Applications of Mathematics

Lubomír Kubáček

Linearized regression model with constraints of type II

Applications of Mathematics, Vol. 48 (2003), No. 3, 175–191

Persistent URL: http://dml.cz/dmlcz/134526

LINEARIZED REGRESSION MODEL WITH CONSTRAINTS OF TYPE II*

Lubomír Kubáček, Olomouc

(Received February 15, 2001)

Abstract. A linearization of the nonlinear regression model causes a bias in estimators of model parameters. It can be eliminated, e.g., either by a proper choice of the point where the model is developed into the Taylor series or by quadratic corrections of linear estimators. The aim of the paper is to obtain formulae for biases and variances of estimators in linearized models and also for corrected estimators.

Keywords: nonlinear regression model, linearization, constraints of type II

MSC 2000: 62J05, 62F10

1. Introduction

A model of many experiments can be written in the form

(1)
$$\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}_1), \boldsymbol{\Sigma}), \quad \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} \in \mathcal{V},$$
$$\mathcal{V} = \left\{ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} : \ \mathbf{u} \in \mathbb{R}^{k_1}, \ \ \mathbf{v} \in \mathbb{R}^{k_2}, \ \ \mathbf{h}(\mathbf{u}, \mathbf{v}) = \mathbf{0}_{q,1} \right\}.$$

Here $\mathbf{Y}$ is an n-dimensional random vector (observation vector) normally distributed with the mean value equal to $\mathbf{f}(\boldsymbol{\beta}_1)$ and a covariance matrix equal to $\boldsymbol{\Sigma}$. The unknown k-dimensional vector $\boldsymbol{\beta}$ is an element of the parametric space $\mathcal{V}$, $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')$, $\boldsymbol{\beta}_1$ is k_1-dimensional and $\boldsymbol{\beta}_2$ is k_2-dimensional, $k_1 + k_2 = k$. The covariance matrix $\boldsymbol{\Sigma}$ is known. The constraints $\mathbf{h}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \mathbf{0}$ will be called the constraints of type II.

* This work was supported by Grant No. 201/99/0327 and by Council of the Czech Government J14/98:153100011.

This kind of constraints occurs frequently in chemistry but not only there. An example of a utilization of the regression model with constraints of type II in metrology is presented in [3] and [5]. Constraints of type II are different from constraints of type I (a model with constraints of type I is $\mathbf{Y} \sim N_n(\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Sigma})$, $\{\boldsymbol{\beta}\colon\ \mathbf{h}(\boldsymbol{\beta}) = \mathbf{0}\}$), since in constraints of type II the subvector $\boldsymbol{\beta}_2$ of the vector parameter $\boldsymbol{\beta}$ occurs in the constraints only. The author has not been able to investigate a model with constraints of type II as a special case of the model with constraints of type I and therefore it is studied separately.

In the following let such good approximations $\boldsymbol{\beta}_1^{(0)}$ and $\boldsymbol{\beta}_2^{(0)}$ of the vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, respectively, be known that we can use the quadratic approximation of the functions $\mathbf{b}(\cdot)$ and $\mathbf{h}(\cdot, \cdot\cdot)$.

Let

$$
\begin{aligned}
\boldsymbol{\beta}_1 &= \boldsymbol{\beta}_1^{(0)} + \delta\boldsymbol{\beta}_1, \\
\mathbf{f}(\boldsymbol{\beta}_1) &= \mathbf{f}_0 + \mathbf{F}\delta\boldsymbol{\beta}_1 + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}_1), \\
\mathbf{f}_0 &= \mathbf{f}(\boldsymbol{\beta}_1^{(0)}), \\
\mathbf{F} &= \partial\mathbf{f}(\mathbf{u})/\partial\mathbf{u}'|_{u=\beta_1^{(0)}}, \\
\boldsymbol{\kappa}(\delta\boldsymbol{\beta}_1) &= (\kappa_1(\delta\boldsymbol{\beta}_1), \ldots, \kappa_n(\delta\boldsymbol{\beta}_1))', \\
\kappa_i(\delta\boldsymbol{\beta}_1) &= \partial^2 f_i(\mathbf{u})/\partial\mathbf{u}\partial\mathbf{u}'|_{u=\beta_1^{(0)}}, \quad i = 1, \ldots, n, \\
\mathbf{h}(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) &= \mathbf{H}_1\delta\boldsymbol{\beta}_1 + \mathbf{H}_2\delta\boldsymbol{\beta}_2 + \frac{1}{2}\boldsymbol{\omega}(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2), \\
\mathbf{H}_1 &= \partial\mathbf{h}(\mathbf{u}, \mathbf{v})/\partial\mathbf{u}'|_{u=\beta_1^{(0)},\, v=\beta_2^{(0)}}, \\
\mathbf{H}_2 &= \partial\mathbf{h}(\mathbf{u}, \mathbf{v})/\partial\mathbf{v}'|_{u=\beta_1^{(0)},\, v=\beta_2^{(0)}}, \\
\boldsymbol{\omega}'(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2) &= (\omega_1(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2), \ldots, \omega_q(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2)), \\
\omega_i(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2) &= \begin{pmatrix} \frac{\partial^2 h_i(\mathbf{u},\mathbf{v})}{\partial\mathbf{u}\partial\mathbf{u}'}, & \frac{\partial^2 h_i(\mathbf{u},\mathbf{v})}{\partial\mathbf{u}\partial\mathbf{v}'} \\ \frac{\partial^2 h_i(\mathbf{u},\mathbf{v})}{\partial\mathbf{v}\partial\mathbf{u}'}, & \frac{\partial^2 h_i(\mathbf{u},\mathbf{v})}{\partial\mathbf{v}\partial\mathbf{v}'} \end{pmatrix}\Bigg|_{u=\beta_1^{(0)},\, v=\beta_2^{(0)}}, \quad i = 1, \ldots, q.
\end{aligned}
$$

The choice of $\boldsymbol{\beta}_1^{(0)}$ and $\boldsymbol{\beta}_2^{(0)}$ is such that $\mathbf{h}(\boldsymbol{\beta}_1^{(0)}, \boldsymbol{\beta}_2^{(0)}) = \mathbf{0}$.

The linearized version of (1) is

$$
\mathbf{Y} - \mathbf{f}_0 \sim N_n(\mathbf{F}\delta\boldsymbol{\beta}_1, \boldsymbol{\Sigma}), \quad \mathbf{H}_1\delta\boldsymbol{\beta}_1 + \mathbf{H}_2\delta\boldsymbol{\beta}_2 = \mathbf{0} \tag{2}
$$

and the quadratized version of (1) is

$$
\begin{aligned}
&\mathbf{Y} - \mathbf{f}_0 \sim N_n(\mathbf{F}\delta\boldsymbol{\beta}_1 + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}_1), \boldsymbol{\Sigma}), \\
&\qquad\qquad \mathbf{H}_1\delta\boldsymbol{\beta}_1 + \mathbf{H}_2\delta\boldsymbol{\beta}_2 + \frac{1}{2}\boldsymbol{\omega}(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2) = \mathbf{0}.
\end{aligned} \tag{3}
$$

In the following it is assumed that the rank $r(\mathbf{F})$ of the matrix $\mathbf{F}$ is $r(\mathbf{F}) = k_1 < n$, $r(\mathbf{H}_1, \mathbf{H}_2) = q < k_1 + k_2$ and $r(\mathbf{H}_2) = k_2 < q$. The covariance matrix $\mathbf{\Sigma}$ is known and positive definite.

2. Auxiliary statements

Proofs of the lemmas in this section can be found, e.g., in [2] or [5].

Lemma 2.1. *The* BLUE *(best linear unbiased estimator) of the parameters* $\delta\boldsymbol{\beta}_1$ *and* $\delta\boldsymbol{\beta}_2$ *in the model* (2) *is*

$$
\begin{aligned}
\delta\hat{\hat{\boldsymbol{\beta}}}_1 &= \delta\hat{\boldsymbol{\beta}}_1 - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1\delta\hat{\boldsymbol{\beta}}_1,\\
\delta\hat{\hat{\boldsymbol{\beta}}}_2 &= -\big[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}\big]'\mathbf{H}_1\delta\hat{\boldsymbol{\beta}}_1,\\
\delta\hat{\boldsymbol{\beta}}_1 &= \mathbf{C}^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{f}_0),\\
\mathbf{C} &= \mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F},\\
\mathbf{M}_{H_2} &= \mathbf{I} - \mathbf{H}_2(\mathbf{H}_2'\mathbf{H}_2)^{-1}\mathbf{H}_2',\\
\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1) &= \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1\mathbf{C}^{-1},\\
\operatorname{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2) &= -\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2\\
&\quad\times[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1},\\
\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2) &= [\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1} - \mathbf{I}.
\end{aligned}
$$

The symbol $\mathbf{I}$ *means the identity matrix, the notation* $(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}$ *means the Moore-Penrose generalized inverse of the matrix* $\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2}$ *and* $(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}$ *means the minimum* $\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'$*-seminorm g-inverse of the matrix* $\mathbf{H}_2'$ *(in more detail cf.* [7]*). Further,*

$$
\begin{aligned}
(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+} &= (\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1} - (\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2\\
&\quad\times[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\\
&\quad\times\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1},\\
(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')} &= (\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2\\
&\quad\times[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}.
\end{aligned}
$$

Remark 2.1. The equality

$$\delta\hat{\hat{\boldsymbol{\beta}}}_2 = -[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'\mathbf{H}_1\delta\hat{\hat{\boldsymbol{\beta}}}_1 = -[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'\mathbf{H}_1\delta\hat{\boldsymbol{\beta}}_1$$

can be easily proved. If the aim of the calculation is to determine $\delta\hat{\hat{\boldsymbol{\beta}}}_2$ only, then the last equality may be of some use.

Lemma 2.2. *The estimators from Lemma 2.1 are biased in the model (3) and*

$$E(\delta\hat{\hat{\boldsymbol\beta}}_1) = \delta\boldsymbol\beta_1 + \frac{1}{2}[\mathbf{I} - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1]\mathbf{C}^{-1}\mathbf{F}'\boldsymbol\Sigma^{-1}\boldsymbol\kappa(\delta\boldsymbol\beta_1)$$
$$+ \frac{1}{2}\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\boldsymbol\omega(\delta\boldsymbol\beta_1, \delta\boldsymbol\beta_2),$$
$$= \delta\boldsymbol\beta_1 + \begin{pmatrix} (\delta\boldsymbol\beta_1', \delta\boldsymbol\beta_2')\mathbf{B}^{(e_1)}\begin{pmatrix}\delta\boldsymbol\beta_1\\ \delta\boldsymbol\beta_2\end{pmatrix} \\ \vdots \\ (\delta\boldsymbol\beta_1', \delta\boldsymbol\beta_2')\mathbf{B}^{(e_{k_1})}\begin{pmatrix}\delta\boldsymbol\beta_1\\ \delta\boldsymbol\beta_2\end{pmatrix}\end{pmatrix} = \delta\boldsymbol\beta_1 + \mathbf{b}_1,$$

$$E(\delta\hat{\hat{\boldsymbol\beta}}_2) = \delta\boldsymbol\beta_2 + \frac{1}{2}[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'[\boldsymbol\omega(\delta\boldsymbol\beta_1, \delta\boldsymbol\beta_2) - \mathbf{H}_1\mathbf{C}^{-1}\mathbf{F}'\boldsymbol\Sigma^{-1}\boldsymbol\kappa(\delta\boldsymbol\beta_1)]$$
$$= \delta\boldsymbol\beta_2 + \begin{pmatrix} (\delta\boldsymbol\beta_1', \delta\boldsymbol\beta_2')\mathbf{D}^{(f_1)}\begin{pmatrix}\delta\boldsymbol\beta_1\\ \delta\boldsymbol\beta_2\end{pmatrix} \\ \vdots \\ (\delta\boldsymbol\beta_1', \delta\boldsymbol\beta_2')\mathbf{D}^{(f_{k_2})}\begin{pmatrix}\delta\boldsymbol\beta_1\\ \delta\boldsymbol\beta_2\end{pmatrix}\end{pmatrix} = \delta\boldsymbol\beta_2 + \mathbf{b}_2.$$

Here

$$\mathbf{B}^{(e_i)} = \frac{1}{2}\begin{pmatrix} \sum_{j=1}^{n}\{\mathbf{P}^{C}_{C^{-1}H_1'M_{H_2}}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol\Sigma^{-1}\}_{i,j}\frac{\partial f_j}{\partial\boldsymbol\beta_1\partial\boldsymbol\beta_1'}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0}\end{pmatrix}$$
$$+ \frac{1}{2}\sum_{j=1}^{q}\{\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\}_{i,j}\frac{\partial^2 h_j}{\partial\begin{pmatrix}\boldsymbol\beta_1\\ \boldsymbol\beta_2\end{pmatrix}\partial(\boldsymbol\beta_1', \boldsymbol\beta_2')},$$
$$i = 1, \dots, k_1,$$
$$\mathbf{P}^{C}_{C^{-1}H_1'M_{H_2}} = \mathbf{I} - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1,$$
$$\mathbf{D}^{(f_i)} = -\frac{1}{2}\begin{pmatrix} \sum_{j=1}^{n}\{[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'\mathbf{H}_1\mathbf{C}^{-1}\mathbf{F}'\boldsymbol\Sigma^{-1}\}_{i,j}\frac{\partial^2 f_j}{\partial\boldsymbol\beta_1\partial\boldsymbol\beta_1'}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0}\end{pmatrix}$$
$$+ \frac{1}{2}\sum_{j=1}^{q}\{[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'\}_{i,j}\frac{\partial^2 h_j}{\partial\begin{pmatrix}\boldsymbol\beta_1\\ \boldsymbol\beta_2\end{pmatrix}\partial(\boldsymbol\beta_1', \boldsymbol\beta_2')}, \quad i = 1, \dots, k_2,$$

and

$$\{\mathbf{e}_i\}_j = \delta_{i,j} \quad \textit{Kronecker delta}, \quad \mathbf{e}_i \in \mathbb{R}^{k_1},$$
$$\{\mathbf{f}_i\}_j = \delta_{i,j} \quad \textit{Kronecker delta}, \quad \mathbf{f}_i \in \mathbb{R}^{k_2}.$$

Lemma 3.1. *Generalized inverses of the matrices* $\text{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)$ *and* $\text{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})$*, respectively, are* $\mathbf{C}$ *and* $\begin{pmatrix} \mathbf{C}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0} \end{pmatrix}$*, respectively.*

Proof. It is sufficient to show the equality

$$[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^+\mathbf{H}_1\mathbf{C}^{-1}]\mathbf{C}$$
$$\times [\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^+\mathbf{H}_1\mathbf{C}^{-1}]$$
$$= [\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^+\mathbf{H}_1\mathbf{C}^{-1}],$$

which is elementary. Further, it is necessary to verify the equality

$$\text{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}) \begin{pmatrix} \mathbf{C}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{0} \end{pmatrix} \text{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}) = \text{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}).$$

□

Lemma 2.2 is a basis for the determination of linearization regions, i.e. such sets in the parametric space in which shifts of the parameters do not cause any essential damage of the estimators. It will be formulated more precisely in the sequel.

To follow the idea of [1] let us define a measure of nonlinearity $C_{II,\delta\beta_1}^{(\text{par})}(\boldsymbol{\beta}_0)$.

Definition 3.1.

$$C_{II,\delta\beta_1}^{(par)}(\boldsymbol{\beta}_0) = \sup\left\{ \frac{\sqrt{Q_1 + Q_2}}{\delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s}} : \ \delta\mathbf{s} \in \mathbb{R}^{k_1+k_2-q} \right\},$$

where

$$Q_1 = \boldsymbol{\kappa}'(\mathbf{K}_1\delta\mathbf{s})\boldsymbol{\Sigma}^{-1}\mathbf{P}_{FM_{H_1'M_{H_2}}}^{\Sigma^{-1}}\boldsymbol{\kappa}(\mathbf{K}_1\delta\mathbf{s}),$$
$$Q_2 = \boldsymbol{\omega}'(\mathbf{K}\delta\mathbf{s})\mathbf{T}'\mathbf{C}\mathbf{P}_{C^{-1}H_1'M_{H_2}}^{C}\mathbf{T}\boldsymbol{\omega}(\mathbf{K}\delta\mathbf{s}),$$
$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \end{pmatrix},$$
$$\mathcal{M}(\mathbf{K}) = \text{Ker}(\mathbf{H}_1, \mathbf{H}_2),$$
$$\mathcal{M}(\mathbf{K}_1) = \mathcal{M}(\mathbf{M}_{H_1'M_{H_2}}),$$
$$\mathcal{M}(\mathbf{K}_2) = \mathcal{M}(\mathbf{M}_{H_2'M_{H_1}}),$$
$$\begin{pmatrix} \mathbf{T} \\ \mathbf{U} \end{pmatrix} = (\mathbf{H}_1, \mathbf{H}_2)^-,$$
$$\mathbf{T} = \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1},$$
$$\mathbf{U} = \mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1},$$
$$4\mathbf{b}_1'\mathbf{C}\mathbf{b}_1 = Q_1 + Q_2.$$

R e m a r k 3.1. It is necessary to make some comment to the above definition.

In a model $\mathbf{Y}-\mathbf{f}_0 \sim N_n(\mathbf{F}\delta\boldsymbol{\beta}+\frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}),\boldsymbol{\Sigma})$, $\delta\boldsymbol{\beta}\in\mathbb{R}^k$ (the model without constraints on the parameter $\boldsymbol{\beta}$) the Bates and Watts parametric curvature is defined as

$$K^{(\mathrm{par})}(\boldsymbol{\beta}_0)=\sup\left\{\frac{\sqrt{\boldsymbol{\kappa}(\mathbf{u})'\boldsymbol{\Sigma}^{-1}\mathbf{P}^{\boldsymbol{\Sigma}^{-1}}\boldsymbol{\kappa}(\mathbf{u})}}{\mathbf{u}'\mathbf{C}\mathbf{u}}\colon\ \mathbf{u}\in\mathbb{R}^k\right\}.$$

The nominator in this relation can be expressed also as $2\sqrt{\mathbf{b}'[\mathrm{Var}(\delta\hat{\boldsymbol{\beta}})]^{-1}\mathbf{b}}$, where $\mathbf{b}=E(\delta\hat{\boldsymbol{\beta}})-\delta\boldsymbol{\beta}=\frac{1}{2}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})$. In the model without constraints the covariance matrix $\mathrm{Var}(\delta\hat{\boldsymbol{\beta}})$ is $\mathbf{C}^{-1}$ and it is regular. Thus it is natural to study the size of the bias $\mathbf{b}$ using the expression $\mathbf{b}'\mathbf{C}\mathbf{b}$.

However, the covariance matrices $\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)$ and $\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)$ are not regular in the case of the model with constraints of the type II. It seems that the expression $\mathbf{b}_1'[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)]^-\mathbf{b}_1$ could be used in the definition of the parameter curvature. However, the last expression is not invariant with respect to the choice of the generalized inverse of the matrix $\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)$, since $\mathbf{b}_1\in\mathcal{M}(\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1))$ need not be valid. The positive definite, i.e. regular, version of the generalized inverse is $\mathbf{C}$ (cf. Lemma 3.1).

As far as the denominator in the definition of $C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)$ is concerned it seems that the quantity $\delta\boldsymbol{\beta}'[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]^-\delta\boldsymbol{\beta}$ should be used. Since $\delta\boldsymbol{\beta}=\mathbf{K}\delta\mathbf{s}$ plus terms of the second order and $\mathcal{M}(\mathbf{K})=\mathcal{M}[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]$, the quantity $\delta\boldsymbol{\beta}'[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]^-\delta\boldsymbol{\beta}$ can be expressed as $\delta\mathbf{s}'\mathbf{K}'\begin{pmatrix}\mathbf{C}, & \mathbf{0}\\ \mathbf{0}, & \mathbf{0}\end{pmatrix}\delta\mathbf{s}=\delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s}$.

In the case $\mathbf{b}_1\notin\mathcal{M}(\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)$, it is of no sense to compare the value $\mathbf{h}_1'\mathbf{b}_1$ with the value $\mathbf{h}_1'\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)\mathbf{h}_1$ (it is to be remarked that for $\mathbf{h}_1\perp\mathcal{M}(\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1))$ we have $\mathbf{h}_1'\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)\mathbf{h}_1=\mathrm{Var}(\mathbf{h}_1'\delta\hat{\hat{\boldsymbol{\beta}}}_1)=0$). However, it is reasonable to compare the value $\mathbf{h}_1'\mathbf{b}_1$ with the value $\mathbf{h}_1'\mathbf{C}^{-1}\mathbf{h}_1$, even though $\mathbf{h}_1'\mathbf{C}^{-1}\mathbf{h}_1>\mathrm{Var}(\mathbf{h}_1'\delta\hat{\hat{\boldsymbol{\beta}}}_1)$. That is why Definition 3.1 was used.

One version how to eliminate the constraints in the model (3) is to write

$$\begin{pmatrix}\delta\boldsymbol{\beta}_1\\ \delta\boldsymbol{\beta}_2\end{pmatrix}=\begin{pmatrix}\mathbf{K}_1\\ \mathbf{K}_2\end{pmatrix}\delta\mathbf{s}-\frac{1}{2}\begin{pmatrix}\mathbf{T}\\ \mathbf{U}\end{pmatrix}\boldsymbol{\omega}(\mathbf{K}_1\delta\mathbf{s},\mathbf{K}_2\delta\mathbf{s}),\quad \delta\mathbf{s}\in\mathbb{R}^{k_1+k_2-q}.$$

Now the model (3) can be rewritten as

$$\mathbf{Y}-\mathbf{f}_0\sim N_n[\mathbf{F}(\mathbf{K}_1\delta\mathbf{s}-\frac{1}{2}\mathbf{T}\boldsymbol{\omega}(\mathbf{K}\delta\mathbf{s}))+\frac{1}{2}\boldsymbol{\kappa}(\mathbf{K}_1\delta\mathbf{s}),\boldsymbol{\Sigma}],\quad \delta\mathbf{s}\in\mathbb{R}^{k_1+k_2-q}.$$

Then the following theorem can be proved (in detail cf. [4] and [5]).

Theorem 3.1. *Let*

$$\mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} = \begin{pmatrix} E(\hat{\hat{\boldsymbol{\beta}}}_1) - \boldsymbol{\beta}_1 \\ E(\hat{\hat{\boldsymbol{\beta}}}_2) - \boldsymbol{\beta}_2 \end{pmatrix} \quad \text{(cf. Lemma 2.2).}$$

If

$$\delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s} \leqslant \frac{2\varepsilon}{C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}^{(0)})},$$

then

$$\forall\{\mathbf{h}_1 \in \mathbb{R}^{k_1}\}|\mathbf{h}_1'\mathbf{b}_1| \leqslant \varepsilon\sqrt{\mathbf{h}_1'\mathbf{C}^{-1}\mathbf{h}_1}.$$

Thus the ε-linearization region for the parameter $\boldsymbol{\beta}_1$ can be defined as the set

$$\left\{\delta\boldsymbol{\beta}_1 \colon\ \delta\boldsymbol{\beta}_1 = \mathbf{K}_1\delta\mathbf{s},\ \ \delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s} \leqslant \frac{2\varepsilon}{C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}^{(0)})}\right\}.$$

Definition 3.2.

$$C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}^{(0)}) = \sup\left\{\frac{\sqrt{\mathbf{q}'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{P}_{H_2}^{(H_1C^{-1}H_1'+H_2H_2')^{-1}}\mathbf{q}}}{\delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s}} \colon\ \delta\mathbf{s} \in \mathbb{R}^{k_1+k_2-q}\right\},$$

where

$$\mathbf{q} = \boldsymbol{\omega}(\mathbf{K}\delta\mathbf{s}) - \mathbf{H}_1\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}(\mathbf{K}_1\delta\mathbf{s}).$$

R e m a r k 3.2. It can be easily shown that

$$\frac{\sqrt{\mathbf{q}'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{P}_{H_2}^{(H_1C^{-1}H_1'+H_2H_2')^{-1}}\mathbf{q}}}{\delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s}} = 2\,\frac{\sqrt{\mathbf{b}_2'\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2\mathbf{b}_2}}{\delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s}}.$$

R e m a r k 3.3. Analogously to Remark 3.1 it is necessary to make some comment to Definition 3.2. Since $\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)$ is not regular, it is not suitable to study the size of the bias $\mathbf{b}_2$ using the expression $\mathbf{b}_2'[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)]^-\mathbf{b}_2$ and to compare the value $\mathbf{h}_2'\mathbf{b}_2$ with the value $\mathbf{h}_2'\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)\mathbf{h}_2$. It seems to be reasonable to study the size of the bias $\mathbf{b}_2$ using a norm $\|\mathbf{b}_2\| = \sqrt{\mathbf{b}_2'\mathbf{A}\mathbf{b}_2}$, where $\mathbf{A}$ is a positive definite matrix. In what follows the matrix $\mathbf{A}$ is chosen as $\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2$, i.e. $\mathbf{b}_2'\mathbf{A}\mathbf{b}_2 = \mathbf{b}_2'\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2\mathbf{b}_2$ (the square root of this expression is the nominator in Definition 3.2 divided by 2). In this case it is quite natural to compare the quantity $\mathbf{h}_2'\mathbf{b}_2$ with $\mathbf{h}_2'[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\mathbf{h}_2$. The term $[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}$ is the main term in the expression for $\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)$ $(= [\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1} - \mathbf{I})$.

Theorem 3.2. *If*

$$\delta\mathbf{s}'\mathbf{K}_1'\mathbf{C}\mathbf{K}_1\delta\mathbf{s} \leqslant \frac{2\varepsilon}{C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}^{(0)})},$$

then

$$\forall\{\mathbf{h}_2 \in \mathbb{R}^{k_2}\}|\mathbf{h}_2'\mathbf{b}_2| \leqslant \varepsilon\sqrt{\mathbf{h}_2'[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1' + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\mathbf{h}_2}.$$

P r o o f. Proof is also given in [4] and [5].

4. Quadratic corrections of the estimators

The constraints of the parameters, i.e.

$$\mathbf{H}_1\delta\boldsymbol{\beta}_1 + \mathbf{H}_2\delta\boldsymbol{\beta}_2 + \frac{1}{2}\boldsymbol{\omega}(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2) = \mathbf{0},$$

cannot be satisfied by the linear estimators $\delta\hat{\hat{\boldsymbol{\beta}}}_1$ and $\delta\hat{\hat{\boldsymbol{\beta}}}_2$ exactly, since $\mathbf{H}_1\delta\hat{\hat{\boldsymbol{\beta}}}_1 + \mathbf{H}_2\delta\hat{\hat{\boldsymbol{\beta}}}_2 = \mathbf{0}$. If $\frac{1}{2}\boldsymbol{\omega}(\delta\boldsymbol{\beta}_1, \delta\boldsymbol{\beta}_2)$ cannot be neglected (the linearization region is too small, the nonlinearity of $\boldsymbol{\omega}(\cdot,\cdot\cdot)$ is too large, etc.), hence it is necessary to use other estimators. In the simplest case it is possible to correct the linear estimators by quadratic terms.

Since for any $(k_1 + k_2) \times (k_1 + k_2)$ symmetric matrix $\mathbf{U}$

$$\begin{aligned} E(\delta\hat{\hat{\boldsymbol{\beta}}}{}'\mathbf{U}\delta\hat{\hat{\boldsymbol{\beta}}}) &= [E(\delta\hat{\hat{\boldsymbol{\beta}}})]'\mathbf{U}E(\delta\hat{\hat{\boldsymbol{\beta}}}) + \mathrm{Tr}[\mathbf{U}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ &= \delta\boldsymbol{\beta}'\mathbf{U}\delta\boldsymbol{\beta} + \mathrm{Tr}[\mathbf{U}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ &\quad + \text{terms of the third orders in } \delta\boldsymbol{\beta}, \end{aligned}$$

the quadratic estimators can be written in the form

$$\delta\tilde{\hat{\hat{\boldsymbol{\beta}}}}_1 = \delta\hat{\hat{\boldsymbol{\beta}}}_1 - \begin{pmatrix} \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{B}^{(e_1)}\delta\hat{\hat{\boldsymbol{\beta}}} \\ \vdots \\ \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{B}^{(e_{k_1})}\delta\hat{\hat{\boldsymbol{\beta}}} \end{pmatrix} + \begin{pmatrix} \mathrm{Tr}[\mathbf{B}^{(e_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{B}^{(e_{k_1})}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix},$$

$$\delta\tilde{\hat{\hat{\boldsymbol{\beta}}}}_2 = \delta\hat{\hat{\boldsymbol{\beta}}}_2 - \begin{pmatrix} \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{D}^{(f_1)}\delta\hat{\hat{\boldsymbol{\beta}}} \\ \vdots \\ \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{D}^{(f_{k_2})}\delta\hat{\hat{\boldsymbol{\beta}}} \end{pmatrix} + \begin{pmatrix} \mathrm{Tr}[\mathbf{D}^{(f_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{D}^{(f_{k_2})}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix}.$$

These estimators are unbiased as far as the second order terms are concerned, however, the terms

$$\begin{pmatrix} \mathrm{Tr}[\mathbf{B}^{(e_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{B}^{(e_{k_1})}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix}, \quad \begin{pmatrix} \mathrm{Tr}[\mathbf{D}^{(f_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{D}^{(f_{k_2})}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix}$$

make some problems in the constraints as the following theorem shows.

Theorem 4.1. *Let*

$$\delta\bar{\bar{\boldsymbol{\beta}}}_1 = \delta\tilde{\hat{\hat{\boldsymbol{\beta}}}}_1 - \begin{pmatrix} \mathrm{Tr}[\mathbf{B}^{(e_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{B}^{(e_{k_1})}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix},$$

$$\delta\bar{\bar{\boldsymbol{\beta}}}_2 = \delta\tilde{\hat{\hat{\boldsymbol{\beta}}}}_2 - \begin{pmatrix} \mathrm{Tr}[\mathbf{D}^{(f_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{D}^{(f_{k_2})}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix}.$$

Then

$$\mathbf{H}_1\delta\bar{\bar{\boldsymbol{\beta}}}_1 + \mathbf{H}_2\delta\bar{\bar{\boldsymbol{\beta}}}_2 + \frac{1}{2}\boldsymbol{\omega}(\delta\bar{\bar{\boldsymbol{\beta}}}_1, \delta\bar{\bar{\boldsymbol{\beta}}}_2) = \frac{1}{2}\boldsymbol{\omega}(\delta\bar{\bar{\boldsymbol{\beta}}}_1, \delta\bar{\bar{\boldsymbol{\beta}}}_2) - \frac{1}{2}\boldsymbol{\omega}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2).$$

Proof. Obviously $\mathbf{H}_1\delta\hat{\hat{\boldsymbol{\beta}}}_1 + \mathbf{H}_2\delta\hat{\hat{\boldsymbol{\beta}}}_2 = \mathbf{0}$. Further,

$$-\{\mathbf{H}_1\}_{s,\cdot} \begin{pmatrix} \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{B}^{(e_1)}\delta\hat{\hat{\boldsymbol{\beta}}} \\ \vdots \\ \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{B}^{(e_{k_1})}\delta\hat{\hat{\boldsymbol{\beta}}} \end{pmatrix} - \{\mathbf{H}_2\}_{s,\cdot} \begin{pmatrix} \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{D}^{(f_1)}\delta\hat{\hat{\boldsymbol{\beta}}} \\ \vdots \\ \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{D}^{(f_{k_2})}\delta\hat{\hat{\boldsymbol{\beta}}} \end{pmatrix}$$

$$= -\delta\hat{\hat{\boldsymbol{\beta}}}'\sum_{r=1}^{k_1}\{\mathbf{H}_1\}_{s,r}\mathbf{B}^{(e_r)}\delta\hat{\hat{\boldsymbol{\beta}}} - \delta\hat{\hat{\boldsymbol{\beta}}}'\sum_{r=1}^{k_2}\{\mathbf{H}_2\}_{s,r}\mathbf{D}^{(f_r)}\delta\hat{\hat{\boldsymbol{\beta}}}$$

$$= -\delta\hat{\hat{\boldsymbol{\beta}}}_1'\sum_{r=1}^{k_1}\{\mathbf{H}_1\}_{s,r}\frac{1}{2}\sum_{j=1}^{n}\{[\mathbf{I} - \mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1]$$

$$
\begin{aligned}
&\times \mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\}_{r,j}\frac{\partial^2 f_r}{\partial\boldsymbol{\beta}_1\partial\boldsymbol{\beta}_1'}\delta\hat{\hat{\boldsymbol{\beta}}}_1\\
&-\delta\hat{\hat{\boldsymbol{\beta}}}'\sum_{r=1}^{k_1}\{\mathbf{H}_1\}_{s,r}\frac{1}{2}\sum_{j=1}^{q}\{\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\}_{r,j}\frac{\partial^2 h_r}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\delta\hat{\hat{\boldsymbol{\beta}}}\\
&-\delta\hat{\hat{\boldsymbol{\beta}}}_1'\sum_{r=1}^{k_2}\{\mathbf{H}_2\}_{s,r}\Big(-\frac{1}{2}\Big)\sum_{j=1}^{n}\{[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'\mathbf{H}_1\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\}_{r,j}\\
&\times\frac{\partial^2 f_r}{\partial\boldsymbol{\beta}_1\partial\boldsymbol{\beta}_1'}\delta\hat{\hat{\boldsymbol{\beta}}}_1-\delta\hat{\hat{\boldsymbol{\beta}}}'\sum_{r=1}^{k_2}\{\mathbf{H}_2\}_{s,r}\frac{1}{2}\sum_{j=1}^{q}\{[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'\}_{r,j}\frac{\partial^2 h_r}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\delta\hat{\hat{\boldsymbol{\beta}}}.
\end{aligned}
$$

We also have (terms with $\delta\hat{\hat{\boldsymbol{\beta}}}_1$)

$$
\begin{aligned}
&-\sum_{r=1}^{k_1}\{\mathbf{H}_1\}_{s,r}\frac{1}{2}\sum_{j=1}^{n}\{[\mathbf{I}-\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1]\\
&\times\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\}_{r,j}\\
&-\sum_{r=1}^{k_2}\{\mathbf{H}_2\}_{s,r}\Big(-\frac{1}{2}\Big)\sum_{j=1}^{n}\{[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1')}]'\mathbf{H}_1\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\}_{r,j}\\
=&-\frac{1}{2}\sum_{j=1}^{n}\{[\mathbf{H}_1-\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1]\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\}_{s,j}\\
&+\frac{1}{2}\sum_{j=1}^{n}\{\mathbf{H}_2[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\\
&\times\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\}_{s,j}=\mathbf{0},
\end{aligned}
$$

since

$$
\begin{aligned}
&[\mathbf{H}_1-\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\mathbf{H}_1]\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\\
&\quad=\mathbf{H}_2[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\\
&\quad\times\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}.
\end{aligned}
$$

As far as the terms with $\delta\hat{\hat{\boldsymbol{\beta}}}$ are concerned, we have

$$
\begin{aligned}
&-\sum_{r=1}^{k_1}\{\mathbf{H}_1\}_{s,r}\frac{1}{2}\sum_{j=1}^{q}\{\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^{+}\}_{r,j}\\
&-\sum_{r=1}^{k_2}\{\mathbf{H}_2\}_{s,r}\frac{1}{2}\sum_{j=1}^{q}\{[(\mathbf{H}_2')^{-}_{m(H_1C^{-1}H_1'}]'\}_{r,j}
\end{aligned}
$$

$$
\begin{aligned}
&= -\frac{1}{2}\sum_{j=1}^{q}\{\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'(\mathbf{M}_{H_2}\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'\mathbf{M}_{H_2})^+\}_{s,j}\\
&\quad -\frac{1}{2}\sum_{j=1}^{q}\{\mathbf{H}_2[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\\
&\quad \times \mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\}_{s,j}\\
&= -\frac{1}{2}\sum_{j=1}^{q}\{(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')[(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\\
&\quad -(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\\
&\quad \times \mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}]\}_{s,j}\\
&\quad -\frac{1}{2}\sum_{j=1}^{q}\{\mathbf{H}_2[\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2]^{-1}\\
&\quad \times \mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\}_{s,j}\\
&= -\frac{1}{2}\sum_{j=1}^{q}\{\mathbf{I}\}_{s,j}.
\end{aligned}
$$

Thus

$$
\mathbf{H}_1\delta\overline{\overline{\boldsymbol{\beta}}}_1+\mathbf{H}_2\delta\overline{\overline{\boldsymbol{\beta}}}_2+\frac{1}{2}\boldsymbol{\omega}(\delta\overline{\overline{\boldsymbol{\beta}}})=\frac{1}{2}\boldsymbol{\omega}(\delta\overline{\overline{\boldsymbol{\beta}}})-\frac{1}{2}\boldsymbol{\omega}(\delta\hat{\hat{\boldsymbol{\beta}}}).
$$

□

The bias is better eliminated in the estimators $\delta\tilde{\tilde{\boldsymbol{\beta}}}_1$ and $\delta\tilde{\tilde{\boldsymbol{\beta}}}_2$, however the estimators $\delta\overline{\overline{\boldsymbol{\beta}}}_1$ and $\delta\overline{\overline{\boldsymbol{\beta}}}_2$ better satisfy the constraints

$$
\mathbf{H}_1\delta\boldsymbol{\beta}_1+\mathbf{H}_2\delta\boldsymbol{\beta}_2+\frac{1}{2}\boldsymbol{\omega}(\delta\boldsymbol{\beta}_1,\delta\boldsymbol{\beta}_2)=\mathbf{0}.
$$

Remark 4.1. A decision on the choice of estimators, i.e. $\delta\hat{\hat{\boldsymbol{\beta}}}, \delta\tilde{\tilde{\boldsymbol{\beta}}}$ and $\delta\overline{\overline{\boldsymbol{\beta}}}$, must be made by the user. In practice the constraints are much more important than the bias. Thus in the first step the user must decide whether the term $\frac{1}{2}\boldsymbol{\omega}(\delta\hat{\hat{\boldsymbol{\beta}}}_1,\delta\hat{\hat{\boldsymbol{\beta}}}_2)$ can be neglected and at the same time the conditions of Theorem 3.1 and Theorem 3.2 be satisfied. Then the estimator $\delta\hat{\hat{\boldsymbol{\beta}}}$ is preferred. If the answer is negative, then it must be decided whether the term

$$
\mathbf{H}_1\begin{pmatrix}\operatorname{Tr}[\mathbf{B}^{(e_1)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\\ \vdots\\ \operatorname{Tr}[\mathbf{B}^{(e_{k_1})}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\end{pmatrix}+\mathbf{H}_2\begin{pmatrix}\operatorname{Tr}[\mathbf{D}^{(f_1)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\\ \vdots\\ \operatorname{Tr}[\mathbf{D}^{(f_{k_2})}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\end{pmatrix}\tag{4}
$$

can be neglected. In this case the estimator $\delta\tilde{\tilde{\boldsymbol{\beta}}}$ is to be preferred. If the term (4) cannot be neglected, then $\delta\overline{\overline{\boldsymbol{\beta}}}$ must be used. However, at the same time the term $\frac{1}{2}[\boldsymbol{\omega}(\delta\overline{\overline{\boldsymbol{\beta}}}_1,\delta\overline{\overline{\boldsymbol{\beta}}}_2)-\boldsymbol{\omega}(\delta\hat{\hat{\boldsymbol{\beta}}}_1,\delta\hat{\hat{\boldsymbol{\beta}}}_2)]$ must be negligible.

For the sake of simplicity only the estimators $\delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_1$ and $\delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_2$ will be considered.

Lemma 4.1.

$$E(\delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_1) - \delta\boldsymbol{\beta}_1 = -\begin{pmatrix} \mathrm{Tr}[\mathbf{B}^{(e_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{B}^{(e_{k_1})}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix} + \ \textit{terms of the third order in}\ \ \delta\boldsymbol{\beta},$$

$$E(\delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_2) - \delta\boldsymbol{\beta}_2 = -\begin{pmatrix} \mathrm{Tr}[\mathbf{D}^{(f_1)}\,\mathrm{Var}(\hat{\hat{\boldsymbol{\beta}}})] \\ \vdots \\ \mathrm{Tr}[\mathbf{D}^{(f_{k_2})}\,\mathrm{Var}(\hat{\hat{\boldsymbol{\beta}}})] \end{pmatrix} + \ \textit{terms of the third order in}\ \ \delta\boldsymbol{\beta}.$$

P r o o f is obvious. □

Lemma 4.2.

$$\begin{aligned}
\mathrm{Var}(\delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_1) = {} & \mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1) - 2\begin{pmatrix} (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(e_1)} \\ \vdots \\ (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(e_{k_1})} \end{pmatrix} \begin{pmatrix} \mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1) \\ \mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_2, \delta\hat{\hat{\boldsymbol{\beta}}}_1) \end{pmatrix} \\
& - 2[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1), \mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2)][\mathbf{B}^{(e_1)}(\delta\boldsymbol{\beta}+\mathbf{b}), \dots, \mathbf{B}^{(e_{k_1})}(\delta\boldsymbol{\beta}+\mathbf{b})] \\
& + \{4(\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(e_i)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{B}^{(e_j)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
& + 2\,\mathrm{Tr}[\mathbf{B}^{(e_i)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{B}^{(e_j)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\}_{i,j=1,\dots,k_1}, \\
\mathrm{cov}(\delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_1, \delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_2) = {} & \mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2) \\
& - 2\begin{pmatrix} (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(e_1)} \\ \vdots \\ (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(e_{k_1})} \end{pmatrix} \begin{pmatrix} \mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2) \\ \mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2) \end{pmatrix} \\
& - 2[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1), \mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2)] \\
& \times [\mathbf{D}^{(f_1)}(\delta\boldsymbol{\beta}+\mathbf{b}), \dots, \mathbf{D}^{(f_{k_2})}(\delta\boldsymbol{\beta}+\mathbf{b})] \\
& + \{4(\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(e_i)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{D}^{(f_j)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
& + 2\,\mathrm{Tr}[\mathbf{B}^{(e_i)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{D}^{(f_j)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\}_{i,j=1,\dots,k_1}, \\
\mathrm{Var}(\delta\overline{\overline{\overline{\boldsymbol{\beta}}}}_2) = {} & \mathrm{Var}(\hat{\hat{\boldsymbol{\beta}}}_2) - 2\begin{pmatrix} (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{D}^{(f_1)} \\ \vdots \\ (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{D}^{(f_{k_2})} \end{pmatrix} \begin{pmatrix} \mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \hat{\hat{\boldsymbol{\beta}}}_2) \\ \mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2) \end{pmatrix} \\
& - 2[\mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_2, \delta\hat{\hat{\boldsymbol{\beta}}}_1), \mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)][\mathbf{D}^{(f_1)}(\delta\boldsymbol{\beta}+\mathbf{b}), \dots, \mathbf{D}^{(f_{k_2})}(\delta\boldsymbol{\beta}+\mathbf{b})] \\
& + \{4(\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{D}^{(f_i)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{D}^{(f_j)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
& + 2\,\mathrm{Tr}[\mathbf{D}^{(f_i)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{D}^{(f_j)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\}_{i,j=1,\dots,k_1}.
\end{aligned}$$

P r o o f. Let $\delta\hat{\hat{\boldsymbol{\beta}}} \sim N_k[\delta\boldsymbol{\beta}+\mathbf{b}, \mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]$ and let $\mathbf{L}$ be any k-dimensional vector and $\mathbf{S}_1$, $\mathbf{S}_2$ any $k \times k$ symmetric matrices. Then the statements are direct consequences of the relations

$$\begin{aligned}\mathrm{cov}(\mathbf{L}'\delta\hat{\hat{\boldsymbol{\beta}}}, \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{S}_i\delta\hat{\hat{\boldsymbol{\beta}}}) &= 2\mathbf{L}'\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{S}_i(\delta\boldsymbol{\beta}+\mathbf{b}), \quad i = 1,2,\\ \mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{S}_1\delta\hat{\hat{\boldsymbol{\beta}}}, \delta\hat{\hat{\boldsymbol{\beta}}}'\mathbf{S}_2\delta\hat{\hat{\boldsymbol{\beta}}}) &= 2\,\mathrm{Tr}[\mathbf{S}_1\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{S}_2\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\\ &\quad + 4(\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{S}_1\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{S}_2(\delta\boldsymbol{\beta}+\mathbf{b}).\end{aligned}$$

□

5. MSE of the estimators $\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1$ and $\mathbf{h}_2'\delta\bar{\bar{\boldsymbol{\beta}}}_2$

In the following text the notation

$$\mathbf{B}^{(h_1)} = \sum_{i=1}^{k_1}\{\mathbf{h}_1\}_i\mathbf{B}^{(e_i)}, \quad \mathbf{D}^{(h_2)} = \sum_{i=1}^{k_2}\{\mathbf{h}_2\}_i\mathbf{D}^{(f_i)},$$

where $\mathbf{h}_1 \in \mathbb{R}^{k_1}, \mathbf{h}_2 \in \mathbb{R}^{k_2}$ are any vectors, will be used.

Thus we can write

$$\begin{aligned}\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1 &= \mathbf{h}_1'\delta\hat{\hat{\boldsymbol{\beta}}}_1 - (\delta\hat{\hat{\boldsymbol{\beta}}}_1', \delta\hat{\hat{\boldsymbol{\beta}}}_2')\mathbf{B}^{(h_1)}\begin{pmatrix}\delta\hat{\hat{\boldsymbol{\beta}}}_1\\ \delta\hat{\hat{\boldsymbol{\beta}}}_2\end{pmatrix},\\ \mathbf{h}_2'\delta\bar{\bar{\boldsymbol{\beta}}}_2 &= \mathbf{h}_2'\delta\hat{\hat{\boldsymbol{\beta}}}_2 - (\delta\hat{\hat{\boldsymbol{\beta}}}_1', \delta\hat{\hat{\boldsymbol{\beta}}}_2')\mathbf{D}^{(h_2)}\begin{pmatrix}\delta\hat{\hat{\boldsymbol{\beta}}}_1\\ \delta\hat{\hat{\boldsymbol{\beta}}}_2\end{pmatrix}.\end{aligned}$$

Lemma 5.1.

$$\begin{aligned}\mathrm{MSE}(\mathbf{h}_1'\delta\hat{\hat{\boldsymbol{\beta}}}_1) &= (\mathbf{h}_1'\mathbf{b}_1)^2 + \mathrm{Var}(\mathbf{h}_1'\delta\hat{\hat{\boldsymbol{\beta}}}_1)\\ &= (\delta\boldsymbol{\beta}'\mathbf{B}^{(h_1)}\delta\boldsymbol{\beta})^2 + \mathbf{h}_1'\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)\mathbf{h}_1,\\ \mathrm{MSE}(\mathbf{h}_2'\delta\hat{\hat{\boldsymbol{\beta}}}_2) &= (\mathbf{h}_2'\mathbf{b}_2)^2 + \mathrm{Var}(\mathbf{h}_2'\delta\hat{\hat{\boldsymbol{\beta}}}_2)\\ &= (\delta\boldsymbol{\beta}'\mathbf{D}^{(h_2)}\delta\boldsymbol{\beta})^2 + \mathbf{h}_2'\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)\mathbf{h}_2.\end{aligned}$$

P r o o f. It is a direct consequence of Lemma 2.2. □

Lemma 5.2.

$$\begin{aligned}E(\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1) - \mathbf{h}_1'\delta\boldsymbol{\beta}_1 &= \mathbf{h}_1'\mathbf{b}_1 - (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(h_1)}(\delta\boldsymbol{\beta}+\mathbf{b}) - \mathrm{Tr}[\mathbf{B}^{(h_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]\\ &= -2\delta\boldsymbol{\beta}'\mathbf{B}^{(h_1)}\mathbf{b} - \mathbf{b}'\mathbf{B}^{(h_1)}\mathbf{b} - \mathrm{Tr}[\mathbf{B}^{(h_1)}\,\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})],\end{aligned}$$

$$
\begin{aligned}
E(\mathbf{h}_2'\delta\bar{\bar{\boldsymbol{\beta}}}_2) - \mathbf{h}_2'\delta\boldsymbol{\beta}_2 &= \mathbf{h}_2'\mathbf{b}_2 - (\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{D}^{(h_2)}(\delta\boldsymbol{\beta}+\mathbf{b}) - \mathrm{Tr}[\mathbf{D}^{(h_2)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})] \\
&= -2\delta\boldsymbol{\beta}'\mathbf{D}^{(h_2)}\mathbf{b} - \mathbf{b}'\mathbf{D}^{(h_2)}\mathbf{b} - \mathrm{Tr}[\mathbf{D}^{(h_2)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})], \\
\operatorname{Var}(\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1) &= \mathbf{h}_1'\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)\mathbf{h}_1 \\
&\quad - 4\mathbf{h}_1'[\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1), \operatorname{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2)]\mathbf{B}^{(h_1)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
&\quad + 4(\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(h_1)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
&\quad + 2\,\mathrm{Tr}[\mathbf{B}^{(h_1)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{B}^{(h_1)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})], \\
\operatorname{cov}(\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1, \mathbf{h}_2'\delta\bar{\bar{\boldsymbol{\beta}}}_2) &= \mathbf{h}_1'\operatorname{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2)\mathbf{h}_2 \\
&\quad - 2(\delta\boldsymbol{\beta}+\mathbf{b})\mathbf{B}^{(h_1)}\begin{pmatrix}\operatorname{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2) \\ \operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)\end{pmatrix}\mathbf{h}_2 \\
&\quad - 2\mathbf{h}_1'[\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1), \operatorname{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1, \delta\hat{\hat{\boldsymbol{\beta}}}_2)]\mathbf{D}^{(h_2)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
&\quad + 4(\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{B}^{(h_1)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{D}^{(h_2)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
&\quad + 2\,\mathrm{Tr}[\mathbf{B}^{(h_1)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{D}^{(h_2)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})], \\
\operatorname{Var}(\mathbf{h}_2'\delta\bar{\bar{\boldsymbol{\beta}}}_2) &= \mathbf{h}_2'\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)\mathbf{h}_2 \\
&\quad - 4\mathbf{h}_2'[\operatorname{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_2, \delta\hat{\hat{\boldsymbol{\beta}}}_1), \operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)]\mathbf{D}^{(h_2)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
&\quad + 4(\delta\boldsymbol{\beta}+\mathbf{b})'\mathbf{D}^{(h_2)}(\delta\boldsymbol{\beta}+\mathbf{b}) \\
&\quad + 2\,\mathrm{Tr}[\mathbf{D}^{(h_2)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})\mathbf{D}^{(h_2)}\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})].
\end{aligned}
$$

Proof. It is a direct consequence of Lemma 4.2. □

Since only linear and quadratic estimators are studied, it does not seem to be important to give terms of all powers (in $\delta\boldsymbol{\beta}$) in the expressions for the MSEs. Of course this is true; all terms are given for the sake of completeness only.

6. Upper bounds of MSEs

With the help of the lemmas from Section 5 we can easily compare the values

$$\mathrm{MSE}(\mathbf{h}_1'\delta\hat{\hat{\boldsymbol{\beta}}}_1) = \mathbf{h}_1'\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)\mathbf{h}_1 + (\delta\boldsymbol{\beta}'\mathbf{B}^{(h_1)}\delta\boldsymbol{\beta})^2$$

versus

$$\mathrm{MSE}(\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1) = \operatorname{Var}(\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1) + [E(\mathbf{h}_1'\delta\bar{\bar{\boldsymbol{\beta}}}_1) - \mathbf{h}_1'\delta\boldsymbol{\beta}_1]^2$$

and

$$\mathrm{MSE}(\mathbf{h}_2'\delta\hat{\hat{\boldsymbol{\beta}}}_2) = \mathbf{h}_2'\operatorname{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_2)\mathbf{h}_2 + (\delta\boldsymbol{\beta}'\mathbf{D}^{(h_2)}\delta\boldsymbol{\beta})^2$$

versus

$$\text{MSE}(\mathbf{h}_2'\delta\overline{\overline{\boldsymbol{\beta}}}_2) = \text{Var}(\mathbf{h}_2'\delta\overline{\overline{\boldsymbol{\beta}}}_2) + [E(\mathbf{h}_2'\delta\overline{\overline{\boldsymbol{\beta}}}_2) - \mathbf{h}_2'\delta\boldsymbol{\beta}_2]^2.$$

If the dimension $k = k_1 + k_2$ of the vector $\delta\boldsymbol{\beta} = (\delta\boldsymbol{\beta}_1', \delta\boldsymbol{\beta}_2')'$ is relatively small, then it is possible to calculate the values of MSE in different directions of the shift $\delta\boldsymbol{\beta}$ and to decide whether the quadratic corrections are useful or not.

However, in the case of a large number k this procedure is extremely tedious. Thus it can be useful to know the upper bounds of the MSE values on a boundary of a suitable set, e.g. the confidence region in the linearized model (2). In our case it is given by the relationships

$$P\{\delta\boldsymbol{\beta} \in \mathcal{E}_{1-\alpha}\} = 1 - \alpha,$$

$$\begin{aligned}
\mathcal{E}_{1-\alpha} = &\left\{ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} : \begin{pmatrix} \mathbf{u} - \delta\hat{\hat{\boldsymbol{\beta}}}_1 \\ \mathbf{v} - \delta\hat{\hat{\boldsymbol{\beta}}}_2 \end{pmatrix}' [\text{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]^- \begin{pmatrix} \mathbf{u} - \delta\hat{\hat{\boldsymbol{\beta}}}_1 \\ \mathbf{v} - \delta\hat{\hat{\boldsymbol{\beta}}}_2 \end{pmatrix} \leqslant \chi^2_{k_1+k_2-q}(1-\alpha) \right\} \\
&= \{\mathbf{K}\delta\mathbf{s}\colon \delta\mathbf{s} \in \mathbb{R}^{k_1+k_2-q}, (\mathbf{K}\delta\mathbf{s} - \delta\hat{\hat{\boldsymbol{\beta}}})'[\text{Var}(\delta\hat{\hat{\boldsymbol{\beta}}})]^-(\mathbf{K}\delta\mathbf{s} - \delta\hat{\hat{\boldsymbol{\beta}}}) \\
&\qquad \leqslant \chi^2_{k_1+k_2-q}(1-\alpha)\} \\
&= \{\mathbf{K}\delta\mathbf{s}\colon \delta\mathbf{s} \in \mathbb{R}^{k_1+k_2-q}, (\mathbf{K}_1\delta\mathbf{s} - \delta\hat{\hat{\boldsymbol{\beta}}}_1)'\mathbf{C}(\mathbf{K}_1\delta\mathbf{s} - \delta\hat{\hat{\boldsymbol{\beta}}}_1) \\
&\qquad \leqslant \chi^2_{k_1+k_2-q}(0; 1-\alpha)\} \subset \text{Ker}(\mathbf{H}_1, \mathbf{H}_2),
\end{aligned}$$

where $\chi^2_{k_1+k_2-q}(1-\alpha)$ is the $(1-\alpha)$-quantile of the random variable with the central chi-square distribution with $k_1 + k_2 - q$ degrees of freedom and the matrix $\mathbf{K}$ is given in Section 3.

However, because of the definition of the quantities

$$C^{(\text{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0) \quad \text{and} \quad C^{(\text{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)$$

it will be more suitable to investigate the upper bound of MSE on the boundary of the ellipsoid

$$\mathcal{E}_{c^2} = \{\delta\boldsymbol{\beta}\colon \delta\boldsymbol{\beta}'\mathbf{V}^{-1}\delta\boldsymbol{\beta} = c^2\},$$

where $\mathbf{V}^{-1} = \begin{pmatrix} \mathbf{C}, & \mathbf{0} \\ \mathbf{0}, & \mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1 + \mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2 \end{pmatrix}$. Obviously $\mathcal{E}_{1-\alpha} \supset \mathcal{E}_{c^2}$ for $c^2 = \chi^2_{k_1+k_2-q}(1-\alpha)$. Thus c^2 should be chosen larger than $\chi^2_{k_1+k_2-q}(1-\alpha)$ in such a way that $\mathcal{E}_{1-\alpha} \subset \mathcal{E}_{c^2}$.

Theorem 6.1. *If the value c^2 is chosen, then (expressions for $\delta\boldsymbol{\beta}_1$)*

$$
\begin{aligned}
(\delta\boldsymbol{\beta}'\mathbf{B}^{(h_1)}\delta\boldsymbol{\beta})^2 &\leqslant c^4\,\mathrm{Tr}(\mathbf{B}^{(h_1)}\mathbf{V}\mathbf{B}^{(h_1)}\mathbf{V}),\\
\mathbf{L}' &= -\,4\mathbf{h}_1'[\mathrm{Var}(\delta\hat{\hat{\boldsymbol{\beta}}}_1),\mathrm{cov}(\delta\hat{\hat{\boldsymbol{\beta}}}_1)]\mathbf{B}^{(h_1)},\\
|\mathbf{L}'\delta\boldsymbol{\beta}| &\leqslant c\sqrt{\mathbf{L}'\mathbf{V}\mathbf{L}},\\
|\mathbf{L}'\mathbf{b}| &\leqslant \frac{1}{2}c^2\sqrt{(C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0))^2+(C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0))^2}\sqrt{\mathbf{L}'\mathbf{V}\mathbf{L}},\\
|4\delta\boldsymbol{\beta}'\mathbf{B}^{(h_1)}\delta\boldsymbol{\beta}| &\leqslant 4c^2\,\mathrm{Tr}(\mathbf{B}^{(h_1)}\mathbf{V}\mathbf{B}^{(h_1)}\mathbf{V}),\\
|8\mathbf{b}'\mathbf{B}^{(h_1)}\delta\boldsymbol{\beta}| &\leqslant 4c^3\sqrt{(C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0))^2+(C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0))^2}\sqrt{\mathrm{Tr}[(\mathbf{B}^{(h_1)}\mathbf{V})^2]},\\
|4\mathbf{b}'\mathbf{B}^{(h_1)}\mathbf{b}| &\leqslant c^4[(C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0))^2+(C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0))^2]\sqrt{\mathrm{Tr}(\mathbf{B}^{(h_1)}\mathbf{V}\mathbf{B}^{(h_1)}\mathbf{V})}.
\end{aligned}
$$

Expressions for $\delta\boldsymbol{\beta}_2$ can be found analogously.

P r o o f. The inequalities are based on the Schwarz inequality and on the definitions of the quantities $C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0)$ and $C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0)$. For example (the matrix $\mathbf{A}$ is symmetric),

$$
\begin{aligned}
|\delta\boldsymbol{\beta}'\mathbf{A}\mathbf{b}| &= |\delta\boldsymbol{\beta}\mathbf{V}^{-1/2}\mathbf{V}^{1/2}\mathbf{A}\mathbf{V}^{1/2}\mathbf{V}^{-1/2}\mathbf{b}|\\
&= |\,\mathrm{Tr}(\delta\boldsymbol{\beta}\mathbf{V}^{-1/2}\mathbf{V}^{1/2}\mathbf{A}\mathbf{V}^{1/2}\mathbf{V}^{-1/2}\mathbf{b})|\\
&= |\,\mathrm{Tr}(\mathbf{V}^{1/2}\mathbf{A}\mathbf{V}^{1/2}\mathbf{V}^{-1/2}\mathbf{b}\delta\boldsymbol{\beta}'\mathbf{V}^{-1/2})|\\
&\leqslant \sqrt{\mathrm{Tr}(\mathbf{V}^{1/2}\mathbf{A}\mathbf{V}^{1/2}\mathbf{V}^{1/2}\mathbf{A}\mathbf{V}^{1/2})}\\
&\quad\times\sqrt{\mathrm{Tr}(\mathbf{V}^{-1/2}\mathbf{b}\delta\boldsymbol{\beta}'\mathbf{V}^{-1/2}\mathbf{V}^{-1/2}\delta\boldsymbol{\beta}\mathbf{b}'\mathbf{V}^{-1/2})}\\
&= \sqrt{\mathrm{Tr}(\mathbf{V}\mathbf{A}\mathbf{V}\mathbf{A})}\sqrt{\delta\boldsymbol{\beta}'\mathbf{V}^{-1}\delta\boldsymbol{\beta}\mathbf{b}'\mathbf{V}^{-1}\mathbf{b}}\\
&\leqslant c\sqrt{\mathrm{Tr}(\mathbf{V}\mathbf{A}\mathbf{V}\mathbf{A})}\\
&\quad\times\sqrt{\mathbf{b}_1'\mathbf{C}\mathbf{b}_1+\mathbf{b}_2'\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2\mathbf{b}_2}\\
&\leqslant \frac{1}{2}c\sqrt{\mathrm{Tr}(\mathbf{V}^{-1}\mathbf{A}\mathbf{V}^{-1}\mathbf{A})}\sqrt{(C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0))^2+(C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0))^2}\\
&\quad\times\delta\boldsymbol{\beta}_1\mathbf{C}\delta\boldsymbol{\beta}_1\\
&\leqslant \frac{1}{2}c^3\sqrt{\mathrm{Tr}(\mathbf{V}^{-1}\mathbf{A}\mathbf{V}^{-1}\mathbf{A})}\sqrt{(C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0))^2+(C^{(\mathrm{par})}_{II,\delta\beta_2}(\boldsymbol{\beta}_0))^2}.
\end{aligned}
$$

Here the inequalities

$$
\begin{aligned}
\mathbf{b}_1'\mathbf{C}\mathbf{b}_1 &\leqslant \frac{1}{4}(\delta\boldsymbol{\beta}_1'\mathbf{C}\delta\boldsymbol{\beta}_1)^2(C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0))^2,\\
\mathbf{b}_2'\mathbf{H}_2'(\mathbf{H}_1\mathbf{C}^{-1}\mathbf{H}_1'+\mathbf{H}_2\mathbf{H}_2')^{-1}\mathbf{H}_2\mathbf{b}_2 &\leqslant \frac{1}{4}(\delta\boldsymbol{\beta}_1'\mathbf{C}\delta\boldsymbol{\beta}_1)^2(C^{(\mathrm{par})}_{II,\delta\beta_1}(\boldsymbol{\beta}_0))^2
\end{aligned}
$$

were used. In a similar way all the desired inequalities can be proved. □

References

[1] *D. M. Bates, D. G. Watts*: Relative curvature measures of nonlinearity. J. Roy. Statist. Soc. *B42* (1980), 1–25.

[2] *L. Kubáček, L. Kubáčková, J. Volaufová*: Statistical Models with Linear Structures. Veda, Bratislava, 1995.

[3] *L. Kubáček*: One of the calibration problems. Acta Univ. Palack. Olomuc., Mathematica 36 (1997), 117–130.

[4] *L. Kubáček, L. Kubáčková*: Regression models with a weak nonlinearity. Technical Report Nr. 1998.1. Universität Stuttgart, 1998, pp. 1–67.

[5] *L. Kubáček, L. Kubáčková*: Statistics and Metrology. Palacký University in Olomouc–Publishing House, 2000. (In Czech.)

[6] *C. R. Rao*: Unified theory of linear estimation. Sankhya A *33* (1971), 371–394.

[7] *C. R. Rao*: Generalized Inverse of Matrices and Its Applications. J. Wiley, N. York-London-Sydney-Toronto, 1971.

Author's address: *L. Kubáček*, Department of Mathematical Analysis and Applied Mathematics, Faculty of Science, Palacký University, Tomkova 40, 779 00 Olomouc, Czech Republic, e-mail: `kubacekl@risc.upol.cz`.