Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **41** (2002) 67–81

Weakly Nonlinear Constraints in Regression Models *

LUBOMÍR KUBÁČEK [1], EVA TESAŘÍKOVÁ [2]

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

(Received January 31, 2002)

Abstract

Linear estimators in nonlinear regression models with nonlinear constraints can suffer more or less mainly by bias. The quadratic corrections can help however not in every situation. Some numerical studies of the problem are presented in the paper.

Key words: Nonlinear regression model, nonlinear constraints, linearization, quadratization.

2000 Mathematics Subject Classification: 62J05, 62F10

0 Introduction

Nonlinear regression models with nonlinear constraints create problems in a construction of estimators. Standard method of a linearization can lead to non-negligible bias and can also worsen other properties of estimators. It seems that quadratic corrections should make estimators better. It is really so as far as the bias is concerned. However a quadratic term can produce a large dispersion

*Supported by the grant No. 201/99/0327 of the Grant Agency of the Czech Republic and by the Council of the Czech Government J 14/98: 153 100011.

such that the mean square error characterization of the linear estimator can be better than the mean square error of the quadratic estimator. A theoretical basis for such an investigation is given in [3], however no numerical results are given there. Therefore a follow-up to that paper from the numerical viewpoint is presented here.[1]

1 Notation

The considered model is given in the form

$$\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Sigma}), \quad \boldsymbol{\beta} \in \mathcal{V} = \{\boldsymbol{\beta} : \mathbf{g}(\boldsymbol{\beta}) = \mathbf{0}\}. \tag{1}$$

Here $\mathbf{Y}$ is a normally distributed n-dimensional random vector with the mean value equal to $\mathbf{f}(\boldsymbol{\beta}) = (f_1(\boldsymbol{\beta}), \dots, f_n(\boldsymbol{\beta}))'$ and with the covariance matrix $\boldsymbol{\Sigma}$. The k-dimensional vector $\boldsymbol{\beta}$ is unknown and the parameter space is $\mathcal{V}$. The q-dimensional function $\mathbf{g}(\cdot)$ can be expressed as $\mathbf{g}(\boldsymbol{\beta}^{(0)}) + \mathbf{G}\delta\boldsymbol{\beta} + \frac{1}{2}\gamma(\delta\boldsymbol{\beta})$, where $\boldsymbol{\beta}^{(0)}$ is an approximate value of the actual value $\boldsymbol{\beta}^*$ of the vector $\boldsymbol{\beta}$ and $\delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}^{(0)}$. The $q \times k$ matrix $\mathbf{G}$ is $\partial \mathbf{g}(\mathbf{u})/\partial \mathbf{u}'|_{u=\beta^{(0)}}$, $\gamma(\delta\boldsymbol{\beta}) = (\delta\boldsymbol{\beta}'\mathbf{G}_1\delta\boldsymbol{\beta}, \dots, \delta\boldsymbol{\beta}'\mathbf{G}_q\delta\boldsymbol{\beta})'$, $\mathbf{G}_i = \partial^2 g_i(\mathbf{u})/\partial\mathbf{u}\partial\mathbf{u}'|_{u=\beta^{(0)}}$, $i = 1, \dots, q$. Further it is assumed that the function $\mathbf{f}(\cdot)$ can be expressed analogously, i.e.

$$\mathbf{f}(\boldsymbol{\beta}) = \mathbf{f}(\boldsymbol{\beta}^{(0)}) + \mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\kappa(\delta\boldsymbol{\beta}), \quad \kappa(\delta\boldsymbol{\beta}) = (\delta\boldsymbol{\beta}'\mathbf{F}_1\delta\boldsymbol{\beta}, \dots, \delta\boldsymbol{\beta}'\mathbf{F}_n\delta\boldsymbol{\beta})'.$$

Here

$$\mathbf{F} = \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}'|_{u=\beta^{(0)}}, \qquad \mathbf{F}_i = \partial^2 f_i(\mathbf{u})/\partial\mathbf{u}\partial\mathbf{u}'|_{u=\beta^{(0)}}, \quad i = 1, \dots, n.$$

In the following the vector $\mathbf{g}(\boldsymbol{\beta}^{(0)})$ is assumed to be zero vector (in more detail cf. [2]).

The linearized version of the model (1) is

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

where $\mathbf{f}_0 = \mathbf{f}(\boldsymbol{\beta}^{(0)})$ and the quadratized version of the model (1) is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n(\mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\kappa(\delta\boldsymbol{\beta}), \boldsymbol{\Sigma}), \quad \mathbf{G}\delta\boldsymbol{\beta} + \frac{1}{2}\gamma(\delta\boldsymbol{\beta}) = \mathbf{0}. \tag{3}$$

2 Linear estimator and its bias

In what follows it is assumed that the rank $r(\mathbf{F})$ of the matrix $\mathbf{F}$ is $k < n$, $r(\mathbf{G}) = q < k$ and the matrix $\boldsymbol{\Sigma}$ is positive definite. Under these assumptions the following statements are valid.

[1] A special software Tesaříková, E. and Kubáček, L: Properties of estimators in models with constraints of the type I (in Czech), prepared for the purpose was utilized here.

Statement 2.1 *The best linear unbiased estimator $\delta\hat{\hat{\beta}}$ in (2) is*

$$\begin{aligned}\delta\hat{\hat{\beta}} &= [\mathbf{I} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}]\delta\hat{\beta} \\ &\sim N_k(\delta\beta, \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}),\end{aligned} \tag{4}$$

where $\mathbf{C} = \mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F}$, $\delta\hat{\beta} = \mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0)$ *(the best linear unbiased estimator in the model (2) when the constraints $\mathbf{G}\delta\beta = \mathbf{0}$ are neglected).*

Proof cf. [2]. □

Statement 2.2 *The bias of the estimator (4) in the model (3) is*

$$\begin{aligned}E(\delta\hat{\hat{\beta}}) - \delta\beta &= \mathbf{b} = \frac{1}{2}\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\gamma(\delta\beta) \\ &\quad + \frac{1}{2}[\mathbf{I} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}]\mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\kappa(\delta\beta).\end{aligned} \tag{5}$$

Proof cf. [2]. □

It is quite clear from the relationship (5) how the nonlinearity of the constraints influences the bias of the linear estimator.

The bias from Statement 2.2, i.e. the nonlinear terms in (3) can be neglected, if it is known that the shift $\delta\beta^* = \beta^* - \beta^{(0)}$ (β^* is an actual value of the parameter β in the experiment) is so small that it leads to the inequality $\sqrt{\mathbf{b}'\mathbf{C}\mathbf{b}} < \varepsilon$ for sufficiently small $\varepsilon > 0$.

Statement 2.3 *The set*

$$\left\{\beta \in R^k : ||\mathbf{P}_{F(\beta)}^{\Sigma^{-1}}[\mathbf{Y} - \mathbf{f}(\beta)]||_{\Sigma^{-1}}^2 \leq \chi_k^2(1-\alpha)\right\}$$

is a $(1-\alpha)$-confidence region in the model $\mathbf{Y} \sim N_n[\mathbf{f}(\beta), \mathbf{\Sigma}^{-1}]$. Here $\mathbf{P}_{F(\beta)}^{\Sigma^{-1}}$ is the projection matrix in the Mahalanobis norm $||\mathbf{x}||_{\Sigma^{-1}} = \sqrt{\mathbf{x}'\mathbf{\Sigma}^{-1}\mathbf{x}}$, $\mathbf{x} \in R^n$, on the column space of the matrix $\mathbf{F}(\beta) = \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}|_{u=\beta}$.

Proof cf. Proposition 2.6.1 in [4]. □

Statement 2.4 *In the model*

$$\mathbf{Y} \sim N_n(\mathbf{f}(\beta), \mathbf{\Sigma}), \quad \mathbf{g}(\beta) = \mathbf{0} \tag{6}$$

the $(1-\alpha)$-confidence region is

$$\left\{\beta : \mathbf{g}(\beta) = \mathbf{0}, [\mathbf{Y} - \mathbf{f}(\beta)]'\mathbf{U}[\mathbf{Y} - \mathbf{f}(\beta)] \leq \chi_{k-q}^2(1-\alpha)\right\},$$

where

$$\begin{aligned}\mathbf{U} &= \mathbf{\Sigma}^{-1}\mathbf{F}(\beta)\operatorname{Var}(\hat{\hat{\beta}})\mathbf{F}'(\beta)\mathbf{\Sigma}^{-1}, \\ \mathbf{F}(\beta) &= \partial\mathbf{f}(\beta)/\partial\beta', \\ \mathbf{C}(\beta) &= \mathbf{F}'(\beta)\mathbf{\Sigma}^{-1}\mathbf{F}(\beta), \\ \operatorname{Var}(\hat{\hat{\beta}}) &= \mathbf{C}^{-1}(\beta) - \mathbf{C}^{-1}(\beta)\mathbf{G}'(\beta)[\mathbf{G}(\beta)\mathbf{C}^{-1}(\beta)\mathbf{G}'(\beta)]^{-1}\mathbf{G}(\beta)\mathbf{C}^{-1}(\beta), \\ \mathbf{G}(\beta) &= \partial\mathbf{g}(\beta)/\partial\beta'.\end{aligned}$$

Proof The projection matrix on the tangential space at the point $\mathbf{f}(\beta)$ of the mean value manifold $\{\mathbf{f}(\beta) : \mathbf{g}(\beta) = \mathbf{0}\}$ of the model (6) is

$$\begin{aligned}\mathbf{P}^{\Sigma^{-1}}_{F(\beta)M_{G'(\beta)}} &= \mathbf{F}(\beta)\mathbf{M}_{G'(\beta)}[\mathbf{M}_{G'(\beta)}\mathbf{C}(\beta)\mathbf{M}_{G'(\beta)}]^{+}\mathbf{M}_{G'(\beta)}\mathbf{F}'(\beta)\mathbf{\Sigma}^{-1}\\ &= \mathbf{F}(\beta)\,\mathrm{Var}(\hat{\hat{\beta}})\mathbf{F}'(\beta)\mathbf{\Sigma}^{-1},\end{aligned}$$

where $\mathbf{M}_{G'(\beta)} = \mathbf{I} - \mathbf{P}_{G'(\beta)}$ and $\mathbf{P}_{G'(\beta)}$ is the projection matrix in the Euclidean norm on the column space of the matrix $\mathbf{G}'(\beta)$ and $[\mathbf{M}_{G'(\beta)}\mathbf{C}(\beta)\mathbf{M}_{G'(\beta)}]^{+}$ is the Moore–Penrose generalized inverse of the matrix $\mathbf{M}_{G'(\beta)}\mathbf{C}(\beta)\mathbf{M}_{G'(\beta)}$ (cf. [5]). Now it is sufficient to use Statement 2.3. □

Thus if

$$[\mathbf{Y} - \mathbf{f}(\beta^{(0)})]'\mathbf{\Sigma}^{-1}\mathbf{F}\,\mathrm{Var}(\hat{\hat{\beta}})\mathbf{F}'\mathbf{\Sigma}^{-1}[\mathbf{Y} - \mathbf{f}(\beta^{(0)})] \le \chi^2_{k-q}(1-\alpha), \tag{7}$$

then $\beta^{(0)}$ is covered by the $(1-\alpha)$-confidence region for β^*. The actual value β^* is covered by the $(1-\alpha)$-confidence ellipsoid

$$\mathcal{E} = \left\{\mathbf{u} : (\mathbf{u} - \hat{\hat{\beta}})'[\mathrm{Var}(\hat{\hat{\beta}})]^{-}(\mathbf{u} - \hat{\hat{\beta}}) \le \chi^2_{r[\mathrm{Var}(\hat{\hat{\beta}})]}(1-\alpha)\right\}$$

(if the linearization is possible)). Since the matrix $\mathbf{C}$ is one version of the g-inverse of the matrix $\mathrm{Var}(\hat{\hat{\beta}})$ and $r[\mathrm{Var}(\hat{\hat{\beta}})] = k - q$,

$$\mathcal{E} = \left\{\mathbf{u} : (\mathbf{u} - \hat{\hat{\beta}})'\mathbf{C}(\mathbf{u} - \hat{\hat{\beta}}) \le \chi^2_{k-q}(1-\alpha)\right\}.$$

In an addition the following statement is valid.

Statement 2.5 *Let* $\mathbf{K}_G$ *be a* $k \times (k-q)$ *matrix with the property*

$$\mathcal{M}(\mathbf{K}_G) = \{\mathbf{K}_G\mathbf{u} : \mathbf{u} \in R^{k-q}\} = \mathcal{K}er(\mathbf{G}) = \{\mathbf{v} : \mathbf{G}\mathbf{v} = \mathbf{0}\}$$

and let

$$C^{(par)}_{I,\delta\beta} = \sup\left\{\frac{\sqrt{A+B}}{\mathbf{u}'\mathbf{K}'_G\mathbf{C}\mathbf{K}_G\mathbf{u}} : \mathbf{u} \in R^{k-q}\right\},$$

where

$$\begin{aligned}A &= \boldsymbol{\kappa}'(\mathbf{K}_G\mathbf{u})\mathbf{\Sigma}^{-1}\mathbf{P}^{\Sigma^{-1}}_{FK_G}\boldsymbol{\kappa}(\mathbf{K}_G\mathbf{u}),\\ B &= \boldsymbol{\gamma}'(\mathbf{K}_G\mathbf{u})(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\boldsymbol{\gamma}(\mathbf{K}_G\mathbf{u}),\\ \mathbf{P}^{\Sigma^{-1}}_{FK_G} &= \mathbf{F}\mathbf{K}_G(\mathbf{K}'_G\mathbf{C}\mathbf{K}_G)^{-1}\mathbf{K}'_G\mathbf{F}'\mathbf{\Sigma}^{-1} \quad \textit{(projection matrix on}\\ &\qquad \mathcal{M}(\mathbf{F}\mathbf{K}_G) \textit{ in the norm } \|\mathbf{u}\|_{\Sigma^{-1}} = \sqrt{\mathbf{u}'\mathbf{\Sigma}^{-1}\mathbf{u}}).\end{aligned}$$

If

$$\delta\beta \in \left\{\mathbf{K}_G\mathbf{u} : \mathbf{u}'\mathbf{K}'_G\mathbf{C}\mathbf{K}_G\mathbf{u} \le \frac{2\varepsilon}{C^{(par)}_{I,\delta\beta}}\right\},$$

then

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

Proof cf. in [2]. □

Now if $\chi^2_{k-q}(1-\alpha) \ll 2\varepsilon / C^{(par)}_{I,\delta\beta}$, the model (6) can be linearized (with respect to the bias) at the point $\beta^{(0)}$. However we need the linearization at the point β^* which is unknown. If $\beta^{(0)}$ is chosen in such a way that (7) is satisfied, then β^* is in the neighbourhood of the point $\beta^{(0}$. It can be assumed that the property "can be linearized" does not change in a small neighbourhood of a point of the parametric space $\mathcal{V}$. It would be better to investigate such a neighbourhood in more detail, however in practice it is sufficient to fulfil two conditions

$$\chi^2_{k-q}(1-\alpha) \ll \frac{2\varepsilon}{C^{(par)}_{I,\delta\beta}} \quad \text{and} \quad (7). \tag{8}$$

Since $\mathbf{h}'\mathbf{b} = E(\mathbf{h}'\delta\hat{\hat{\beta}}) - \mathbf{h}'\delta\beta$ and

$$\mathrm{Var}(\mathbf{h}'\delta\hat{\hat{\beta}}) = \mathbf{h}'[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]\mathbf{h} \leq \mathbf{h}'\mathbf{C}^{-1}\mathbf{h},$$

Statement 2.3 enables us to choose for a given vector

$$\mathbf{h} \in \mathcal{M}(\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1})$$

the value of ε in such a way that

$$|\mathbf{h}'\mathbf{b}| \leq \varepsilon_1 \sqrt{\mathbf{h}'[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]\mathbf{h}},$$

i.e. $\varepsilon = \varepsilon_1 \sqrt{\mathbf{h}'[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]\mathbf{h}} / \sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}$. A comparison of the bias with the standard deviation is of no sense in the case

$$\mathbf{h} \perp \mathcal{M}(\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}),$$

since in this case $\mathbf{h}'\,\mathrm{Var}(\delta\hat{\hat{\beta}})\mathbf{h} = 0$. Thus, for practical purposes, a comparison with $\sqrt{\mathbf{h}'\mathbf{C}^{-1}\mathbf{h}}$ seems to be reasonable.

Remark 2.6 The model (3) can be, in some sense, represented by the model without constraints

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\Big(\mathbf{F}\mathbf{K}_G\delta\mathbf{u} + \frac{1}{2}[\boldsymbol{\kappa}(\mathbf{K}_g\delta\mathbf{u}) - \mathbf{F}\mathbf{G}^-_{m(C)}\gamma(\mathbf{K}_G\delta\mathbf{u})], \boldsymbol{\Sigma}\Big), \quad \delta\mathbf{u} \in R^{k-q}.$$

Then the Bates and Watts measures of nonlinearity (cf. [1]) are

$$\begin{aligned} K_I^{(int)} &= \sup\left\{ \frac{\sqrt{\mathbf{k}'\boldsymbol{\Sigma}^{-1}\mathbf{M}^{\Sigma^{-1}}_{FK_G}\mathbf{k}}}{\delta\mathbf{u}'\mathbf{K}'_G\mathbf{C}\mathbf{K}_G\delta\mathbf{u}} : \delta\mathbf{u} \in R^{k-q} \right\} \\ &= \sup\left\{ \frac{\sqrt{\mathbf{k}'\boldsymbol{\Sigma}^{-1}\mathbf{k} - \boldsymbol{\kappa}'(\mathbf{K}_G\delta\mathbf{u})\boldsymbol{\Sigma}^{-1}\mathbf{P}^{\Sigma^{-1}}_{FK_G}\boldsymbol{\kappa}(\mathbf{K}_G\delta\mathbf{u})}}{\delta\mathbf{u}'\mathbf{K}'_G\mathbf{C}\mathbf{K}_G\delta\mathbf{u}} : \delta\mathbf{u} \in R^{k-q} \right\} \end{aligned}$$

and

$$K_I^{(par)} = \sup\left\{\frac{\sqrt{\mathbf{k}'\mathbf{\Sigma}^{-1}\mathbf{P}^{\Sigma^{-1}}_{FK_G}\mathbf{k}}}{\delta\mathbf{u}'\mathbf{K}'_G\mathbf{C}\mathbf{K}_G\delta\mathbf{u}} : \delta\mathbf{u}\in R^{k-q}\right\}$$
$$= \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\mathbf{K}_G\delta\mathbf{u})\mathbf{\Sigma}^{-1}\mathbf{P}^{\Sigma^{-1}}_{FK_G}\boldsymbol{\kappa}(\mathbf{K}_G\delta\mathbf{u})}}{\delta\mathbf{u}'\mathbf{K}'_G\mathbf{C}\mathbf{K}_G\delta\mathbf{u}} : \delta\mathbf{u}\in R^{k-q}\right\},$$

where $\mathbf{k} = \boldsymbol{\kappa}(\mathbf{K}_G\delta\mathbf{u}) - \mathbf{F}\mathbf{G}^-_{m(C)}\gamma(\mathbf{K}_G\delta\mathbf{u})$.

These measures can be used in an investigation of other statistical properties of the model (3). In more detail cf. [2].

3 A quadratic correction of the linear estimator

If the two conditions (8) from the end of the preceding section are not satisfied, the bias can be eliminated by a quadratic correction, i.e. the estimator

$$\begin{aligned}\delta\tilde{\tilde{\beta}} &= \delta\hat{\hat{\beta}} - \frac{1}{2}\left[\mathbf{I}-\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\right]\mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\hat{\hat{\beta}})\\
&- \frac{1}{2}\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\gamma(\delta\hat{\hat{\beta}})\\
&+ \frac{1}{2}\left[\mathbf{I}-\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\right]\mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\begin{pmatrix}\mathrm{Tr}[\mathbf{F}_1\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\\ \vdots\\ \mathrm{Tr}[\mathbf{F}_n\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\end{pmatrix}\\
&+ \frac{1}{2}\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\begin{pmatrix}\mathrm{Tr}[\mathbf{G}_1\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\\ \vdots\\ \mathrm{Tr}[\mathbf{G}_q\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\end{pmatrix}\end{aligned} \tag{9}$$

can be used. Some caution is necessary, since quadratic terms in the estimator can produce a nonnegligible enlargement of the variance. Some comments to it are given in the following text.

Since

$$\mathbf{G}\delta\tilde{\tilde{\beta}} + \frac{1}{2}\gamma(\delta\tilde{\tilde{\beta}}) = -\frac{1}{2}\gamma(\delta\hat{\hat{\beta}}) + \frac{1}{2}\begin{pmatrix}\mathrm{Tr}[\mathbf{G}_1\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\\ \vdots\\ \mathrm{Tr}[\mathbf{G}_q\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\end{pmatrix} + \frac{1}{2}\gamma(\delta\tilde{\tilde{\beta}})$$
$$= \frac{1}{2}\begin{pmatrix}\mathrm{Tr}[\mathbf{G}_1\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\\ \vdots\\ \mathrm{Tr}[\mathbf{G}_q\,\mathrm{Var}(\delta\hat{\hat{\beta}})]\end{pmatrix} + \text{terms of the 3th and higher order,}$$

the estimator $\delta\tilde{\tilde{\beta}}$ does not satisfy the constraints as far as the terms

$$\mathrm{Tr}[\mathbf{G}_i \,\mathrm{Var}(\delta\hat{\hat{\beta}})], \quad i = 1, \ldots, q,$$

are concerned. (It is to be said that the mean value of the estimator $\delta\tilde{\tilde{\beta}}$ from (9) satisfies the constraints $\mathbf{G}\delta\beta + \frac{1}{2}\gamma(\delta\beta) = \mathbf{0}$.)

Now it is to be decided whether satisfying the constraints is more important than the bias due to the terms $\mathrm{Tr}[\mathbf{G}_i \,\mathrm{Var}(\delta\hat{\hat{\beta}})]$, $i = 1, \ldots, q$. In practice satisfying the constraints is preferred. Thus in the following, the estimator

$$\begin{aligned}
\delta\overline{\overline{\beta}} &= \delta\hat{\hat{\beta}} - \frac{1}{2}\left[\mathbf{I} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\right]\mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\hat{\hat{\beta}}) \\
&\quad - \frac{1}{2}\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\gamma(\delta\hat{\hat{\beta}}) \\
&\quad + \frac{1}{2}\left[\mathbf{I} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\right]\mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\begin{pmatrix} \mathrm{Tr}[\mathbf{F}_1 \,\mathrm{Var}(\delta\hat{\hat{\beta}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{F}_n \,\mathrm{Var}(\delta\hat{\hat{\beta}})] \end{pmatrix}
\end{aligned}$$

will be also considered.

Let $h(\delta\beta) = \mathbf{h}'\delta\beta$, where $\mathbf{h} \in R^k$ be given vector. Then the notation

$$\begin{aligned}
\mathbf{F}_{h(\cdot)} &= \sum_{i=1}^{n} \left\{\mathbf{h}'\frac{1}{2}\Big[\mathbf{I} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\Big]\mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\right\}_i \mathbf{F}_i, \\
\mathbf{G}_{h(\cdot)} &= \sum_{i=1}^{q} \left\{\mathbf{h}'\frac{1}{2}\mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\right\}_i \mathbf{G}_i,
\end{aligned}$$

will be used. Let $\mathbf{e}_j \in R^k$ and $\{\mathbf{e}_j\}_i = \delta_{i,j}$ (the Kronecker delta). Then

$$E(\delta\hat{\hat{\beta}}) - \delta\beta = \mathbf{b} = \begin{pmatrix} \delta\beta'\mathbf{A}_1\delta\beta \\ \vdots \\ \delta\beta'\mathbf{A}_k\delta\beta \end{pmatrix},$$

where

$$\mathbf{A}_i = \mathbf{F}_{e_i(\cdot)} + \mathbf{G}_{e_i(\cdot)}, \quad i = 1, \ldots, k.$$

Thus

$$\begin{aligned}
|E(\mathbf{h}'\delta\hat{\hat{\beta}}) - \mathbf{h}'\delta\beta| &= |\delta\beta'\mathbf{A}_{h(\cdot)}\delta\beta|, \quad \mathbf{A}_{h(\cdot)} = \sum_{i=1}^{k}\{\mathbf{h}\}_i\mathbf{A}_i, \\
|E(\mathbf{h}'\delta\overline{\overline{\beta}}) - \mathbf{h}'\delta\beta| &= |\,\mathrm{Tr}[\mathbf{G}_{h(\cdot)}\,\mathrm{Var}(\delta\hat{\hat{\beta}})] + 2\mathbf{b}'\mathbf{A}_{h(\cdot)}\delta\beta + \mathbf{b}'\mathbf{A}_{h(\cdot)}\mathbf{b}|
\end{aligned}$$

and

$$MSE(\mathbf{h}'\delta\hat{\hat{\beta}}) = \mathbf{h}'\,\mathrm{Var}(\delta\hat{\hat{\beta}})\mathbf{h} + \Big[E(\mathbf{h}'\delta\hat{\hat{\beta}}) - \mathbf{h}'\delta\beta\Big]^2,$$

$$\begin{aligned} MSE(\mathbf{h}'\delta\overline{\overline{\beta}}) &= \mathbf{h}'\operatorname{Var}(\delta\overline{\overline{\beta}})\mathbf{h} + \left[E(\mathbf{h}'\delta\overline{\overline{\beta}}) - \mathbf{h}'\delta\beta\right]^2, \\ \mathbf{h}'\operatorname{Var}(\delta\overline{\overline{\beta}})\mathbf{h} &= \mathbf{h}'\operatorname{Var}(\delta\hat{\hat{\beta}}) + 2\operatorname{Tr}\left\{[\mathbf{A}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\beta}})]^2\right\} \\ &\quad - 4\mathbf{h}'\operatorname{Var}(\delta\hat{\hat{\beta}})\mathbf{A}_{h(\cdot)}\delta\beta - 4\mathbf{h}'\operatorname{Var}(\delta\hat{\hat{\beta}})\mathbf{A}_{h(\cdot)}\mathbf{b} \\ &\quad + 4\delta\beta'\mathbf{A}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\beta}})\mathbf{A}_{h(\cdot)}\delta\beta \\ &\quad + 8\mathbf{b}'\mathbf{A}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\beta}})\mathbf{A}_{h(\cdot)}\delta\beta \\ &\quad + 4\mathbf{b}'\mathbf{A}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\beta}})\mathbf{A}_{h(\cdot)}\mathbf{b}. \end{aligned}$$

If the estimator $\mathbf{h}'\delta\overline{\overline{\beta}}$ is used, it is necessary to check whether the ratio

$$\operatorname{Tr}[\mathbf{G}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\beta}})]/\sqrt{\mathbf{h}'[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]\mathbf{h}} \qquad (10)$$

is sufficiently small.

4 Numerical examples

Example 4.1 Let

$$f(x) = \begin{cases} l_1(x;\beta_1,\beta_2) = \beta_1 + \beta_2 x, & x \le 0, \\ l_2(x;\beta_1,\beta_3) = \beta_1\exp(\beta_3 x), & x \ge 0, \end{cases},$$

$$g(\beta_1,\beta_2,\beta_3) = \partial l_2(x;\beta_1,\beta_3)/\partial x|_{x=0} - \partial l_1(x;\beta_1,\beta_2)/\partial x|_{x=0} = \beta_1\beta_3 - \beta_2 = 0,$$

$$\beta_1 = 5, \quad \beta_2 = -4, \quad \beta_3 = -0.8.$$

The values β_1, β_2 and β_3 are chosen as $\beta_1^{(0)}, \beta_2^{(0)}$ and $\beta_3^{(0)}$, respectively.

x	-3	-2	-1	1	2	3
$l_1(x)$	17	13	9			
$l_2(x)$				2.247	1.009	0.454
y	17.1	12.8	8.9	2.3	1.1	0.4

$$\operatorname{Var}(\mathbf{Y}) = \sigma^2\mathbf{I}_{6,6}, \quad \beta_1^{(0)} = 5, \quad \beta_2^{(0)} = -4, \quad \beta_3^{(0)} = -0.8,$$

$$\partial l_1(x,\beta_1,\beta_2,\beta_3)/\partial(\beta_1,\beta_2,\beta_3) = (1, x, 0),$$

$$\partial l_2(x,\beta_1,\beta_2,\beta_3)/\partial(\beta_1,\beta_2,\beta_3) = (\exp(\beta_3 x), 0, \beta_1 x\exp(\beta_3 x)),$$

$$\partial g(\beta_1,\beta_2,\beta_3)/\partial(\beta_1,\beta_2,\beta_3) = (\beta_3, -1, \beta_1).$$

Let $\sigma = 0.2$, $\mathbf{h} = (1,0,0)'$. Then we obtain

$$\mathbf{K}_G = \begin{pmatrix} 0.98791, & 0.00000 \\ -0.03040, & 0.98058 \\ 0.15199, & 0.19612 \end{pmatrix},$$

$$
\begin{aligned}
\delta\hat{\beta} &= (-0.23583, -0.086786, 0.053987)', \\
\mathbf{C}^{-1} &= \begin{pmatrix} 0.086383, & 0.037021, & -0.012123 \\ 0.037021, & 0.018723, & -0.005196 \\ -0.012123, & -0.005196, & 0.005345 \end{pmatrix}, \\
\delta\hat{\hat{\beta}} &= (-0.01713, 0.010692, -0.000602)', \\
\mathrm{Var}(\delta\hat{\hat{\beta}}) &= \begin{pmatrix} 0.019518, & 0.007219, & 0.004567 \\ 0.007219, & 0.005440, & 0.002243 \\ 0.004567, & 0.002243, & 0.001178 \end{pmatrix}, \\
\mathbf{d} &= (0.1397, 0.0738, 0.0343)', \\
\delta\tilde{\tilde{\beta}} &= (-0.01524, 0.01154, 0.00078)'.
\end{aligned}
$$

Here

$$
\mathbf{d} = \left(\sqrt{\mathrm{Var}(\delta\hat{\hat{\beta}}_1)}, \sqrt{\mathrm{Var}(\delta\hat{\hat{\beta}}_2)}, \sqrt{\mathrm{Var}(\delta\hat{\hat{\beta}}_3)}\right)'. \tag{11}
$$

Further

$$
K_I^{(par)} = 0.104963, \quad C_{I,\delta\beta}^{(par)} = 0.105101.
$$

In this case

$$
\chi_2^2(0.95) = 5.99 < \frac{2c}{C_{I,\delta\beta}^{(par)}} = \frac{1}{0.105101} = 9.515\ .
$$

Thus the linearization is possible.

The relationship (10) is

$$
\frac{\mathrm{Tr}[\mathbf{G}_{h(\cdot)}\,\mathrm{Var}(\delta\hat{\hat{\beta}})]}{\sqrt{\mathbf{h}'[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]\mathbf{h}}} = -0.013108
$$

thus the term $\mathrm{Tr}[\mathbf{G}_{h(\cdot)}\,\mathrm{Var}(\delta\hat{\hat{\beta}})]$ can be neglected. Further

$$
\begin{aligned}
\mathbf{h}'\,\mathrm{Var}(\delta\hat{\hat{\beta}})\mathbf{h} &= 0.019518, \quad MSE(\mathbf{h}'\delta\hat{\hat{\beta}}) = 0.019527, \\
\mathbf{h}'\,\mathrm{Var}(\delta\tilde{\tilde{\beta}})\mathbf{h} &= 0.018197, \quad MSE(\mathbf{h}'\delta\tilde{\tilde{\beta}}) = 0.018199.
\end{aligned}
$$

Since $\hat{\hat{\beta}} = (4.982870, -3.989308, -0.800602)'$,

$$
\begin{aligned}
l_1'(0, \hat{\hat{\beta}}) &= \partial l_1(x, \hat{\hat{\beta}})/\partial x\Big|_{x=0} = -3.989308, \\
l_2'(0, \hat{\hat{\beta}}) &= \partial l_2(x, \hat{\hat{\beta}})/\partial x\Big|_{x=0} = -3.989296
\end{aligned}
$$

and $\tilde{\tilde{\beta}} = (4.984760, -3.988460, -0.799220)'$,

$$
l_1'(0, \tilde{\tilde{\beta}}) = -3.988460, \quad l_2'(0, \tilde{\tilde{\beta}}) = -3.983920.
$$

As far as the constraints are concerned the linear estimator seems to be a little better.

It can be of some interest to give the *UBMSE* (upper bounds for single terms of *MSE*). They are given in the following table for

$$\delta\mathbf{s}_1 = \begin{pmatrix} 0.283 \\ 0.000 \end{pmatrix} \Rightarrow \delta\boldsymbol{\beta} = \mathbf{K}_G\delta\mathbf{s}_1 = \begin{pmatrix} 0.280 \\ -0.009 \\ 0.043 \end{pmatrix}$$

and

$$\delta\mathbf{s}_2 = \begin{pmatrix} 0.000 \\ 0.150 \end{pmatrix} \Rightarrow \delta\boldsymbol{\beta} = \mathbf{K}_G\delta\mathbf{s}_2 = \begin{pmatrix} 0.000 \\ 0.147 \\ 0.029 \end{pmatrix};$$

compare **d** (11).

term	*UBMSE*($\delta\mathbf{s}_1$)	*UBMSE*($\delta\mathbf{s}_2$)
$\delta\boldsymbol{\beta}'\mathbf{A}_{h(\cdot)}\delta\boldsymbol{\beta}$	0.073 296	0.067 076
$-4\mathbf{h}'\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{A}_{h(\cdot)}\delta\boldsymbol{\beta}$	0.014 021	0.013 413
$4\delta\boldsymbol{\beta}'\mathbf{A}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{A}_{h(\cdot)}\delta\boldsymbol{\beta}$	0.002 518	0.002 304
$-4\mathbf{h}'\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{A}_{h(\cdot)}\mathbf{b}$	0.002 152	0.001 970
$8\mathbf{b}'\mathbf{A}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{A}_{h(\cdot)}\delta\boldsymbol{\beta}$	0.000 773	0.000 677
$4\operatorname{Tr}[\mathbf{G}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\mathbf{b}'\mathbf{A}_{h(\cdot)}\delta\boldsymbol{\beta}$	0.000 168	0.000 147
$4\mathbf{b}'\mathbf{A}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{A}_{h(\cdot)}\mathbf{b}$	0.000 059	0.000 050
$2\operatorname{Tr}[\mathbf{G}_{h(\cdot)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\mathbf{b}'\mathbf{A}_{h(\cdot)}\mathbf{b}$	0.000 013	0.000 011
$4(\mathbf{b}'\mathbf{A}_{h(\cdot)}\delta\boldsymbol{\beta})^2$	0.000 506	0.000 388
$4\mathbf{b}'\mathbf{A}_{h(\cdot)}\delta\boldsymbol{\beta}\mathbf{b}'\mathbf{A}_{h(\cdot)}\mathbf{b}$	0.000 078	0.000 057
$(\mathbf{b}'\mathbf{A}_{h(\cdot)}\mathbf{b})^2$	0.000 003	0.000 002

In the case $\sigma = 0.2$ the linearization gives satisfactory results, even the quadratic estimator seems to be a little better. For $\mathbf{h} = (0,1,0)'$ and $\mathbf{h} = (0,0,1)'$ we obtain similar results.

In the case $\sigma = 0.5$

$$K_I^{(par)} = 0.262407 \quad \text{and} \quad C_{I,\delta\beta}^{(par)} = 0.262752.$$

Thus

$$\chi_2^2(0.95) = 5.99 > \frac{2c}{C_{I,\delta\beta}^{(par)}} = \frac{1}{0.262752} = 3.806$$

and linearization cannot be recomended. Thus a decision must be made whether a more precise measurement can be organized or it is less tedious to investigate whether a utilization of quadratic corrections is reasonable with respect to *MSE*.

Example 4.2 Let

$$f(x) = \begin{cases} l_1(x;\beta_1) = \exp(\beta_1 x), & x \le 4, \\ l_2(x;\beta_2,\beta_3) = \beta_2 \exp[-(x-\beta_3)^2], & x \ge 4, \end{cases},$$

$$g_1(\beta_1,\beta_2,\beta_3) = l_1(4;\beta_1) - l_2(4;\beta_2,\beta_3) = \exp(4\beta_1) - \beta_2\exp[-(4-\beta_3)^2] = 0,$$

$$\begin{aligned} g_2(\beta_1,\beta_2,\beta_3) &= l_1'(4;\beta_1) - l_2'(4;\beta_2,\beta_3) \\ &= \partial l_1(x,\beta_1)/\partial x|_{x=4} - \partial l_2(x,\beta_2,\beta_3)/\partial x|_{x=4} \\ &= \beta_1\exp(4\beta_1) + 2\beta_2(4-\beta_3)\exp[-(4-\beta_3)^2] = 0, \end{aligned}$$

$$\beta_1 = 0.5, \quad \beta_2 = 7.865609, \quad \beta_3 = 4.25\,.$$

The values $\beta_1^{(0)}, \beta_2^{(0)}$ and $\beta_3^{(0)}$ are chosen as β_1, β_2 and β_3, respectively.

x	1	2	3	5	6	7
$l_1(x)$	1.649	2.718	4.482			
$l_2(x)$				4.482	0.368	0.004
y	1.8	2.5	4.3	4.6	0.3	0

$$\mathrm{Var}(\mathbf{Y}) = \sigma^2 \mathbf{I}_{6,6}, \quad \beta_1^{(0)} = 0.5, \quad \beta_2^{(0)} = 7.865609, \quad \beta_3^{(0)} = 4.25,$$

$$\begin{aligned} \frac{\partial l_1(x,\beta_1)}{\partial(\beta_1,\beta_2,\beta_3)} &= \Big(x\exp(\beta_1 x), 0, 0\Big), \\ \frac{\partial l_2(x,\beta_2,\beta_3)}{\partial(\beta_1,\beta_2,\beta_3)} &= (0, \exp[-(x-\beta_3)^2], 2\beta_2(x-\beta_3)\exp[-(x-\beta_3)^2]), \\ \frac{\partial g_1(\beta_1,\beta_2,\beta_3)}{\partial(\beta_1,\beta_2,\beta_3)} &= \Big(4\exp(4\beta_1), -\exp[-(4-\beta_3)^2], \\ &\qquad -2\beta_2(4-\beta_3)\exp[-(4-\beta_3)^2]\Big), \\ \frac{\partial g_2(\beta_1,\beta_2,\beta_3)}{\partial(\beta_1,\beta_2,\beta_3)} &= ((1+4\beta_1)\exp(4\beta_1), 2(4-\beta_3)\exp[-(4-\beta_3)^2], \\ &\qquad [-2\beta_2 + 4\beta_2(4-\beta_3)^2]\exp[-(4-\beta_3)^2]). \end{aligned}$$

Let $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, $\sigma = 0.2$, $\mathbf{h} = (1,0,0)'$. Then we obtain

$$\begin{aligned} \mathbf{K}_G &= (0.029898, 0.999440, 0.014949)', \\ K_I^{(par)} &= 0.984855, \quad C_{I,\delta\beta}^{(par)} = 1.945840, \end{aligned}$$

$$\delta\mathbf{s} = 1 \quad \Rightarrow \quad \delta\beta = \mathbf{K}_g\delta\mathbf{s} = \begin{pmatrix} 0.029898 \\ 0.999440 \\ 0.014949 \end{pmatrix},$$

$$\delta\hat{\beta} = (-0.015880, 0.313440, -0.064269)',$$

$$
\begin{aligned}
\mathbf{C}^{-1} &= \begin{pmatrix} 0.000188, & 0.000000, & 0.000000 \\ 0.000000, & 0.495600, & -0.125430 \\ 0.000000, & -0.125430, & 0.000916 \end{pmatrix}, \\
\delta\hat{\hat{\beta}} &= (0.000931, 0.031125, 0.000466)', \\
\mathrm{Var}(\delta\hat{\hat{\beta}}) &= \begin{pmatrix} 0.000009839, & 0.000328900, & 0.000004919 \\ 0.000328900, & 0.010995000, & 0.000164470 \\ 0.000004919, & 0.000164470, & 0.000002458 \end{pmatrix}, \\
\mathbf{d} &= (0.003136, 0.104857, 0.001568)', \\
\delta\tilde{\tilde{\beta}} &= (0.000951, 0.031146, 0.000476)',
\end{aligned}
$$

In this case

$$\chi_1^2(0.95) = 3.84 > \frac{2c}{C_{I,\delta\beta}^{(par)}} = \frac{1}{1.945840} = 0.5139.$$

Thus the linearization cannot be recomended.

Further

$$\frac{\mathrm{Tr}[\mathbf{G}_{h(\cdot)}\,\mathrm{Var}(\delta\hat{\hat{\beta}})]}{\sqrt{\mathbf{h}'[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]\mathbf{h}}} = 0.006137.$$

The term $\mathrm{Tr}[\mathbf{G}_{h(\cdot)}\,\mathrm{Var}(\delta\hat{\hat{\beta}})]$ can be neglected. Further

$$
\begin{aligned}
\mathbf{h}'\,\mathrm{Var}(\delta\hat{\hat{\beta}})\mathbf{h} &= 0.000009839, \quad MSE(\mathbf{h}'\delta\hat{\hat{\beta}}) = 0.000013894, \\
\mathbf{h}'\,\mathrm{Var}(\delta\tilde{\tilde{\beta}})\mathbf{h} &= 0.000007214, \quad MSE(\mathbf{h}'\delta\tilde{\tilde{\beta}}) = 0.000007304, \\
\mathbf{b} &= (0.002014, 0.002061, 0.001007)'.
\end{aligned}
$$

Since $\hat{\hat{\beta}} = (0.500931, 7.896734, 4.250476)'$,

$$
\begin{aligned}
l_1(4, \hat{\hat{\beta}}) &= 7.416624, \quad l_2(4, \hat{\hat{\beta}}) = 7.400493 \\
l_1'(4, \hat{\hat{\beta}}) &= 3.715217, \quad l_2'(4, \hat{\hat{\beta}}) = 3.715325
\end{aligned}
$$

and $\tilde{\tilde{\beta}} = (0.500951, 7.896755, 4.250476)'$,

$$
\begin{aligned}
l_1(4, \tilde{\tilde{\beta}}) &= 7.417218, \quad l_2(4, \tilde{\tilde{\beta}}) = 7.416548, \\
l_1'(4, \tilde{\tilde{\beta}}) &= 3.715663, \quad l_2'(4, \tilde{\tilde{\beta}}) = 3.715334\,.
\end{aligned}
$$

In the cases $\mathbf{h} = (0,1,0)'$ and $\mathbf{h} = (0,0,1)'$ similar results are obtained. The model is nonlinear, however not weakly, mainly due to the constraints. The linearization is not suitable, however the quadratic estimation can be satisfactory. Also the constraints are satisfied a little better in the case of quadratic estimators than in the case of linear estimators.

Example 4.3 Let

$$f(x) = \begin{cases} l_1(x;\beta_1) = \beta_1 x, & x \le 5, \\ l_2(x;\beta_2,\beta_3) = \beta_2 \exp(\beta_3 x), & x \ge 5, \end{cases},$$

$$g(\beta_1,\beta_2,\beta_3) = f_1(5;\beta_1) - f_2(5;\beta_2,\beta_3) = 5\beta_1 - \beta_2\exp(\beta_3 5),$$

$$\beta_1 = 1.473, \quad \beta_2 = 33, \quad \beta_3 = -0.299954\,.$$

The values $\beta_1^{(0)}, \beta_2^{(0)}$ and $\beta_3^{(0)}$ are chosen as β_1, β_2 and β_3, respectively.

x	1	2	3	6	7	8
$l_1(x)$	1.473	2.945	4.418			
$l_2(x)$				5.455	4.041	2.994
y	1.2	3.2	4.9	5.1	3.8	2.5

$$\mathrm{Var}(\mathbf{Y}) = \sigma^2\mathbf{I}_{6,6}, \quad \beta_1^{(0)} = 1.473, \quad \beta_2^{(0)} = 33, \quad \beta_3^{(0)} = -0.299954,$$

$$\begin{aligned} \partial l_1(x,\beta_1)/\partial(\beta_1,\beta_2,\beta_3) &= (x,0,0), \\ \partial l_2(x,\beta_2,\beta_3)/\partial(\beta_1,\beta_2,\beta_3) &= \Big(0, \exp(\beta_3 x), x\beta_2\exp(\beta_3 x\Big), \\ \partial g(\beta_1,\beta_2,\beta_3)/\partial(\beta_1,\beta_2,\beta_3) &= \Big(5, -\exp(5\beta_3), -5\beta_2\exp(5\beta_3)\Big). \end{aligned}$$

Let $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, $\sigma = 0.5$, $\mathbf{h} = (0,1,0)'$. Then we obtain

$$\begin{aligned} \mathbf{K}_G &= \begin{pmatrix} 0.990910, & 0 \\ 0.000815, & -0.999980 \\ 0.134540, & 0.006061 \end{pmatrix}, \\ K_I^{(par)} &= 5.262598, \quad C_{I,\delta\beta}^{(par)} = 5.5672, \\ \delta \mathrm{s} &= \begin{pmatrix} 0 \\ 10 \end{pmatrix} \quad \Rightarrow \quad \delta\beta = \mathbf{K}_j\delta\mathrm{s} = \begin{pmatrix} 0 \\ -9.9998 \\ 0.0606 \end{pmatrix}, \\ \delta\hat{\beta} &= (0.119860, 5.86137, -0.0389)', \\ \mathbf{C}^{-1} &= \begin{pmatrix} 0.017857, & 0.000000, & 0.000000 \\ 0.000000, & 391.52, & -1.769 \\ 0.000000, & -1.769, & 0.008095 \end{pmatrix}, \\ \delta\hat{\hat{\beta}} &= (0.084899, 14.569, -0.076767)', \\ \mathrm{Var}(\delta\hat{\hat{\beta}}) &= \begin{pmatrix} 0.013544, & 1.0744, & -0.004672 \\ 1.0744, & 123.93, & -0.6052 \\ -0.004672, & -0.6052, & 0.003034 \end{pmatrix}, \\ \mathbf{d} &= (0.11638, 11.132385, 0.05508)', \\ \delta\tilde{\tilde{\beta}} &= (0.086411, 16.4085, -0.079310)', \end{aligned}$$

In this case

$$\chi_2^2(0.95) = 5.99 > \frac{2c}{C_{I,\delta\beta}^{(par)}} = \frac{1}{5.5672} = 0.17962.$$

Thus the linearization is impossible. The value $C_{I,\delta\beta}^{(par)}$ indicates an extremaly high nonlinearity, which is due to the large value of σ.

Further

$$\frac{\mathrm{Tr}[\mathbf{G}_{h(\cdot)} \operatorname{Var}(\delta\hat{\hat{\beta}})]}{\sqrt{\mathbf{h}'[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}'(\mathbf{G}\mathbf{C}^{-1}\mathbf{G}')^{-1}\mathbf{G}\mathbf{C}^{-1}]\mathbf{h}}} = -0.428117\,.$$

The term $\mathrm{Tr}[\mathbf{G}_{h(\cdot)} \operatorname{Var}(\delta\hat{\hat{\beta}})]$ is not sufficiently small and also other characteristics of nonlinearity are bad. Further

$$\begin{aligned}
\mathbf{h}' \operatorname{Var}(\delta\hat{\hat{\beta}})\mathbf{h} &= 123.93, \quad MSE(\mathbf{h}'\delta\hat{\hat{\beta}}) = 129.511,\\
\mathbf{h}' \operatorname{Var}(\delta\tilde{\tilde{\beta}})\mathbf{h} &= 51.0987, \quad MSE(\mathbf{h}'\delta\tilde{\tilde{\beta}}) = 77.7384,\\
\mathbf{b} &= (-0.001942, -2.263, 0.004875)'.
\end{aligned}$$

Since $\delta\hat{\hat{\beta}} = (0.084899, 14.569, -0.076767)'$,

$$l_1(5,\hat{\hat{\beta}}) = 7.789495, \quad l_2(5,\hat{\hat{\beta}}) = 7.232$$

and $\delta\tilde{\tilde{\beta}} = (0.086411, 16.4085, -0.079310)$,

$$l_1(5,\tilde{\tilde{\beta}}) = 7.797055, \quad l_2(5,\tilde{\tilde{\beta}}) = 7.417206,$$

what is a terrible result and thus the calculation cannot be used.

Let in this example $\sigma = 0.01, \delta\mathbf{s} = (0, 0.1)'$ and the measured data be

$$1.5,\ 2.95,\ 4.4,\ 5.46,\ 4.03,\ 2.98\,.$$

Then

$$K_I^{(par)} = 0.105251, \quad C_{I,\delta\beta}^{(par)} = 0.107389,$$

$$\begin{aligned}
\delta\hat{\beta} &= (-0.001571, 0.608090, -0.002987)',\\
\delta\hat{\hat{\beta}} &= (-0.000051, 0.203970, -0.001229)',\\
\mathbf{d} &= (0.002328, 0.222645, 0.001102)',\\
\delta\tilde{\tilde{\beta}} &= (0.000049, 0.204640, -0.001233)',
\end{aligned}$$

In this case

$$\chi_2^2(0.95) = 5.99 < \frac{2c}{C_{I,\delta\beta}^{(par)}} = \frac{1}{0.107389} = 9.312\,.$$

In this case the linearization is possible. However the value $\sigma = 0.01$ is now essentially smaller than the preceding value 0.5.

Further

$$\mathbf{h}' \operatorname{Var}(\delta\hat{\hat{\beta}})\mathbf{h} = 0.000001213 \quad MSE(\mathbf{h}'\delta\hat{\hat{\beta}}) = 0.000001213,$$
$$\mathbf{h}' \operatorname{Var}(\delta\tilde{\tilde{\beta}})\mathbf{h} = 0.000001237, \quad MSE(\mathbf{h}'\delta\tilde{\tilde{\beta}}) = 0.000001237,$$
$$\mathbf{b} = (-0.00000258, 0.00035817, -0.000002098)'.$$

Since $\hat{\hat{\beta}} = (1.473051, 33.203970, -0.301183)'$,

$$l_1(5,\hat{\hat{\beta}}) = 7.365255, \quad l_2(5,\hat{\hat{\beta}}) = 7.365113$$

a $\tilde{\tilde{\beta}} = (1.473049, 33.204640, -0.301187)'$,

$$l_1(5,\tilde{\tilde{\beta}}) = 7.365245, \quad l_2(5,\tilde{\tilde{\beta}}) = 7.365115\,.$$

In this case the linearization is sufficient.

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.: Statistics and Metrology. *Vyd. Univ. Palackého, Olomouc*, 2000 (in Czech).

[3] Kubáček, L.: *Linearized models with constraints of the type I.* Applications of Mathematics (to appear).

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

[5] Rao, C. R., Mitra, S. K.: Generalized Inverse of the Matrix and Its Applications. *J. Wiley, New York*, 1971.