AUSTRIAN JOURNAL OF STATISTICS
Volume 35 (2006), Number 2&3, 245–252

Insensitivity Regions and Outliers in Mixed Models with Constraints

Eva Fišerová and Lubomír Kubáček
Palacký University, Olomouc, Czech Republic

Abstract: A procedure for detecting outliers in regular linear regression models with constraints on mean value parameters is presented. A problem, how unknown variance components influence the optimum quality of used test statistics, is studied by sensitivity analysis. Explicit forms of insensitivity regions for testing hypotheses are given.

Keywords: Regression Model with Constraints, Variance Components, Unbiased Estimator, Hypothesis Testing.

1 Introduction

Let the mixed linear regression model when mean value parameters satisfy some linear constraints be under consideration. To detect rough errors or mistakes (outliers) in observations needs a knowledge of a covariance matrix of an observation vector. If unknown variance components occur in it, some problems arise how to recognize, whether estimates or approximations of them can be used instead of their true values. Insensitivity approach is presented in the paper.

The aim of the paper is to give a procedure how to detect outliers and to determine proper insensitivity regions for testing hypotheses.

2 Models with Outliers

A linear regression model with constraints will be denoted as

$$\mathbf{Y} \sim N_n\left(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}\right), \qquad \boldsymbol{\beta} \in \mathcal{V} = \{\mathbf{u} : \mathbf{b} + \mathbf{B}\mathbf{u} = \mathbf{0}\}. \tag{1}$$

Here $\mathbf{Y}$ is an n-dimensional random vector (observation vector) which is normally distributed, its mean value is $\mathrm{E}(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ and the covariance matrix is $\mathrm{var}(\mathbf{Y}) = \boldsymbol{\Sigma}$. The parametric space for $\boldsymbol{\beta}$ is $\mathcal{V}$, $\boldsymbol{\beta} \in \mathbb{R}^k$ is an unknown parameter, $\mathbf{X}$ and $\mathbf{B}$ are given matrices of the type $n \times k$ and $q \times k$, respectively, $\mathbf{b} \in \mathbb{R}^q$ is a given vector.

The model (1) will be supposed to be regular, i.e., the matrix $\mathbf{X}$ has the full column rank ($\mathrm{rank}(\mathbf{X}) = k < n$), $\boldsymbol{\Sigma}$ is positive definite (p.d.) and $\mathrm{rank}(\mathbf{B}) = q < k$.

There are several procedures to detect outliers in measurements, cf. e.g. Gnanadesikan (1977). Here the approach given by Zvára (1989) is used. At the first step the parameter $\boldsymbol{\beta}$ in the model (1) is estimated as the best linear unbiased estimator (BLUE) $\hat{\boldsymbol{\beta}}$.

Lemma 2.1. *In the model (1) the BLUE $\hat{\boldsymbol{\beta}}$ of the parameter $\boldsymbol{\beta}$ is*

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= (\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} - \mathbf{C}^{-1}\mathbf{B}'(\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{b}\,,\\
\mathrm{var}(\hat{\boldsymbol{\beta}}) &= \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{B}'(\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{B}\mathbf{C}^{-1} = (\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\,,\\
\mathbf{C} &= \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\,.
\end{aligned}$$

(The symbol $(\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^+$ *means the Moore-Penrose generalized inverse of the matrix* $\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'}$ *(cf. Rao and Mitra, 1971). Here* $\mathbf{M}_{B'} = \mathbf{I} - \mathbf{P}_{B'}$, $\mathbf{P}_{B'} = \mathbf{B}'(\mathbf{B}')^+$*.)*

Proof. Proof is given, e.g. in Kubáček et al. (1995, p. 80). □

The residual vector $\mathbf{v} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$ is distributed as

$$\mathbf{v} \sim N_n \left[\mathbf{0}, \boldsymbol{\Sigma} - \mathbf{X}(\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^+\mathbf{X}'\right] .$$

Suspicious measurements y_i, $i = i_1, \dots, i_r$, can be found by testing the null hypothesis $H_0 : \mathrm{E}(\{\mathbf{v}\}_i) = 0$, $i = 1, \dots, n$, against the alternative hypothesis $H_a : \exists i : \mathrm{E}(\{\mathbf{v}\}_i) \neq 0$ by the help of the test statistic

$$T = \mathbf{v}'[\mathrm{var}(\mathbf{v})]^-\mathbf{v} \sim \chi^2_{n+q-k}(\delta_1) , \qquad \delta_1 = \mathrm{E}(\mathbf{v})'[\mathrm{var}(\mathbf{v})]^-\mathrm{E}(\mathbf{v}) .$$

One version of $[\mathrm{var}(\mathbf{v})]^-$ is $\boldsymbol{\Sigma}^{-1}$. With respect to Scheffé (1959) it is valid

$$\forall \mathbf{h} \in \mathbb{R}^n : |\mathbf{h}'\mathbf{v}| \leq \sqrt{\chi^2_{n+q-k}(1-\alpha)}\sqrt{\mathbf{h}'\mathrm{var}(\mathbf{v})\mathbf{h}} \;\Leftrightarrow\; \mathbf{v}'[\mathrm{var}(\mathbf{v})]^-\mathbf{v} \leq \chi^2_{n+q-k}(1-\alpha) ,$$

where $\chi^2_{n+q-k}(1-\alpha)$ is the $(1-\alpha)$ quantile of the chi-square distribution with $n+q-k$ degrees of freedom. Thus if

$$|\{\mathbf{v}\}_i| \geq \sqrt{\chi^2_{n+q-k}(1-\alpha)}\sqrt{\{\mathrm{var}(\mathbf{v})\}_{ii}} , \qquad i \in \{1, \dots, n\}$$

then the ith measurement is considered to be suspicious. However $\chi^2_{n+q-k}(1-\alpha)$ in some cases is rather large and therefore in practice the value $u(1-\alpha/2)$ (the $(1-\alpha/2)$-quantile of the normal distribution $N_1(0,1)$) is used instead of $\sqrt{\chi^2_{n+q-k}(1-\alpha)} > u(1-\alpha/2)$.

If no suspicious large value $|\{\mathbf{v}\}_i|$, i.e., no suspicious measurement, is found, stop this procedure. If r suspicious measurements are found, the test given in the following text is a basis for a decision whether suspicious measurements are outliers or not. Let the model (1) be rewritten in the form

$$\mathbf{Y} \sim N_n \left[(\mathbf{X}, \mathbf{E})\begin{pmatrix}\boldsymbol{\beta}\\ \boldsymbol{\Delta}\end{pmatrix}, \boldsymbol{\Sigma}\right] , \qquad \boldsymbol{\beta} \in \mathcal{V} = \{\mathbf{u} : \mathbf{b} + \mathbf{B}\mathbf{u} = \mathbf{0}\} , \qquad \boldsymbol{\Delta} \in \mathbb{R}^r , \quad (2)$$

where

$$\mathbf{E} = (\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_r}) , \quad \mathbf{e}_{i_j} \in \mathbb{R}^n , \quad j = 1, \dots, r , \quad \{\mathbf{e}_{i_j}\}_k = \begin{cases} 0, & k \neq i_j, \\ 1, & k = i_j, \end{cases}$$

and i_j is the index with suspicious large value $|\{\mathbf{v}\}_{i_j}|$.

Lemma 2.2. *The hypothesis* $H_0 : \boldsymbol{\Delta} = \mathbf{0}$ *versus* $H_a : \boldsymbol{\Delta} \neq \mathbf{0}$ *can be tested in the model (2) if and only if*

$$\mathcal{M}(\mathbf{X}\mathbf{M}_{B'}) \cap \mathcal{M}(\mathbf{E}) = \{\mathbf{0}\} \quad \Leftrightarrow \quad \mathcal{M}\begin{pmatrix}\mathbf{X}\\ \mathbf{B}\end{pmatrix} \cap \mathcal{M}\begin{pmatrix}\mathbf{E}\\ \mathbf{0}\end{pmatrix} = \{\mathbf{0}\} . \qquad (3)$$

(Here $\mathcal{M}(\mathbf{A}_{m,n}) = \{\mathbf{A}\mathbf{u} : \mathbf{u} \in \mathbb{R}^n\}$ *is a linear subspace generated by columns of the matrix* $\mathbf{A}$*.)*

Proof. The hypothesis $\mathbf{0}\boldsymbol{\beta} + \mathbf{I}\boldsymbol{\Delta} = \mathbf{0}$ can be tested iff $\mathcal{M}\left(\begin{smallmatrix}\mathbf{0}\\ \mathbf{I}\end{smallmatrix}\right) \subset \mathcal{M}\begin{pmatrix}\mathbf{X}', & \mathbf{B}'\\ \mathbf{E}', & \mathbf{0}\end{pmatrix}$ (cf. Zvára, 1989). Thus it must be true $\mathcal{M}(\mathbf{E}'\mathbf{M}_{XM_{B'}}) = \mathcal{M}(\mathbf{I}) = \mathbb{R}^r$. Since

$$\operatorname{rank}\begin{pmatrix}\mathbf{M}_{B'}\mathbf{X}'\\ \mathbf{E}'\end{pmatrix} = \operatorname{rank}\left(\mathbf{E}'\mathbf{M}_{XM_{B'}}\right) + \operatorname{rank}(\mathbf{X}\mathbf{M}_{B'})$$

(cf. Rao and Mitra, 1971, p. 137), the equality $\operatorname{rank}(\mathbf{E}'\mathbf{M}_{XM_{B'}}) = r$ can be valid iff $\mathcal{M}(\mathbf{X}\mathbf{M}_{B'}) \cap \mathcal{M}(\mathbf{E}) = \{\mathbf{0}\}$ ($\operatorname{rank}(\mathbf{E}') = r$).

The equivalence (3) is implied by the following equivalence

$$\mathcal{M}(\mathbf{X}\mathbf{M}_{B'}) \cap \mathcal{M}(\mathbf{E}) = \{\mathbf{0}\} \quad \Leftrightarrow \quad \operatorname{rank}\begin{pmatrix}\mathbf{M}_{B'}\mathbf{X}'\\ \mathbf{E}'\end{pmatrix} = \operatorname{rank}(\mathbf{E}') + \operatorname{rank}(\mathbf{M}_{B'}\mathbf{X}')\,,$$

$$\mathcal{M}\begin{pmatrix}\mathbf{X}\\ \mathbf{B}\end{pmatrix} \cap \mathcal{M}\begin{pmatrix}\mathbf{E}\\ \mathbf{0}\end{pmatrix} = \{\mathbf{0}\} \quad \Leftrightarrow \quad \operatorname{rank}\begin{pmatrix}\mathbf{X}, & \mathbf{E}\\ \mathbf{B}, & \mathbf{0}\end{pmatrix} = \operatorname{rank}\begin{pmatrix}\mathbf{X}\\ \mathbf{B}\end{pmatrix} + \operatorname{rank}(\mathbf{E})\,.$$

In both cases the equality $\operatorname{rank}(\mathbf{E}'\mathbf{M}_{XM_{B'}}) = \operatorname{rank}(\mathbf{E}')$ is necessary and sufficient condition for equivalence (3). □

Lemma 2.3. *Let the condition (3) be fulfilled. In the model (2) BLUEs of parameters* $\boldsymbol{\beta}$ *and* $\boldsymbol{\Delta}$ *are*

$$\begin{pmatrix}\hat{\boldsymbol{\beta}}_{out}\\ \hat{\boldsymbol{\Delta}}\end{pmatrix} = \begin{pmatrix}\hat{\boldsymbol{\beta}} - (\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{E}\hat{\boldsymbol{\Delta}}\\ \left[\mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{B'}}\right)^{+}\mathbf{E}\right]^{-1}\mathbf{E}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})\end{pmatrix}.$$

Further

$$\begin{aligned}
\operatorname{var}(\hat{\boldsymbol{\beta}}_{out}) &= \operatorname{var}(\hat{\boldsymbol{\beta}}) + (\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\,\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{E}\\
&\quad\times\left[\mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{B'}}\right)^{+}\mathbf{E}\right]^{-1}\mathbf{E}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\left(\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'}\right)^{+},\\
\operatorname{var}(\hat{\boldsymbol{\Delta}}) &= \left[\mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{B'}}\right)^{+}\mathbf{E}\right]^{-1},\\
\operatorname{cov}(\hat{\boldsymbol{\beta}}_{out}, \hat{\boldsymbol{\Delta}}) &= -(\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{E}\left[\mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{B'}}\right)^{+}\mathbf{E}\right]^{-1}.
\end{aligned}$$

Proof. At first it is to be remarked that regularity of the matrix $\mathbf{E}'(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{B'}})^{+}\mathbf{E}$ is implied by the assumption (3) and $\operatorname{rank}(\mathbf{E}_{n,r}) = r < n$, respectively.

Let $\boldsymbol{\beta}_0$ be any solution of the equation $\mathbf{B}\boldsymbol{\beta} + \mathbf{b} = \mathbf{0}$, i.e., $\boldsymbol{\beta} = \boldsymbol{\beta}_0 + \mathbf{M}_{B'}\boldsymbol{\gamma}$, $\boldsymbol{\gamma} \in \mathbb{R}^k$. Thus we obtain the model without constraints

$$\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0 \sim N_n\left[(\mathbf{X}\mathbf{M}_{B'}, \mathbf{E})\begin{pmatrix}\boldsymbol{\gamma}\\ \boldsymbol{\Delta}\end{pmatrix}, \boldsymbol{\Sigma}\right], \qquad \boldsymbol{\gamma} \in \mathbb{R}^k, \quad \boldsymbol{\Delta} \in \mathbb{R}^r,$$

which is not regular, however the assumption (3) ensures the estimability of vectors $\mathbf{M}_{B'}\boldsymbol{\gamma}$ and $\boldsymbol{\Delta}$. BLUEs of $\mathbf{M}_{B'}\boldsymbol{\gamma}$ and $\boldsymbol{\Delta}$ are

$$\begin{pmatrix}\widehat{\mathbf{M}_{B'}\boldsymbol{\gamma}}\\ \hat{\boldsymbol{\Delta}}\end{pmatrix} = \left[\begin{pmatrix}\mathbf{M}_{B'}\mathbf{X}'\\ \mathbf{E}'\end{pmatrix}\boldsymbol{\Sigma}^{-1}(\mathbf{X}\mathbf{M}_{B'}, \mathbf{E})\right]^{+}\begin{pmatrix}\mathbf{M}_{B'}\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)\\ \mathbf{E}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}_0)\end{pmatrix}.$$

The covariance matrix of the estimator $(\hat{\boldsymbol{\beta}}'_{out}, \hat{\boldsymbol{\Delta}}')'$ is

$$\mathrm{var}\begin{pmatrix}\hat{\boldsymbol{\beta}}_{out}\\ \hat{\boldsymbol{\Delta}}\end{pmatrix} = \left[\begin{pmatrix}\mathbf{M}_{B'}\mathbf{X}'\\ \mathbf{E}'\end{pmatrix}\boldsymbol{\Sigma}^{-1}(\mathbf{X}\mathbf{M}_{B'}, \mathbf{E})\right]^{+} = \begin{pmatrix}\boxed{11}, & \boxed{12}\\ \boxed{21}, & \boxed{22}\end{pmatrix},$$

where

$$\begin{aligned}
\boxed{11} &= (\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+} - (\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{E}\boxed{21},\\
\boxed{12} &= -(\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{E}\left[\mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{B'}}\right)^{+}\mathbf{E}\right]^{-1} = \boxed{21}',\\
\boxed{22} &= \left[\mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{B'}}\right)^{+}\mathbf{E}\right]^{-1},\\
\boldsymbol{\beta}_0 &= -\mathbf{C}^{-1}\mathbf{B}'(\mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{b}\,.
\end{aligned}$$

□

Now outliers among suspicious measurements y_i, $i = i_1, \ldots, i_r$, can be found by testing the null hypothesis $H_0 : \boldsymbol{\Delta} = \mathbf{0}$ against the alternative hypothesis $H_a : \boldsymbol{\Delta} \neq \mathbf{0}$ by the help of the test statistic

$$T_{out} = \widehat{\boldsymbol{\Delta}}'[\mathrm{var}(\widehat{\boldsymbol{\Delta}})]^{-1}\widehat{\boldsymbol{\Delta}} \sim \chi^2_r(\delta_2)\,, \qquad \delta_2 = \boldsymbol{\Delta}'[\mathrm{var}(\widehat{\boldsymbol{\Delta}})]^{-1}\boldsymbol{\Delta}\,.$$

Similarly as in the case of selection of suspicious measurements, if

$$|\{\boldsymbol{\Delta}\}_{i^*}| \geq \sqrt{\chi^2_r(1-\alpha)}\sqrt{\left\{\mathrm{var}(\widehat{\boldsymbol{\Delta}})\right\}_{i^*i^*}}\,, \qquad i^* \in \{i_1, \ldots, i_r\}\,,$$

where i_j is the index with suspicious measurement, then the i^*th measurement is considered to be outlier. The value $u(1-\alpha/2)$ instead of $\sqrt{\chi^2_r(1-\alpha)}$ is used sometimes in practice. Some precaution is necessary in the case

$$u(1-\alpha/2)\sqrt{\left\{\mathrm{var}(\widehat{\boldsymbol{\Delta}})\right\}_{i^*i^*}} \leq \left|\{\widehat{\boldsymbol{\Delta}}\}_{i^*}\right| \leq \sqrt{\chi^2_r(1-\alpha)}\sqrt{\left\{\mathrm{var}(\widehat{\boldsymbol{\Delta}})\right\}_{i^*i^*}}\,.$$

At the last step outliers y_{i^*} are omitted from realization of the observation vector $\mathbf{Y}$ and the whole procedure is repeated.

Remark 2.4. *It is well known that the least squares (LS) methodology leads to an effect that outliers significantly influence estimates in a contradiction to the robust methodology with just opposite effect. Thus in the LS-methodology an outlier can be nonsignificantly overlapped by a residual influenced by the outlier itself.*

3 Problem of Variance Components

In this section we use the insensitivity approach. For more detail we refer to Kubáček and Kubáčková (2000), Lešanská (2001, 2002).

Let the covariance matrix in models (1), (2) be considered in the form $\boldsymbol{\Sigma} = \sum_{i=1}^{p} \vartheta_i \mathbf{V}_i$, $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_p)' \in \underline{\vartheta} \subset \mathbb{R}^p$. Such models are called mixed models. Here except $\boldsymbol{\beta}$ and

$\boldsymbol{\Delta}$ also the vector parameter $\boldsymbol{\vartheta}$ is unknown. $\mathbf{V}_1, \ldots, \mathbf{V}_p$ are known symmetric matrices. The parameter space $\underline{\vartheta}$ is an open set in $\mathbb{R}^p$ with the property that if $\boldsymbol{\vartheta} \in \underline{\vartheta}$, then $\sum_{i=1}^p \vartheta_i \mathbf{V}_i$ is p.d.

If an approximation $\boldsymbol{\vartheta}_0$ of the parameter $\boldsymbol{\vartheta}$ is at our disposal, outliers can be detected by the help of the procedure given in the previous section. In this case estimators are $\boldsymbol{\vartheta}_0$-locally best linear unbiased only.

However the substitution of the true value $\boldsymbol{\vartheta}^*$ by its approximation $\boldsymbol{\vartheta}_0$ can destroy the optimum quality of used statistical inference, namely the risk of tests T and T_{out}. Hence the question arises which values of $\boldsymbol{\vartheta}_0$ make increase of the risk of the test α not larger than tolerable given ε. One possible solution is to find insensitivity regions $\mathcal{N}_\varepsilon$ and $\mathcal{N}_{out,\varepsilon}$ of all points $\boldsymbol{\vartheta}_0 = \boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta}$ such that if $\boldsymbol{\vartheta}_0 \in \mathcal{N}_\varepsilon$ and $\boldsymbol{\vartheta}_0 \in \mathcal{N}_{out,\varepsilon}$, then the risk of tests T and T_{out}, respectively, is not larger than $\alpha + \varepsilon$.

Lemma 3.1. *The infinitesimal approximation of* $T(\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta})$ *is*

$$T(\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta}) \approx T(\boldsymbol{\vartheta}^*) + \boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\,,$$

where

$$\begin{aligned}\{\boldsymbol{\eta}(\boldsymbol{\vartheta}^*)\}_i = \mathbf{v}'(\boldsymbol{\vartheta}^*)\Big\{&2\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\mathbf{X}[\mathbf{M}_{B'}\mathbf{C}(\boldsymbol{\vartheta}^*)\mathbf{M}_{B'}]^+\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\mathbf{V}_i\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*) \\ &- \boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\mathbf{V}_i\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\Big\}\mathbf{v}(\boldsymbol{\vartheta}^*)\,, \quad i = 1, \ldots, p\,.\end{aligned}$$

Further

$$\begin{aligned}\mathrm{E}[\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}] &= -\mathbf{a}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\,, \\ \mathrm{var}[\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}] &= 2\delta\boldsymbol{\vartheta}'\mathbf{S}_K\delta\boldsymbol{\vartheta}\,,\end{aligned}$$

where

$$\{\mathbf{a}(\boldsymbol{\vartheta}^*)\}_i = \mathrm{tr}\left\{\left[\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right]^+\mathbf{V}_i\right\}\,, \quad i = 1, \ldots, p\,,$$

and the (i,j)*th entry of the matrix* $\mathbf{S}_K$, $i,j = 1, \ldots, p$, *is*

$$\{\mathbf{S}_K\}_{i,j} = \mathrm{tr}\left\{\left[\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right]^+\mathbf{V}_i\left[\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right]^+\mathbf{V}_j\right\}\,.$$

Proof. Using Taylor series when the second and higher derivatives are neglected we get

$$T(\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta}) \approx T(\boldsymbol{\vartheta}^*) + \sum_{i=1}^p \frac{\partial T(\boldsymbol{\vartheta}^*)}{\partial\vartheta_i}\delta\vartheta_i\,,$$

$$\{\boldsymbol{\eta}(\boldsymbol{\vartheta}^*)\}_i = \frac{\partial T(\boldsymbol{\vartheta}^*)}{\partial\vartheta_i} = 2\mathbf{v}'(\boldsymbol{\vartheta}^*)\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\frac{\partial\mathbf{v}(\boldsymbol{\vartheta}^*)}{\partial\vartheta_i} + \mathbf{v}'(\boldsymbol{\vartheta}^*)\frac{\partial\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)}{\partial\vartheta_i}\mathbf{v}(\boldsymbol{\vartheta}^*)\,.$$

In the following text the dependence on $\boldsymbol{\vartheta}$ is not written. From equalities

$$\frac{\partial\boldsymbol{\Sigma}^{-1}}{\partial\vartheta_i} = -\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\vartheta_i}\boldsymbol{\Sigma}^{-1}\,, \quad \frac{\partial\boldsymbol{\Sigma}}{\partial\vartheta_i} = \mathbf{V}_i$$

it follows

$$\{\boldsymbol{\eta}\}_i = \frac{\partial T}{\partial \vartheta_i} = \mathbf{v}'\mathbf{A}_i\mathbf{v}\,, \quad i = 1, \dots, p\,,$$
$$\mathbf{A}_i = 2\boldsymbol{\Sigma}^{-1}\mathbf{X}(\mathbf{M}_{B'}\mathbf{C}\mathbf{M}_{B'})^{+}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{V}_i\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\mathbf{V}_i\boldsymbol{\Sigma}^{-1}\,.$$

Thus

$$\mathrm{E}\left[\{\boldsymbol{\eta}\}_i\right] = \mathrm{tr}[\mathbf{A}_i \mathrm{var}(\mathbf{v})]\,,$$
$$\mathrm{cov}\left[\{\boldsymbol{\eta}\}_i, \{\boldsymbol{\eta}\}_j\right] = \mathrm{cov}(\mathbf{v}'\mathbf{A}_i\mathbf{v}, \mathbf{v}'\mathbf{A}_j\mathbf{v}) = 2\mathrm{tr}\left[\mathbf{A}_i\mathrm{var}(\mathbf{v})\mathbf{A}_j\mathrm{var}(\mathbf{v})\right]$$

and the proof can be easily finished. □

Let $\varepsilon > 0$ be a given tolerable increase of the risk α of the test T. Let δ_ε be given as a solution of the equation

$$\mathcal{P}_{H_0}\left\{T(\boldsymbol{\vartheta}^*) + \delta_\varepsilon \geq \chi^2_{n+q-k}(1-\alpha)\right\} = \alpha + \varepsilon\,,$$

i.e.,

$$\delta_\varepsilon = \chi^2_{n+q-k}(1-\alpha) - \chi^2_{n+q-k}(1-\alpha-\varepsilon)\,.$$

The symbol $\mathcal{P}_{H_0}$ means the probability in the case that the null hypothesis H_0 is true. Let $t > 0$ be a such real number that $\mathcal{P}_{H_0}\{\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta} < \delta_\varepsilon\} \approx 1$, i.e.,

$$\mathrm{E}\left[\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right] + t\sqrt{\mathrm{var}\left[\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right]} \leq \delta_\varepsilon \tag{4}$$

where t is sufficiently large. Let

$$\mathcal{N}_\varepsilon = \left\{\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta} : \mathrm{E}\left[\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right] + t\sqrt{\mathrm{var}\left[\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right]} \leq \delta_\varepsilon\right\}$$

be the insensitivity region for the test T. Then

$$\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta} \in \mathcal{N}_\varepsilon \quad \Rightarrow \quad \mathcal{P}_{H_0}\left\{T(\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta}) \geq \chi^2_{n+q-k}(1-\alpha)\right\} \leq \alpha + \varepsilon\,.$$

Theorem 3.2. *The insensitivity region for the test T can be expressed as*

$$\mathcal{N}_\varepsilon = \Big\{\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta} :$$
$$\left[\delta\boldsymbol{\vartheta} - \delta_\varepsilon\mathbf{D}_t^{+}\mathbf{a}(\boldsymbol{\vartheta}^*)\right]'\mathbf{D}_t\left[\delta\boldsymbol{\vartheta} - \delta_\varepsilon\mathbf{D}_t^{+}\mathbf{a}(\boldsymbol{\vartheta}^*)\right] \leq \left[1 + \mathbf{a}'(\boldsymbol{\vartheta}^*)\mathbf{D}_t^{+}\mathbf{a}(\boldsymbol{\vartheta}^*)\right]\delta_\varepsilon^2\Big\}\,,$$

where

$$\mathbf{D}_t = 2t^2\mathbf{S}_K - \mathbf{a}(\boldsymbol{\vartheta}^*)\mathbf{a}'(\boldsymbol{\vartheta}^*)\,.$$

Proof. It is sufficient to check the inequality (4) from which it follows

$$t^2\mathrm{var}[\boldsymbol{\eta}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}] \leq \left[\delta_\varepsilon + \mathbf{a}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right]^2$$
$$\Rightarrow \quad \delta\boldsymbol{\vartheta}'\left(2t^2\mathbf{S}_K - \mathbf{a}(\boldsymbol{\vartheta}^*)\mathbf{a}'(\boldsymbol{\vartheta}^*)\right)\delta\boldsymbol{\vartheta} - 2\delta_\varepsilon\mathbf{a}'(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta} \leq \delta_\varepsilon^2\,.$$

Since $\mathbf{a}(\boldsymbol{\vartheta}^*) \in \mathcal{M}(\mathbf{D}_t)$ (cf. Lešanská, 2001), $\mathcal{N}_\varepsilon$ can be written as in the statement. □

Lemma 3.3. *The infinitesimal approximation of* $T_{out}(\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta})$ *is*

$$T_{out}(\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta}) \approx T_{out}(\boldsymbol{\vartheta}^*) + \boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\,,$$

where

$$\begin{aligned}\{\boldsymbol{\eta}_{out}(\boldsymbol{\vartheta}^*)\}_i &= -\widehat{\boldsymbol{\Delta}}'(\boldsymbol{\vartheta}^*)\mathbf{F}_i\widehat{\boldsymbol{\Delta}}(\boldsymbol{\vartheta}^*) - 2\widehat{\boldsymbol{\Delta}}'(\boldsymbol{\vartheta}^*)\mathbf{G}_i\mathbf{v}_{out}(\boldsymbol{\vartheta}^*)\,, \quad i = 1,\dots,p,\\ \mathbf{F}_i &= \mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right)^+\mathbf{V}_i\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right)^+\mathbf{E}\,,\\ \mathbf{G}_i &= \mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right)^+\mathbf{V}_i\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\,,\\ \mathbf{v}_{out}(\boldsymbol{\vartheta}^*) &= \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_{out}(\boldsymbol{\vartheta}^*) - \mathbf{E}\widehat{\boldsymbol{\Delta}}(\boldsymbol{\vartheta}^*)\,.\end{aligned}$$

Further

$$\begin{aligned}\mathrm{E}[\boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}] &= -\mathbf{a}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\,,\\ \mathrm{var}[\boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}] &= \delta\boldsymbol{\vartheta}'\left(4\mathbf{C}_U - 2\mathbf{S}_Z\right)\delta\boldsymbol{\vartheta}\,,\end{aligned}$$

where

$$\begin{aligned}\{\mathbf{a}_{out}(\boldsymbol{\vartheta}^*)\}_i &= \mathrm{tr}(\mathbf{Z}\mathbf{V}_i)\,, \quad i = 1,\dots,p\,,\\ \{\mathbf{C}_U\}_{i,j} &= \mathrm{tr}\left[\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right)^+\mathbf{V}_i\mathbf{Z}\mathbf{V}_j\right]\,,\\ \{\mathbf{S}_Z\}_{i,j} &= \mathrm{tr}(\mathbf{Z}\mathbf{V}_i\mathbf{Z}\mathbf{V}_j)\,, \quad i,j = 1,\dots,p\,,\end{aligned}$$

and

$$\begin{aligned}\mathbf{Z} = \left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right)^+\mathbf{E}\Big[\mathbf{E}'\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right)^+\mathbf{E}\Big]^{-1}\mathbf{E}'&\\ \times\left(\mathbf{M}_{XM_{B'}}\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\mathbf{M}_{XM_{B'}}\right)^+&\,.\end{aligned}$$

Proof. It can be proved similarly as Lemma 3.1. □

Analogously as for the test T can be stated. Let $\delta_{out,\varepsilon}$ be given by

$$\begin{aligned}&\mathcal{P}_{H_0}\left\{T_{out}(\boldsymbol{\vartheta}^*) + \delta_{out,\varepsilon} \geq \chi^2_r(1-\alpha)\right\} = \alpha + \varepsilon\\ &\Rightarrow \quad \delta_{out,\varepsilon} = \chi^2_r(1-\alpha) - \chi^2_r(1-\alpha-\varepsilon)\,.\end{aligned}$$

Let $t > 0$ be a such real number that $\mathcal{P}_{H_0}\{\boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta} < \delta_{out,\varepsilon}\} \approx 1$, i.e.,

$$\mathrm{E}\left[\boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right] + t\sqrt{\mathrm{var}\left[\boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right]} \leq \delta_{out,\varepsilon}$$

where t is sufficiently large. Let

$$\mathcal{N}_{out,\varepsilon} = \left\{\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta} : \ \mathrm{E}\left[\boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right] + t\sqrt{\mathrm{var}\left[\boldsymbol{\eta}'_{out}(\boldsymbol{\vartheta}^*)\delta\boldsymbol{\vartheta}\right]} \leq \delta_{out,\varepsilon}\right\}$$

be the insensitivity region for the test T_{out}. Then

$$\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta} \in \mathcal{N}_{out,\varepsilon} \quad \Rightarrow \quad \mathcal{P}_{H_0}\left\{T_{out}(\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta}) \geq \chi^2_r(1-\alpha)\right\} \leq \alpha + \varepsilon\,.$$

Theorem 3.4. *The insensitivity region for the test* T_{out} *can be expressed as*

$$\begin{aligned}\mathcal{N}_{out,\varepsilon} = \Big\{\boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta} : \left[\delta\boldsymbol{\vartheta} - \delta_{out,\varepsilon}\mathbf{D}^+_{out,t}\mathbf{a}_{out}(\boldsymbol{\vartheta}^*)\right]'\mathbf{D}_{out,t}\left[\delta\boldsymbol{\vartheta} - \delta_{out,\varepsilon}\mathbf{D}^+_{out,t}\mathbf{a}_{out}(\boldsymbol{\vartheta}^*)\right]&\\ \leq \left[1 + \mathbf{a}'_{out}(\boldsymbol{\vartheta}^*)\mathbf{D}^+_{out,t}\mathbf{a}_{out}(\boldsymbol{\vartheta}^*)\right]\delta^2_{out,\varepsilon}\Big\}&,\end{aligned}$$

$$\mathbf{D}_{out,t} = t^2\left(4\mathbf{C}_U - 2\mathbf{S}_Z\right) - \mathbf{a}_{out}(\boldsymbol{\vartheta}^*)\mathbf{a}'_{out}(\boldsymbol{\vartheta}^*).$$

4 Conclusion

Generally insensitivity regions $\mathcal{N}_\varepsilon$ and $\mathcal{N}_{out,\varepsilon}$ are different. Moreover, $\mathcal{N}_{out,\varepsilon}$ depends on the number of suspicious measurements.

An information on $\mathcal{N}_\varepsilon$ and $\mathcal{N}_{out,\varepsilon}$, respectively, enables us to decide whether approximations of $\boldsymbol{\vartheta}$ can be used in statistical inference or not. Some more detailed analysis and a numerical example exceed the scope of the paper, it is prepared a continuation.

Acknowledgement

This work was supported by the Council of Czech Government MSM 6 198 959 214.

References

Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*. New York: J. Wiley.

Kubáček, L., and Kubáčková, L. (2000). Nonsensitiveness regions in universal models. *Math. Slovaca*, *50*, 219-240.

Kubáček, L., Kubáčková, L., and Volaufová, J. (1995). *Statistical Models with Linear Structures*. Bratislava: Veda.

Lešanská, E. (2001). Insensitivity regions for testing statistical hypotheses in mixed models with constraints. *Mt. Math. Publ.*, *22*, 209-222.

Lešanská, E. (2002). Optimization of the size of nonsensitiveness regions. *Appl. Math.*, *47*, 9-23.

Rao, C. R., and Mitra, S. K. (1971). *Generalized Inverse of Matrices and its Applications*. New York: J. Wiley.

Scheffé, H. (1959). *The Analysis of Variance*. New York: John Wiley & Sons.

Zvára, K. (1989). *Regression Analysis*. Praha: Academia. (in Czech)

Authors' address:

Eva Fišerová and Lubomír Kubáček
Dept. of Math. Anal. and Appl. Math.
Faculty of Science
Palacký University
Tomkova 40
CZ 779 00 Olomouc
Czech Republic

E-mail: fiserova@inf.upol.cz
and kubacekl@inf.upol.cz