\mathcal{E}_{P_{F'}\delta\beta}$ from Lemma 1. Thus it seems to be reasonable to enlarge the ellipsoid $\mathcal{E}_{P_{F'}\delta\beta}$ to $\overline{\mathcal{E}}$ in such a way that $\mathbf{P}_{F'}\delta\boldsymbol{\beta} \in \overline{\mathcal{E}}$ with sufficiently high probability.

In the following text the notation $\mathbf{U} = \operatorname{Var}(\widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})$ will be used.

Definition 1 Let a set $\overline{\mathcal{E}}$ be defined as

$$\overline{\mathcal{E}} = \Big\{\mathbf{P}_{F'}\mathbf{u}\colon \mathbf{u} \in R^k, (\mathbf{P}_{F'}\mathbf{u} - \widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}})'\big[\mathbf{U} + c^2(\mathbf{P}_{F'} - \mathbf{P}_U)\big]^+ \times (\mathbf{P}_{F'}\mathbf{u} - \widehat{\mathbf{P}_{F'}\delta\boldsymbol{\beta}}) \leq \chi^2_{r(F'T-\Sigma)}(0; 1-\alpha)\Big\},$$

where $\mathbf{T} = \boldsymbol{\Sigma} + \mathbf{F}\mathbf{F}'$ and the choice c^2 depends on the opinion of the user (cf. the following remark).

Remark 1 The number c^2 should be comparable with the spectral numbers of the matrix $\mathbf{U}$. The semiaxes of $\overline{\mathcal{E}}$ in the space $\mathcal{M}(\mathbf{P}_{F'} - \mathbf{P}_U)$ have the same size equal to

$$a = c\sqrt{\chi^2_{r(F'T-\Sigma)}(0; 1-\alpha)}.$$

The smaller is c, the smaller is the probability $P\{\mathbf{P}_{F'}\delta\boldsymbol{\beta} \in \overline{\mathcal{E}}\}$. Thus c cannot be smaller than some reasonable bound. If $\mathbf{b} \in \mathcal{M}(\mathbf{U})$, then it can be tolerated in the case $\mathbf{b}'\mathbf{U}^-\mathbf{b} \leq \varepsilon$. Let

$$\mathbf{U} = \sum_{i=1}^{f} \lambda_i \mathbf{f}_i \mathbf{f}_i', \quad f = r(\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma}),$$

be spectral decomposition of the matrix $\mathbf{U}$ and

$$\lambda_{\max} = \max\{\lambda_i\colon i = 1, \dots, r(\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma})\}.$$

If $\mathbf{h} = s\mathbf{f}_{\max}$ (the vector $\mathbf{f}_{\max}$ corresponds to $\lambda_{\max}$), then, regarding the Scheffé theorem [8] $(\mathbf{b}'\mathbf{U}^-\mathbf{b} \leq \varepsilon \Leftrightarrow \forall\{\mathbf{h} \in \mathcal{M}(\mathbf{U})\}|\mathbf{h}'\mathbf{b}| \leq \varepsilon\sqrt{\mathbf{h}'\mathbf{U}\mathbf{h}})$,

$$|\mathbf{h}'\mathbf{b}| = s|\mathbf{f}'_{\max}\mathbf{b}| \leq s\varepsilon\sqrt{\lambda_{\max}}.$$

In the worst case (i.e. $\mathbf{b} = t\mathbf{f}_{\max}$) $\|\mathbf{b}\| = t < \varepsilon\sqrt{\lambda_{\max}}$. It implies that the bias $\mathbf{b}$ with the norm smaller than $\varepsilon\sqrt{\lambda_{\max}}$ can be tolerated and thus the choice $c^2 = \lambda_{\max}$ is reasonable.

Definition 2 Let the measure of nonlinearity for the confidence ellipsoid be

$$C^{(ell)} = \sup\left\{\frac{2\sqrt{\mathbf{b}'(\delta\boldsymbol{\gamma})\big[\mathbf{U} + \lambda_{\max}(\mathbf{P}_{F'} - \mathbf{P}_U)\big]^+\mathbf{b}(\delta\boldsymbol{\gamma})}}{\delta\boldsymbol{\gamma}'\big[\mathbf{U} + \lambda_{\max}(\mathbf{P}_{F'} - \mathbf{P}_U)\big]^+\delta\boldsymbol{\gamma}}\colon \delta\boldsymbol{\gamma} \in R^{r(F)}\right\},$$

where

$$\mathbf{b}(\delta\boldsymbol{\gamma}) = \frac{1}{2}\mathbf{P}_{F'}(\mathbf{F}'\mathbf{T}^-\mathbf{F})^-\mathbf{F}'\mathbf{T}^-\boldsymbol{\kappa}(\delta\boldsymbol{\gamma}).$$

Theorem 1 *If* $\delta\boldsymbol{\beta} \in \mathcal{L}^{(ell)}_{\delta\gamma}$, *where*

$$\mathcal{L}^{(ell)}_{\delta\gamma} = \left\{\delta\boldsymbol{\gamma}\colon \delta\boldsymbol{\gamma} \in \mathcal{M}(\mathbf{F}'), \delta\boldsymbol{\gamma}'\big[\mathbf{U} + \lambda_{\max}(\mathbf{P}_{F'} - \mathbf{P}_U)\big]^-\delta\boldsymbol{\gamma} \leq \frac{2\sqrt{\delta_{\max}}}{C^{(ell)}}\right\},$$

then

$$P\Big\{\delta\boldsymbol{\gamma} \in \overline{\mathcal{E}}\Big\} \geq 1-\alpha-\varepsilon.$$

Here $\delta_{\max}$ *is a solution of the equation*

$$P\Big\{\chi^2_f(\delta_{\max}) \leq \chi^2_f(0;1-\alpha)\Big\} = 1-\alpha-\varepsilon$$

and $f = r(\mathbf{F}'\mathbf{T}^-\boldsymbol{\Sigma})$.

Proof Regarding Definition 6

$$2\sqrt{\mathbf{b}'(\delta\boldsymbol{\gamma})\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+\mathbf{b}(\delta\boldsymbol{\gamma})} \leq \delta\boldsymbol{\gamma}'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^-\delta\boldsymbol{\gamma}C^{(ell)}.$$

Let

$$\delta\boldsymbol{\gamma}'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^-\delta\boldsymbol{\gamma} \leq \frac{2\sqrt{\delta_{\max}}}{C^{(ell)}}.$$

Further

$$\begin{aligned}
&(\widehat{\delta\boldsymbol{\gamma}}-\delta\boldsymbol{\gamma})'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+(\widehat{\delta\boldsymbol{\gamma}}-\delta\boldsymbol{\gamma}) = \\
&\quad= \big[\widehat{\delta\boldsymbol{\gamma}}-E(\widehat{\delta\boldsymbol{\gamma}})+E(\widehat{\delta\boldsymbol{\gamma}})-\delta\boldsymbol{\gamma}\big]'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+ \\
&\quad\quad\times\big[\widehat{\delta\boldsymbol{\gamma}}-E(\widehat{\delta\boldsymbol{\gamma}})+E(\widehat{\delta\boldsymbol{\gamma}})-\delta\boldsymbol{\gamma}\big] \\
&\quad= \big[\widehat{\delta\boldsymbol{\gamma}}-E(\widehat{\delta\boldsymbol{\gamma}})\big]'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+\big[\widehat{\delta\boldsymbol{\gamma}}-E(\widehat{\delta\boldsymbol{\gamma}})\big] \\
&\quad\quad+2\mathbf{b}'(\delta\boldsymbol{\gamma})\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+\big[\widehat{\delta\boldsymbol{\gamma}}-E(\widehat{\delta\boldsymbol{\gamma}})\big] \\
&\quad\quad+\mathbf{b}'(\delta\boldsymbol{\gamma})\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+\mathbf{b}(\delta\boldsymbol{\gamma}) = \chi^2_f(\delta),
\end{aligned}$$

where

$$\delta = \mathbf{b}'(\delta\boldsymbol{\gamma})\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+\mathbf{b}(\delta\boldsymbol{\gamma}),$$

what is implied by Lemma 3. The relation

$$\begin{aligned}
&\big[(\mathbf{Y}-\boldsymbol{\mu})+\boldsymbol{\mu}\big]'\boldsymbol{\Sigma}^+\big[(\mathbf{Y}-\boldsymbol{\mu})+\boldsymbol{\mu}\big] = \\
&\quad= (\mathbf{Y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^-(\mathbf{Y}-\boldsymbol{\mu})+2\boldsymbol{\mu}'\boldsymbol{\Sigma}^+(\mathbf{Y}-\boldsymbol{\mu})+\boldsymbol{\mu}'\boldsymbol{\Sigma}^+\boldsymbol{\mu} = \chi^2_{r(\Sigma)}(\boldsymbol{\mu}'\boldsymbol{\Sigma}^+\boldsymbol{\mu}),
\end{aligned}$$

based on Lemma 3 is used as well.

Thus

$$(\widehat{\delta\boldsymbol{\gamma}}-\delta\boldsymbol{\gamma})'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+(\widehat{\delta\boldsymbol{\gamma}}-\delta\boldsymbol{\gamma}) = \chi^2_f(\delta),$$

where

$$\delta = \mathbf{b}'(\delta\boldsymbol{\gamma})\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^+\mathbf{b}(\delta\boldsymbol{\gamma}).$$

If $\delta \leq \delta_{\max}$, then

$$P\Big\{\chi^2_f(\delta) \leq \chi^2_f(0;1-\alpha)\Big\} \geq P\Big\{\chi^2_f(\delta_{\max}) \leq \chi^2_f(0;1-\alpha)\Big\} = 1-\alpha-\varepsilon,$$

what means $P\Big\{\delta\gamma \in \overline{\mathcal{E}}\Big\} \geq 1-\alpha-\varepsilon$. □

Remark 2 Let us apply Theorem 1 on the regular linearized model. Then $\mathbf{P}_{F'} = \mathbf{P}_U = \mathbf{I}$, $\overline{\mathcal{E}} = \mathcal{E}_{\delta\gamma}$ and $C^{(ell)} = K^{(par)}$, where $K^{(par)}$ is the Bates and Watts parametric curvature

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

(in more detail cf. [1]). In this case the statement

$$\delta\boldsymbol{\beta} \in \left\{\mathbf{u}\colon \mathbf{u}'\mathbf{F}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{u} \le \frac{2\sqrt{\delta_{\max}}}{K^{(par)}}\right\} \Rightarrow P\left\{\mathbf{P}_{F'}\delta\boldsymbol{\beta} \in \mathcal{E}_{P_{F'}\delta\beta}\right\}$$
$$= P\{\delta\boldsymbol{\beta} \in \mathcal{E}_{\delta\beta}\} \ge 1-\alpha-\varepsilon$$

is true (cf. also [4]). Thus Theorem 7 is a reasonable generalization suitable for the singular model.

Remark 3 In the case that only one function of the parameter $\boldsymbol{\beta}$, i.e. $h(\boldsymbol{\gamma}) = \mathbf{h}'\boldsymbol{\gamma}_0 + \mathbf{h}'\delta\boldsymbol{\gamma}$, $\delta\boldsymbol{\gamma} \in \mathcal{M}(\mathbf{F}')$, is important, a very simple procedure can be used. Let in the first case $\mathbf{h}'\mathbf{P}_{F'}\left[(\mathbf{F}'\mathbf{T}^-\mathbf{F})^- - \mathbf{I}\right]\mathbf{P}_{F'}\mathbf{h} > 0$.

Since

$$b_h = E(\widehat{\mathbf{h}'\delta\boldsymbol{\gamma}}) - \mathbf{h}'\delta\boldsymbol{\gamma} = \frac{1}{2}\mathbf{h}'\mathbf{P}_{F'}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\boldsymbol{\kappa}_{\omega_0}(\delta\boldsymbol{\gamma}) = \delta\boldsymbol{\gamma}'\mathbf{A}_h\delta\boldsymbol{\gamma},$$

where

$$\mathbf{A}_h = \sum_{i=1}^n \left\{\frac{1}{2}\mathbf{h}'\mathbf{P}_{F'}\left[(\mathbf{F}')^-_{m(\Sigma)}\right]'\right\}_i \frac{\partial^2 f_i(\mathbf{u}+\boldsymbol{\omega}_0)}{\partial\mathbf{u}\partial\mathbf{u}'}\Big|_{u=\gamma_0},$$

we obtain

$$\delta\boldsymbol{\gamma} \in \mathcal{L}_{h'\delta\gamma} = \left\{\mathbf{u}\colon \mathbf{u} \in \mathcal{M}(\mathbf{F}'), |\mathbf{u}'\mathbf{A}_{h'\delta\beta}\mathbf{u}| \le \sqrt{\delta_{1,\max}}\right\}$$
$$\Rightarrow P\left\{|\mathbf{h}'\delta\boldsymbol{\gamma} - \widehat{\delta\boldsymbol{\gamma}}| \le \sqrt{\chi_1^2(0;1-\alpha)}\sqrt{\mathbf{h}'\mathbf{P}_{F'}\mathbf{U}\mathbf{P}_{F'}\mathbf{h}}\right\} \ge 1-\alpha-\varepsilon.$$

Here $\delta_{1,\max}$ is a solution of the equation

$$P\left\{\chi_1^2(\delta_{1,\max}) \le \chi_1^2(0;1-\alpha)\right\} = 1-\alpha-\varepsilon.$$

If $\mathbf{h}'\mathbf{U}\mathbf{h} = 0$, then

$$P\left\{\mathbf{h}'\widehat{\delta\boldsymbol{\gamma}} - E(\mathbf{h}'\widehat{\delta\boldsymbol{\gamma}}) = 0\right\} = 1$$

and thus

$$P\left\{\mathbf{h}'\widehat{\delta\boldsymbol{\gamma}} = \mathbf{h}'\delta\boldsymbol{\gamma} + \mathbf{h}'\mathbf{b}(\delta\boldsymbol{\gamma})\right\} = 1.$$

Thus

$$\delta\boldsymbol{\gamma} \in \mathcal{L}_{h'\delta\gamma} = \left\{\mathbf{u}\colon \mathbf{u} \in \mathcal{M}(\mathbf{F}'), |\mathbf{u}'\mathbf{A}_h\mathbf{u}| \le \Delta\right\}$$
$$\Rightarrow P\left\{\mathbf{h}'\delta\boldsymbol{\gamma} \in \left\{u\colon u \in R^1, |u - \widehat{\mathbf{h}'\delta\boldsymbol{\gamma}}| \le \Delta\right\}\right\} = 1.$$

It is interesting to compare the linearization regions $\mathcal{L}_{\delta\gamma}$ and $\mathcal{L}_{h'\delta\gamma}$.

4 Numerical example

Let us consider the regression model

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \end{pmatrix} \sim N_6 \left[\begin{pmatrix} \beta_1 \exp(-\beta_3) \\ \beta_1 \exp(-\beta_3) \\ \beta_1 \exp(-\beta_3) \\ \beta_2 \exp(-\beta_3) \\ \beta_2 \exp(-\beta_3) \\ \beta_2 \exp(-\beta_3) \end{pmatrix}, \mathbf{\Sigma}_{6,6} \right], \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \in R^3,$$

$$\mathbf{\Sigma}_{6,6} = \sigma^2 \mathbf{I}_{6,6}, \quad \sigma^2 = (0.5)^2.$$

Then

$$\mathbf{F} = \frac{\partial \mathbf{f}(\mathbf{u} + \boldsymbol{\omega}_0)}{\partial \mathbf{u}'}\Big|_{u=\gamma_0} = \begin{pmatrix} \mathbf{1}_3, & \mathbf{0}, & -\mathbf{1}_3 \\ 0, & \mathbf{1}_3, & -\mathbf{1}_3 \end{pmatrix}, \quad \mathbf{1}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$\mathbf{F}_1 = \mathbf{F}_2 = \mathbf{F}_3 = \begin{pmatrix} 0, & 0, & -1 \\ 0, & 0, & 0 \\ -1, & 0, & 1 \end{pmatrix}, \quad \mathbf{F}_4 = \mathbf{F}_5 = \mathbf{F}_6 = \begin{pmatrix} 0, & 0, & 0 \\ 0, & 0, & -1 \\ 0, & -1, & 1 \end{pmatrix}.$$

Here

$$\mathbf{F}_i = \frac{\partial^2 f_i(\mathbf{u} + \boldsymbol{\omega}_0)}{\partial \mathbf{u} \partial \mathbf{u}'}\Big|_{u=\gamma_0}, \quad i = 1, \dots, 6,$$

$$\mathbf{P}_{F'} = \mathbf{F}'(\mathbf{F}\mathbf{F}')^- \mathbf{F} = \frac{1}{3} \begin{pmatrix} 2, & -1, & -1 \\ -1, & 2, & -1 \\ -1, & -1 & 2 \end{pmatrix},$$

$$\mathrm{Var}(\widehat{\mathbf{P}_{F'} \delta \boldsymbol{\beta}}) = \mathbf{U} = \mathbf{P}_{F'} \Big\{ \big[\mathbf{F}'(\mathbf{\Sigma} + \mathbf{F}\mathbf{F}')^- \mathbf{F} \big]^- - \mathbf{I} \Big\} \mathbf{P}_{F'} = \frac{\sigma^2}{54} \begin{pmatrix} 10, & -8, & -2 \\ -8, & 10, & -2 \\ -2, & -2, & 4 \end{pmatrix},$$

$$\mathbf{P}_U = \mathbf{U}(\mathbf{U}^2)^- \mathbf{U} = \frac{1}{3} \begin{pmatrix} 2, & -1, & -1 \\ -1, & 2, & -1 \\ -1, & -1 & 2 \end{pmatrix},$$

$$\mathbf{U} = \sum_{i=1}^{r[\mathbf{F}'(\Sigma + \mathbf{F}\mathbf{F}')^- \Sigma]} \lambda_i \mathbf{f}_i \mathbf{f}_i' = \sum_{i=1}^{2} \lambda_i \mathbf{f}_i \mathbf{f}_i', \quad \lambda_1 = \frac{1}{3}\sigma^2, \ \lambda_2 = \frac{1}{9}\sigma^2, \ \lambda_{\max} = \frac{1}{3}\sigma^2,$$

$\delta_{\max} = 0.48$ is a solution of the equation

$$P\left\{ \chi_f^2(0; 1-\alpha) \right\} = 1 - \alpha - \varepsilon,$$

and $f = r\big[\mathbf{F}'(\mathbf{\Sigma} + \mathbf{F}\mathbf{F}')^- \mathbf{\Sigma}\big] = 2$, $\alpha = 0.05$, $\varepsilon = 0.04$.

Further

$$C^{(ell)} = \sup\left\{ \frac{2\sqrt{\mathbf{b}'(\delta\boldsymbol{\gamma})\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^{+}\mathbf{b}(\delta\boldsymbol{\gamma})}}{\delta\boldsymbol{\gamma}'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^{+}\delta\boldsymbol{\gamma}} : \delta\boldsymbol{\gamma}\in R^2 \right\}$$
$$= \sigma\cdot 0.191273,$$

where

$$\mathbf{b} = \frac{1}{2}\mathbf{P}_{F'}(\mathbf{F}'\mathbf{T}^{-}\mathbf{F})^{-}\mathbf{F}'\mathbf{T}^{-}\boldsymbol{\kappa}_{\omega_0}(\delta\boldsymbol{\gamma}).$$

The linearization region for $\delta\boldsymbol{\gamma} = \mathbf{P}_{F'}\delta\boldsymbol{\beta}$ is

$$\mathcal{L}_{\delta\gamma} = \left\{ \mathbf{u}\colon \mathbf{u}\in\mathcal{M}(\mathbf{F}'), \mathbf{u}'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^{+}\mathbf{u} \le \frac{2\sqrt{\delta_{\max}}}{C^{(ell)}} \right\}$$

and the set $\overline{\mathcal{E}_{\delta\gamma}}$ is

$$\overline{\mathcal{E}_{\delta\gamma}} = \Big\{ \mathbf{u}\colon \mathbf{u}\in\mathcal{M}(\mathbf{F}'), (\mathbf{u}-\widehat{\delta\boldsymbol{\gamma}})'\big[\mathbf{U}+\lambda_{\max}(\mathbf{P}_{F'}-\mathbf{P}_U)\big]^{+}$$
$$\times(\mathbf{u}-\widehat{\delta\boldsymbol{\gamma}}) \le \chi^2_{r(F'T^{-}\Sigma)}(0;1-\alpha) \Big\}$$

The linearization region $\mathcal{L}_{\delta\boldsymbol{\gamma}}$ is the ellipse in the subspace $\mathcal{M}(\mathbf{F}')$ with the semi-axes

$$a_{\mathcal{L},1} = 1.5539\,\sqrt{\sigma}, \quad a_{\mathcal{L},2} = 0.8972\,\sqrt{\sigma}$$

and $\overline{\mathcal{E}_{\delta\gamma}}$ is the ellipse in $\mathcal{M}(\mathbf{F}')$ with the semi-axes

$$a_{\mathcal{E},1} = 0.2359\,\sigma, \quad a_{\mathcal{E},2} = 0.1362\,\sigma.$$

For $\sigma = 0.5$ it means

$$a_{\mathcal{L},1} = 1.099, \quad a_{\mathcal{L},2} = 0.634$$

and

$$a_{\mathcal{E},1} = 0.118, \quad a_{\mathcal{E},2} = 0.068.$$

Thus the linearization is possible.

As far as the single function of $\boldsymbol{\beta}$ is concerned let us consider $\mathbf{h} = (1,0,0)'$.

$$\mathbf{A}_h = \sum_{s=1}^{6}\left\{ \frac{1}{2}\mathbf{h}'\mathbf{P}_{F'}\big[\mathbf{F}'(\boldsymbol{\Sigma}+\mathbf{F}\mathbf{F}')^{-}\mathbf{F}\big]^{-}\mathbf{F}'(\boldsymbol{\Sigma}+\mathbf{F}\mathbf{F}')^{-} \right\}_s \mathbf{F}_s$$
$$= \frac{1}{18}\begin{pmatrix} 0, & 0, & -6 \\ 0, & 0, & 3 \\ -6, & 3, & 3 \end{pmatrix}$$

and

$$\mathcal{L}_{h'\delta\gamma} = \{\mathbf{u}\colon \mathbf{u}\in\mathcal{M}(\mathbf{F}'), \mathbf{u}'\mathbf{A}_h\mathbf{u} \le \delta_{1,\max}\}$$

where $\delta_{1,\max} = 0.339$ is a solution of the equation

$$P\Big\{\chi_1^2(\delta_{1,\max}) \le \chi_1^2(0; 0.95)\Big\} = 1 - 0.05 - 0.04.$$

The linearization region $\mathcal{L}_{h'\delta\gamma}$ is the hyperbola in $\mathcal{M}(\mathbf{F}')$ with the real semi-axis $a = 1.1768$ and the imaginar bi, $b = 1.714$. Thus the linearization region for the confidence interval for $\delta\gamma_1$ is essentially larger (in the case $\sigma = 0.5$) than the linearization region for the whole vector $\delta\boldsymbol{\gamma}$.

References

[1] Bates, D. M., Watts, D. G.: *Relative curvature measures of nonlinearity.* J. Roy. Stat. Soc. **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.: *Regression models with a weak nonlinearity.* Technical report Nr. 1998.1, Universität Stuttgart, 1998 1–67.

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

[5] Kubáček, L., Tesaříková, E.: *Linearization region for confidence ellispoids.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **47** (2008), 101–113.

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

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

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