VERSITA

TATRA MOUNTAINS
Mathematical Publications
DOI: 10.2478/v10127-012-0004-1
Tatra Mt. Math. Publ. **51** (2012), 33–43

DECOMPOSITION OF MULTIVARIATE STATISTICAL MODELS

Eva Fišerová — Lubomír Kubáček

ABSTRACT. The paper is focused on decomposition of a multivariate model into a system of two simpler models. The multiresponses are considered to be independent with the same covariance matrix. Tests are proposed to identify which of the two models should be used in order to obtain more efficient estimators. In a case of partly known model parameters, a tolerance domain for negligible parameters is given.

1. Introduction

Multivariate statistical models, or the so-called multivariate multiple regression models, are utilized widely for studying relationships between a set of multiresponse data and a set of regressors. A multivariate view in multiresponse multiple regression situations is important since, generally, multiresponse data should be modelled jointly, see, e.g., [3], [4], [9].

In some cases, a multivariate model can be decomposed to a system of simpler models. For example, if each multiresponse is not intercorrelated, the multivariate model can be reduced to a system of independent univariate models. This situation will be discussed in detail in Section 2.

In Section 3, we propose tests for making a decision whether to use a multivariate model or its corresponding system of two simpler multivariate models in order to obtain more efficient estimators. Section 4 is devoted to a decomposition when one part of model parameters is known and, moreover, its values are small. We derive such a tolerance domain that if the parameters lie inside it is better to neglect these parameters and use a system of two simpler multivariate models. The behavior of the obtained theoretical results is studied by simulations in Section 5. A discussion concludes the paper.

2010 Mathematics Subject Classification: Primary 62H12; Secondary 62J05, 62H15.
Keywords: multivariate model, decomposition of model, unbiased estimator, test for decomposition, tolerance domain for negligible parameters.
The research was supported by the Council of Czech Government MSM 6198959214.

2. Multivariate linear model

A multivariate model for p-dimensional multiresponse data with n observations and a k-dimensional set of regressors (including an intercept) can be written as a system of n equations ($i = 1, 2, \dots, n$)

$$(Y_{i1}, Y_{i2}, \dots, Y_{ip}) = (x_{i1}, x_{i2}, \dots, x_{ik})(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_p) + (\varepsilon_{i1}, \varepsilon_{i2}, \dots, \varepsilon_{ip}),$$

or, equivalently, in a matrix form

$$\underline{\boldsymbol{Y}}_{(n\times p)} = \boldsymbol{X}_{(n\times k)} \boldsymbol{B}_{(k\times p)} + \underline{\boldsymbol{\varepsilon}}_{(n\times p)}. \tag{1}$$

Here $\underline{\boldsymbol{Y}}$ is a random matrix (observation matrix), $\boldsymbol{X}$ is a known design matrix, $\boldsymbol{B}$ is a matrix of unknown parameters and $\underline{\boldsymbol{\varepsilon}}$ is a random error matrix. We will denote a multiresponse as $\underline{\boldsymbol{Y}}_{i\cdot} = (Y_{i1}, Y_{i2}, \dots, Y_{ip})'$. In the following text we will assume that $\boldsymbol{X}$ is of full column rank and the multiresponses are independent with the same covariance matrix $\boldsymbol{\Sigma}$ which is positive definite.

If the multiresponses are not intercorrelated, i.e., $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ (identity matrix of order $p \times p$), the multivariate model (1) can be reduced into p independent univariate models

$$\boldsymbol{Y}_j = \boldsymbol{X}\boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_j, \quad \boldsymbol{Y}_j = (Y_{1j}, Y_{2j}, \dots, Y_{nj})', \qquad j = 1, 2, \dots, p.$$

Let us consider a partitioned multivariate model

$$\left(\underset{(n\times p_1)}{\underline{\boldsymbol{Y}}^1}, \underset{(n\times p_2)}{\underline{\boldsymbol{Y}}^2}\right) = \left(\underset{(n\times k_1)}{\boldsymbol{X}_1}, \underset{(n\times k_2)}{\boldsymbol{X}_2}\right)\begin{pmatrix} \underset{(k_1\times p_1)}{\boldsymbol{B}_{11}}, & \underset{(k_1\times p_2)}{\boldsymbol{B}_{12}} \\ \underset{(k_2\times p_1)}{\boldsymbol{B}_{21}}, & \underset{(k_2\times p_2)}{\boldsymbol{B}_{22}} \end{pmatrix} + \left(\underset{(n\times p_1)}{\underline{\boldsymbol{\varepsilon}}_1}, \underset{(n\times p_2)}{\underline{\boldsymbol{\varepsilon}}_2}\right). \tag{2}$$

Then, the covariance matrix $\boldsymbol{\Sigma}$ of the multiresponse $\underline{\boldsymbol{Y}}_{i\cdot}$ is partitioned in the same way, i.e.,

$$\operatorname{var}\begin{pmatrix} \underline{\boldsymbol{Y}}_{i\cdot}^1 \\ \underline{\boldsymbol{Y}}_{i\cdot}^2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{11}, & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21}, & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \qquad i = 1, 2, \dots, n. \tag{3}$$

If the vectors $\underline{\boldsymbol{Y}}_{i\cdot}^1$ and $\underline{\boldsymbol{Y}}_{i\cdot}^2$ are not correlated, i.e., $\boldsymbol{\Sigma}_{12} = \boldsymbol{0}$, $\boldsymbol{\Sigma}_{21} = \boldsymbol{0}$, the model (2) can be decomposed into two independent multivariate models

$$\underline{\boldsymbol{Y}}^1 = \boldsymbol{X}_1 \boldsymbol{B}_{11} + \boldsymbol{X}_2 \boldsymbol{B}_{21} + \underline{\boldsymbol{\varepsilon}}_1, \quad \underline{\boldsymbol{Y}}^2 = \boldsymbol{X}_1 \boldsymbol{B}_{12} + \boldsymbol{X}_2 \boldsymbol{B}_{22} + \underline{\boldsymbol{\varepsilon}}_2 .$$

If, moreover, $\boldsymbol{B}_{12} = \boldsymbol{0}$, $\boldsymbol{B}_{21} = \boldsymbol{0}$, we obtain the following two independent models

$$\underline{\boldsymbol{Y}}^1 = \boldsymbol{X}_1 \boldsymbol{B}_{11} + \underline{\boldsymbol{\varepsilon}}_1, \quad \underline{\boldsymbol{Y}}^2 = \boldsymbol{X}_2 \boldsymbol{B}_{22} + \underline{\boldsymbol{\varepsilon}}_2 . \tag{4}$$

However, in the case when $\underline{\boldsymbol{Y}}_{i\cdot}^1$ and $\underline{\boldsymbol{Y}}_{i\cdot}^2$ are correlated and $\boldsymbol{B}_{12} = \boldsymbol{0}$, $\boldsymbol{B}_{21} = \boldsymbol{0}$, the system of models (4) represents a special case of the so-called seemingly unrelated equations, or SUR models [10]. These models should be estimated together. If each model is estimated separately, estimates are consistent, although not efficient. Some special methods for estimation with explicit formulas for estimators can be found in [5].

3. Tests for decomposition of multivariate models

Let us consider a multivariate model (2) with a covariance matrix $\boldsymbol{\Sigma}$ of each multiresponse given by (3). Further, let us consider a system of two simpler multivariate models

$$\underset{(n\times p_1)}{\underline{\boldsymbol{Y}}^1} = \underset{(n\times k_1)}{\boldsymbol{X}_1}\,\underset{(k_1\times p_1)}{\boldsymbol{B}_1} + \underset{(n\times p_1)}{\underline{\boldsymbol{\varepsilon}}_1}, \qquad \underset{(n\times p_2)}{\underline{\boldsymbol{Y}}^2} = \underset{(n\times k_2)}{\boldsymbol{X}_2}\,\underset{(k_2\times p_2)}{\boldsymbol{B}_2} + \underset{(n\times p_2)}{\underline{\boldsymbol{\varepsilon}}_2} \tag{5}$$

with the same covariance matrix. The problem is to decide which of the models (2) and (5) should be chosen for modelling in order to obtain more efficient estimators.

Many statements in multivariate theory can be obtained directly from univariate theory. The multivariate model (2) can be rewritten in a suitable univariate form

$$\begin{aligned}\begin{pmatrix}\mathrm{vec}(\underline{\boldsymbol{Y}}^1)\\ \mathrm{vec}(\underline{\boldsymbol{Y}}^2)\end{pmatrix} &= \begin{pmatrix}\boldsymbol{I}_{p_1}\otimes\boldsymbol{X}_1, & \boldsymbol{0}, & \boldsymbol{I}_{p_1}\otimes\boldsymbol{X}_2, & \boldsymbol{0}\\ \boldsymbol{0}, & \boldsymbol{I}_{p_2}\otimes\boldsymbol{X}_2, & \boldsymbol{0}, & \boldsymbol{I}_{p_2}\otimes\boldsymbol{X}_1\end{pmatrix}\\ &\quad\times\left[\mathrm{vec}(\boldsymbol{B}_{11})', \mathrm{vec}(\boldsymbol{B}_{22})', \mathrm{vec}(\boldsymbol{B}_{21})', \mathrm{vec}(\boldsymbol{B}_{12})'\right]'\\ &\quad+\begin{pmatrix}\mathrm{vec}(\underline{\boldsymbol{\varepsilon}}^1)\\ \mathrm{vec}(\underline{\boldsymbol{\varepsilon}}^2)\end{pmatrix}.\end{aligned} \tag{6}$$

Here, the symbol $\mathrm{vec}(\underline{\boldsymbol{Y}}^1)$ denotes the column vector composed of the columns of $\underline{\boldsymbol{Y}}^1$. The notation $\otimes$ means the Kronecker multiplication of matrices [8]. The corresponding covariance matrix $\boldsymbol{V}$ is

$$\boldsymbol{V} = \mathrm{var}\begin{pmatrix}\mathrm{vec}(\underline{\boldsymbol{Y}}^1)\\ \mathrm{vec}(\underline{\boldsymbol{Y}}^2)\end{pmatrix} = \begin{pmatrix}\boldsymbol{\Sigma}_{11}\otimes\boldsymbol{I}_n, & \boldsymbol{\Sigma}_{12}\otimes\boldsymbol{I}_n\\ \boldsymbol{\Sigma}_{21}\otimes\boldsymbol{I}_n, & \boldsymbol{\Sigma}_{22}\otimes\boldsymbol{I}_n\end{pmatrix}. \tag{7}$$

If the covariance matrix $\boldsymbol{\Sigma}$ is known, the best linear unbiased estimators (BLUEs) of $\boldsymbol{B}_{11}$, $\boldsymbol{B}_{12}$, $\boldsymbol{B}_{21}$ and $\boldsymbol{B}_{22}$ in model (2) and their covariance matrices are given as [4]

$$\begin{aligned}\widehat{\boldsymbol{B}}_{11} &= (\boldsymbol{X}_1'\boldsymbol{M}_{X_2}\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{M}_{X_2}\underline{\boldsymbol{Y}}^1,\ \mathrm{var}\big[\mathrm{vec}(\widehat{\boldsymbol{B}}_{11})\big] = \boldsymbol{\Sigma}_{11}\otimes(\boldsymbol{X}_1'\boldsymbol{M}_{X_2}\boldsymbol{X}_1)^{-1},\\ \widehat{\boldsymbol{B}}_{12} &= (\boldsymbol{X}_1'\boldsymbol{M}_{X_2}\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{M}_{X_2}\underline{\boldsymbol{Y}}^2,\ \mathrm{var}\big[\mathrm{vec}(\widehat{\boldsymbol{B}}_{12})\big] = \boldsymbol{\Sigma}_{22}\otimes(\boldsymbol{X}_1'\boldsymbol{M}_{X_2}\boldsymbol{X}_1)^{-1},\\ \widehat{\boldsymbol{B}}_{21} &= (\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\underline{\boldsymbol{Y}}^1,\ \mathrm{var}\big[\mathrm{vec}(\widehat{\boldsymbol{B}}_{21})\big] = \boldsymbol{\Sigma}_{11}\otimes(\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1},\\ \widehat{\boldsymbol{B}}_{22} &= (\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\underline{\boldsymbol{Y}}^2,\ \mathrm{var}\big[\mathrm{vec}(\widehat{\boldsymbol{B}}_{22})\big] = \boldsymbol{\Sigma}_{22}\otimes(\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1},\end{aligned}$$

where $\boldsymbol{M}_{X_i} = \boldsymbol{I}_n - \boldsymbol{P}_{X_i}$, $\boldsymbol{P}_{X_i} = \boldsymbol{X}_i(\boldsymbol{X}_i'\boldsymbol{X}_i)^{-1}\boldsymbol{X}_i'$, $i = 1, 2$.

Similarly, we can rewrite the model (5) in the univariate form

$$\begin{pmatrix} \operatorname{vec}(\underline{\boldsymbol{Y}}^1) \\ \operatorname{vec}(\underline{\boldsymbol{Y}}^2) \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_1, & \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_2 \end{pmatrix} \begin{pmatrix} \operatorname{vec}(\boldsymbol{B}_1) \\ \operatorname{vec}(\boldsymbol{B}_2) \end{pmatrix} + \begin{pmatrix} \operatorname{vec}(\underline{\boldsymbol{\varepsilon}}^1) \\ \operatorname{vec}(\underline{\boldsymbol{\varepsilon}}^2) \end{pmatrix} \tag{8}$$

with the covariance matrix (7).

Now, applying results from univariate theory [1], [6], explicit formulas can be derived for the BLUEs of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ in model (5) of the form

$$\begin{aligned} \operatorname{vec}(\widehat{\boldsymbol{B}}_1) &= \Big[(\boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_1')\boldsymbol{U}_1^{-1}(\boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_1)\Big]^{-1}(\boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_1')\boldsymbol{U}_1^{-1} \\ &\quad \times \Big\{\operatorname{vec}(\underline{\boldsymbol{Y}}^1) - \Big[(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}) \otimes \boldsymbol{M}_{X_2}\Big]\operatorname{vec}(\underline{\boldsymbol{Y}}^2)\Big\}, \end{aligned}$$

$$\begin{aligned} \operatorname{vec}(\widehat{\boldsymbol{B}}_2) &= \Big[(\boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_2')\boldsymbol{U}_2^{-1}(\boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_2)\Big]^{-1}(\boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_2')\boldsymbol{U}_2^{-1} \\ &\quad \times \Big\{\operatorname{vec}(\underline{\boldsymbol{Y}}^2) - \Big[(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}) \otimes \boldsymbol{M}_{X_1}\Big]\operatorname{vec}(\underline{\boldsymbol{Y}}^1)\Big\}, \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{U}_1 &= \boldsymbol{\Sigma}_{11.2} \otimes \boldsymbol{I}_n + (\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}) \otimes \boldsymbol{P}_{X_2}, \\ \boldsymbol{U}_2 &= \boldsymbol{\Sigma}_{22.1} \otimes \boldsymbol{I}_n + (\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}) \otimes \boldsymbol{P}_{X_1} \end{aligned}$$

and

$$\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}, \quad \boldsymbol{\Sigma}_{22.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}.$$

The covariance matrices are

$$\begin{aligned} \operatorname{var}\Big[\operatorname{vec}(\widehat{\boldsymbol{B}}_1)\Big] &= \Big[(\boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_1')\boldsymbol{U}_1^{-1}(\boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_1)\Big]^{-1}, \\ \operatorname{var}\Big[\operatorname{vec}(\widehat{\boldsymbol{B}}_2)\Big] &= \Big[(\boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_2')\boldsymbol{U}_2^{-1}(\boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_2)\Big]^{-1}. \end{aligned}$$

Note that the number of parameters in the model (5) is less than in (2) and therefore the estimators are more efficient, particularly

$$\operatorname{var}\Big[\operatorname{vec}(\widehat{\boldsymbol{B}}_{11})\Big] \geq_{\mathrm{L}} \operatorname{var}\Big[\operatorname{vec}(\widehat{\boldsymbol{B}}_1)\Big], \quad \operatorname{var}\Big[\operatorname{vec}(\widehat{\boldsymbol{B}}_{22})\Big] \geq_{\mathrm{L}} \operatorname{var}\Big[\operatorname{vec}(\widehat{\boldsymbol{B}}_2)\Big],$$

where the symbol $\geq_L$ means the Loevner ordering.

If the parameter matrices $\boldsymbol{B}_{12}$ and $\boldsymbol{B}_{21}$ in model (2) are zeros, the parameter matrices $\boldsymbol{B}_{11}$, $\boldsymbol{B}_{22}$ in model (2) and $\boldsymbol{B}_1$, $\boldsymbol{B}_2$ in model (5), respectively, are the same, however, the estimators in model (5) are more efficient. Thus it is reasonable to neglect sufficiently small parameters $\boldsymbol{B}_{12}$ and $\boldsymbol{B}_{21}$ and test the hypothesis that "a system of simpler multivariate models (5) is a true model", i.e., test $\boldsymbol{B}_{12} = \boldsymbol{0}$ and $\boldsymbol{B}_{21} = \boldsymbol{0}$. For the sake of simplicity, test statistics based on the estimators of $\boldsymbol{B}_{12}$ and $\boldsymbol{B}_{21}$, respectively, in model (2) are used. The explicit formulas are given in the following theorem.

Theorem 1. *Let* $\mathrm{vec}(\underline{\boldsymbol{Y}})$ *be normally distributed and* $\boldsymbol{\Sigma}$ *be a known covariance matrix of multiresponse* $\underline{\boldsymbol{Y}}_{i.}$. *Then it holds that*

(1) *under* $\boldsymbol{B}_{21} = \mathbf{0}$,

$$T_{21} = \mathrm{Tr}\Big[\big(\underline{\boldsymbol{Y}}^1\big)'\boldsymbol{M}_{\mathrm{X}_1}\boldsymbol{X}_2\big(\boldsymbol{X}_2'\boldsymbol{M}_{\mathrm{X}_1}\boldsymbol{X}_2\big)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_{\mathrm{X}_1}\underline{\boldsymbol{Y}}^1\boldsymbol{\Sigma}_{11}^{-1}\Big] \sim \chi^2_{\mathrm{p_1k_2}};$$

(2) *under* $\boldsymbol{B}_{12} = \mathbf{0}$,

$$T_{12} = \mathrm{Tr}\Big[\big(\underline{\boldsymbol{Y}}^2\big)'\boldsymbol{M}_{\mathrm{X}_2}\boldsymbol{X}_1\big(\boldsymbol{X}_1'\boldsymbol{M}_{\mathrm{X}_2}\boldsymbol{X}_1\big)^{-1}\boldsymbol{X}_1'\boldsymbol{M}_{\mathrm{X}_2}\underline{\boldsymbol{Y}}^2\boldsymbol{\Sigma}_{22}^{-1}\Big] \sim \chi^2_{\mathrm{p_2k_1}}.$$

The symbol $\mathrm{Tr}(\boldsymbol{\Sigma})$ *denotes trace of the matrix* $\boldsymbol{\Sigma}$.

P r o o f. Under the null hypothesis $\boldsymbol{B}_{21} = \mathbf{0}$, the random vector

$$\mathrm{vec}\big(\widehat{\boldsymbol{B}}_{21}\big) = \Big\{\boldsymbol{I}_{p_1} \otimes \big[(\boldsymbol{X}_2\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\big]\Big\}\mathrm{vec}\big(\underline{\boldsymbol{Y}}^1\big)$$

is normally distributed as $N_{k_2p_1}\big[\mathbf{0}, \boldsymbol{\Sigma}_{11} \otimes (\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\big]$, and thus the random variable $T_{21} = \mathrm{vec}(\widehat{\boldsymbol{B}}_{21})'\big[\boldsymbol{\Sigma}_{11} \otimes (\boldsymbol{X}_2'\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}\big]^{-1}\mathrm{vec}(\widehat{\boldsymbol{B}}_{21})$ has a $\chi^2_{k_2p_1}$ distribution. According to [7], T_{21} is equivalent to

$$T_{21} = \mathrm{Tr}\Big[\widehat{\boldsymbol{B}}_{21}'\boldsymbol{X}_2'\boldsymbol{M}_{\mathrm{X}_1}\boldsymbol{X}_2\widehat{\boldsymbol{B}}_{21}\boldsymbol{\Sigma}_{11}^{-1}\Big]$$

and using the expression for the BLUE of $\boldsymbol{B}_{21}$ we obtain the first statement. Similarly, we can proceed with the test statistic T_{12}. □

Now, we can test the hypothesis $\boldsymbol{B}_{21} = \mathbf{0}$ and $\boldsymbol{B}_{12} = \mathbf{0}$ in model (2) using the statistics T_{21} and T_{12}. With respect to the Bonferroni inequality [2], if

$$T_{21} \leq \chi^2_{p_1k_2}(1-\alpha/2) \quad \text{and} \quad T_{12} \leq \chi^2_{p_2k_1}(1-\alpha/2),$$

where $\chi^2_{p_1k_2}(1-\alpha/2)$ denotes the $(1-\alpha/2)$-quantile of a $\chi^2_{p_1k_2}$ distribution, both hypotheses $\boldsymbol{B}_{21} = \mathbf{0},\ \boldsymbol{B}_{12} = \mathbf{0}$ cannot be rejected on the significance level α.

To construct the test of the hypothesis that $\boldsymbol{B}_{21} = \mathbf{0}$, $\boldsymbol{B}_{12} = \mathbf{0}$ in model (2) as a single statistic is more complicated, and therefore the matter is not considered here.

Unknown covariance matrix $\boldsymbol{\Sigma}$

Let us consider that the covariance matrix $\boldsymbol{\Sigma}$ of multiresponses $\underline{\boldsymbol{Y}}_{i.}$ is unknown. An unbiased estimator of $\boldsymbol{\Sigma}$ in model (2) is [3]

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_{11}, & \widehat{\boldsymbol{\Sigma}}_{12} \\ \widehat{\boldsymbol{\Sigma}}_{21}, & \widehat{\boldsymbol{\Sigma}}_{22} \end{pmatrix} = \frac{1}{n-k_1-k_2}\begin{pmatrix} \big(\underline{\boldsymbol{Y}}^1\big)' \\ \big(\underline{\boldsymbol{Y}}^2\big)' \end{pmatrix}\boldsymbol{M}_{(X_1,X_2)}\big(\underline{\boldsymbol{Y}}^1, \underline{\boldsymbol{Y}}^2\big).$$

With $n-k_1-k_2 \geq p_1+p_2$, $\widehat{\boldsymbol{\Sigma}}$ is nonsingular with probability 1 and $(n-k_1-k_2)\widehat{\boldsymbol{\Sigma}}$ has the Wishart distribution with $n-k_1-k_2$ degrees of freedom and with the parameters equal to the entries of the matrix $\boldsymbol{\Sigma}$. We will write $(n-k_1-k_2)\widehat{\boldsymbol{\Sigma}} \sim W_{p_1+p_2}[n-k_1-k_2, \boldsymbol{\Sigma}]$.

To construct the test of the hypothesis about $\boldsymbol{B}_{12}$ and $\boldsymbol{B}_{21}$ we can proceed similarly as with a known covariance matrix $\boldsymbol{\Sigma}$. Proper statistics F_{12} and F_{21} are derived in the following Theorem 2. If $F_{12} \leq \chi^2_{p_2k_1}(1-\alpha/2)$ and $F_{21} \leq \chi^2_{p_1k_2}(1-\alpha/2)$, both hypotheses $\boldsymbol{B}_{12} = \boldsymbol{0}$, $\boldsymbol{B}_{21} = \boldsymbol{0}$ cannot be rejected on the asymptotic significance level α.

Theorem 2. *Let* $\operatorname{vec}(\underline{\boldsymbol{Y}})$ *be normally distributed and* $\boldsymbol{\Sigma}$ *be an unknown covariance matrix of multiresponse* $\underline{\boldsymbol{Y}}_{i.}$.

(1) *Under* $\boldsymbol{B}_{21} = \boldsymbol{0}$*, the statistic* F_{21}*, given by*

$$F_{21} = -\left[n - k_1 - \frac{p_1 + k_2 + 1}{2}\right] \times \log \frac{\det\left[(\underline{\boldsymbol{Y}}^1)'\boldsymbol{M}_{(X_1,X_2)}\underline{\boldsymbol{Y}}^1\right]}{\det\left[(\underline{\boldsymbol{Y}}^1)'\boldsymbol{M}_{(X_1,X_2)}\underline{\boldsymbol{Y}}^1 + \widehat{\boldsymbol{B}}'_{21}\boldsymbol{X}'_2\boldsymbol{M}_{X_1}\boldsymbol{X}_2\widehat{\boldsymbol{B}}_{21}\right]},$$

is asymptotically distributed as $\chi^2_{k_2p_1}$.

(2) *Under* $\boldsymbol{B}_{12} = \boldsymbol{0}$*, the statistic* F_{12}*, given by*

$$F_{12} = -\left[n - k_2 - \frac{p_2 + k_1 + 1}{2}\right] \times \log \frac{\det\left[(\underline{\boldsymbol{Y}}^2)'\boldsymbol{M}_{(X_1,X_2)}\underline{\boldsymbol{Y}}^2\right]}{\det\left[(\underline{\boldsymbol{Y}}^2)'\boldsymbol{M}_{(X_1,X_2)}\underline{\boldsymbol{Y}}^2 + \widehat{\boldsymbol{B}}'_{12}\boldsymbol{X}'_1\boldsymbol{M}_{X_2}\boldsymbol{X}_1\widehat{\boldsymbol{B}}_{12}\right]},$$

is asymptotically distributed as $\chi^2_{k_1p_2}$.

Proof. If the matrix $\boldsymbol{B}_{21}$ is zero, the estimator $\operatorname{vec}(\widehat{\boldsymbol{B}}_{21})$ has k_2p_1-dimensional normal distribution with zero mean value and the covariance matrix equal to

$$\boldsymbol{\Sigma}_{11} \otimes (\boldsymbol{X}'_2\boldsymbol{M}_{X_1}\boldsymbol{X}_2)^{-1}.$$

Hence,

$$\widehat{\boldsymbol{B}}_{21}\boldsymbol{X}'_2\boldsymbol{M}_{X_1}\boldsymbol{X}_2\widehat{\boldsymbol{B}}_{21} \sim W_{p_1}(k_2, \boldsymbol{\Sigma}_{11}). \tag{9}$$

An unbiased estimator of $\boldsymbol{\Sigma}_{11}$ is

$$(\underline{\boldsymbol{Y}}^1)'\boldsymbol{M}_{(X_1,X_2)}\underline{\boldsymbol{Y}}^1/(n - k_1 - k_2)$$

and

$$\big(\underline{\boldsymbol{Y}}^1\big)'\boldsymbol{M}_{(X_1,X_2)}\underline{\boldsymbol{Y}}^1 \sim W_{p_1}(n - k_1 - k_2, \boldsymbol{\Sigma}_{11}). \tag{10}$$

Since

$$\boldsymbol{M}_{(X_1,X_2)} = \boldsymbol{M}_{X_1} - \boldsymbol{P}_{M_{X_1}X_2},$$

the matrices (9) and (10) are independent, and thus by the Wilks-Bartlett theorem ([3, p. 300]) we obtain the statistic F_{21}. The second statement can be proved analogously. □

4. Tolerance domain for negligible parameters

In this section we will consider a completely different situation. Let us assume that the multivariate model (2) is the true model and the parameter matrices $\boldsymbol{B}_{12}$ and $\boldsymbol{B}_{21}$ are known, e.g., they represent some physical or geodetical constants and, moreover, their values are small. The studied problem is whether these parameters can be neglected or not, i.e., whether the multivariate model (2) can be approximated by the system of two simpler multivariate models (5) or not.

From univariate theory [1] it is known that estimators of unknown model parameters from the underparametrized model (5) are biased in the true model (2). Let us consider the model (2) rewritten into the univariate form (6) with the known covariance matrix $\boldsymbol{V}$ given by (7). Let us denote

$$\boldsymbol{F} = \begin{pmatrix} \boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_1, & \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_2 \end{pmatrix}, \qquad \boldsymbol{S} = \begin{pmatrix} \boldsymbol{I}_{p_1} \otimes \boldsymbol{X}_2, & \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{I}_{p_2} \otimes \boldsymbol{X}_1 \end{pmatrix}$$

and $\boldsymbol{\kappa} = \big(\mathrm{vec}(\boldsymbol{B}_{21})', \mathrm{vec}(\boldsymbol{B}_{12})'\big)'$. Then, using univariate theory, the mean square error (MSE) of the BLUEs $\mathrm{vec}(\widehat{\boldsymbol{B}}_1)$, $\mathrm{vec}(\widehat{\boldsymbol{B}}_2)$ from the approximate models (5) in model (2) is [1]

$$\begin{aligned} \mathrm{MSE}\begin{pmatrix} \mathrm{vec}\big(\widehat{\boldsymbol{B}}_1\big) \\ \mathrm{vec}\big(\widehat{\boldsymbol{B}}_2\big) \end{pmatrix} &= \big(\boldsymbol{F}'\boldsymbol{V}^{-1}\boldsymbol{F}\big)^{-1} \\ &\quad + \big(\boldsymbol{F}'\boldsymbol{V}^{-1}\boldsymbol{F}\big)^{-1}\boldsymbol{F}'\boldsymbol{V}^{-1}\boldsymbol{S}\boldsymbol{\kappa}\boldsymbol{\kappa}'\boldsymbol{S}'\boldsymbol{V}^{-1}\boldsymbol{F}\big(\boldsymbol{F}'\boldsymbol{V}^{-1}\boldsymbol{F}\big)^{-1}. \end{aligned}$$

The estimators of $\boldsymbol{B}_{11}$ and $\boldsymbol{B}_{22}$ are unbiased in model (2). Thus, if

$$\mathrm{MSE}\begin{pmatrix} \mathrm{vec}\big(\widehat{\boldsymbol{B}}_1\big) \\ \mathrm{vec}\big(\widehat{\boldsymbol{B}}_2\big) \end{pmatrix} \leq_L \mathrm{var}\begin{pmatrix} \mathrm{vec}\big(\widehat{\boldsymbol{B}}_{11}\big) \\ \mathrm{vec}\big(\widehat{\boldsymbol{B}}_{22}\big) \end{pmatrix}, \tag{11}$$

the estimators of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ from the approximate model (5) are better (more efficient) than the estimators of $\boldsymbol{B}_{11}$ and $\boldsymbol{B}_{22}$ from model (2). Again applying univariate theory, the inequality (11) is true if and only if the vector $\boldsymbol{\kappa}$ is included in the tolerance domain [1]

$$\mathcal{T} = \Big\{\boldsymbol{\kappa} : \boldsymbol{\kappa}'\boldsymbol{S}'\big(\boldsymbol{M}_F\boldsymbol{V}\boldsymbol{M}_F\big)^{+}\boldsymbol{S}\boldsymbol{\kappa} \leq 1\Big\}. \tag{12}$$

Geometrically, a tolerance domain is a $(k_1p_2 + k_2p_1)$-dimensional ellipsoid. Its measure depends on the covariance matrix $\boldsymbol{\Sigma}$ of multiresponses and the smaller the standard deviation, the smaller the tolerance domain $\mathcal{T}$. Particularly, k^2-multiple of $\boldsymbol{\Sigma}$ makes homothetic change of the boundary of $\mathcal{T}$ in the ratio $1 : k$.

If the covariance matrix $\boldsymbol{\Sigma}$ of the multiresponse is unknown, one can use the empirical version of the tolerance domain with estimated covariance matrix for a raw analysis of negligible parameters.

5. Simulation study

Using simulations we will study the behavior of the proposed tests for decomposition of the multivariate model (2) based on statistics T_{12}, T_{21} (a known covariance matrix $\mathbf{\Sigma}$ of the multiresponse) and F_{12}, F_{21} (an unknown $\mathbf{\Sigma}$), respectively, for different choices of the covariance matrix, parameter matrices, and true model, i.e., the multivariate model (2) or the system of simpler multivariate models (5).

We have considered $n = 48$ observations, a multiresponse $\underline{\boldsymbol{Y}}^j_{i\cdot}$, $j = 1, 2$, with dimensions $p_1 = 3$ and $p_2 = 4$, and the number of regressors equal to $k_1 = 2$ and $k_2 = 2$. The parameter matrices have been chosen as

$$\boldsymbol{B}_1 = \begin{pmatrix} 3, & 2, & 2 \\ 2, & 3, & 3 \end{pmatrix}, \quad \boldsymbol{B}_2 = \begin{pmatrix} 2, & 4, & 4, & 1.5 \\ 4, & 2, & 4, & 4 \end{pmatrix}, \tag{13}$$

$$\boldsymbol{B}_{12} = \begin{pmatrix} 1, & 3, & 1, & 15 \\ 2, & 7, & 8, & 3 \end{pmatrix}, \quad \boldsymbol{B}_{21} = \begin{pmatrix} 4, & 4, & 1 \\ 2, & 8, & 3 \end{pmatrix}. \tag{14}$$

The observation matrices $\underline{\boldsymbol{Y}}^1$ and $\underline{\boldsymbol{Y}}^2$ were generated in a natural way, a normally distributed error term was added to the true mean. The multiresponses were independent and each had the same covariance matrix $\mathbf{\Sigma}$ considered in the following six forms:

$$\mathbf{\Sigma}_1 = \boldsymbol{I}_7; \qquad \mathbf{\Sigma}_2 = 100\boldsymbol{I}_7; \quad \mathbf{\Sigma}_3$$

was partitioned as

$$\mathbf{\Sigma}_{11} = \boldsymbol{I}_3, \qquad \mathbf{\Sigma}_{22} = \boldsymbol{I}_4$$

and

$$\mathbf{\Sigma}_{12} = \big(0.6\boldsymbol{I}_3, \mathbf{0}_{(3\times 1)}\big); \quad \mathbf{\Sigma}_4 = 100\mathbf{\Sigma}_3; \quad \mathbf{\Sigma}_5$$

was decomposed as

$$\mathbf{\Sigma}_{11} = \begin{pmatrix} 1 & 0.35 & 0.14 \\ 0.35 & 2 & 0.16 \\ 0.14 & 0.16 & 2 \end{pmatrix}, \qquad \mathbf{\Sigma}_{22} = \begin{pmatrix} 8 & 0.7 & 0.35 & 0.71 \\ 0.7 & 1 & 0.46 & 0.3 \\ 0.35 & 0.46 & 2 & 0.3 \\ 0.71 & 0.3 & 0.3 & 4 \end{pmatrix},$$

$$\mathbf{\Sigma}_{12} = \begin{pmatrix} 0.8 & 0.4 & 0.7 & 0.6 \\ 0.12 & 0.3 & 0.3 & 0.05 \\ 0.25 & 0.56 & 0.1 & 0.42 \end{pmatrix};$$

and $\mathbf{\Sigma}_6 = 100\mathbf{\Sigma}_5$.

50 000 simulations was done for covariance matrix $\mathbf{\Sigma}$. Then, all simulated observation matrices $\underline{\boldsymbol{Y}}^1$ and $\underline{\boldsymbol{Y}}^2$ were used in the two proposed tests for decomposition of the model (T_{12}, T_{21} for a known $\mathbf{\Sigma}$, and F_{12}, F_{21} for an estimated $\mathbf{\Sigma}$).

First, data were simulated from the system of two simpler models (5), i.e., for matrices $\boldsymbol{B}_1$, $\boldsymbol{B}_2$ given by (13) and for zero matrices $\boldsymbol{B}_{12}$, $\boldsymbol{B}_{21}$. The obtained results are shown in Tables 1 and 2. We can see that both proposed tests are conservative. The true model (5) was only rejected in 2.5 % (0.5 %) of cases for the significance level $\alpha = 5\,\%$ ($\alpha = 1\,\%$) in case of a known covariance matrix $\boldsymbol{\Sigma}$, and in 1 % (0.1 %) of cases for an estimated $\boldsymbol{\Sigma}$. On the significance level $\alpha = 5\,\%$ ($\alpha = 1\,\%$), the test based on statistics T_{12} and T_{21} distinguished the correct model in 97.5 % (99.5 %) cases; the test based on statistics F_{12} and F_{21} distinguished the correct model in 99 % (99.9 %) cases.

Table 1. Empirical probabilities (in %) of rejecting the hypothesis "true model is the system of two simpler models (5)" on the significance level α. Data simulated from model (5), used statistics T_{12}, T_{21} for a known $\boldsymbol{\Sigma}$.

Parameter matrices	α	$\boldsymbol{\Sigma}_1$	$\boldsymbol{\Sigma}_2$	$\boldsymbol{\Sigma}_3$	$\boldsymbol{\Sigma}_4$	$\boldsymbol{\Sigma}_5$	$\boldsymbol{\Sigma}_6$
(13)	5	2.61	2.47	2.46	2.57	2.46	2.51
(13)	1	0.52	0.47	0.49	0.51	0.52	0.48
$100 \cdot (13)$	5	2.46	2.56	2.41	2.45	2.47	2.52
$100 \cdot (13)$	1	0.52	0.50	0.46	0.50	0.54	0.49

Table 2. Empirical probabilities (in %) of rejecting the hypothesis "true model is the system of two simpler models (5)" on the significance level α. Data simulated from model (5), used statistics F_{12}, F_{21} for an estimated $\boldsymbol{\Sigma}$.

Parameter matrices	α	$\boldsymbol{\Sigma}_1$	$\boldsymbol{\Sigma}_2$	$\boldsymbol{\Sigma}_3$	$\boldsymbol{\Sigma}_4$	$\boldsymbol{\Sigma}_5$	$\boldsymbol{\Sigma}_6$
(13)	5	1.08	1.03	0.99	1.03	0.96	1.03
(13)	1	0.10	0.10	0.12	0.08	0.10	0.10
$100 \cdot (13)$	5	1.03	1.05	1.02	1.02	1.03	0.99
$100 \cdot (13)$	1	0.12	0.13	0.10	0.12	0.13	0.12

Results for data simulated from the model (2), i.e., for matrices $\boldsymbol{B}_1, \boldsymbol{B}_2$ and $\boldsymbol{B}_{12}$, $\boldsymbol{B}_{21}$ given by (13) and (14), respectively, are presented in Tables 3 and 4. We can see that both proposed tests are sensitive to the relative accuracy of observations (whether the tested parameters $\boldsymbol{B}_{12}$ and $\boldsymbol{B}_{21}$ are estimated with sufficient precision or not). If we have relatively very precise observations, tests always recognized the correct model (2). However, when a relative precision of observations decreases, the probability of rejecting the decomposition of model (2) into incorrect model (5) also decreases.

Table 3. Empirical probabilities (in %) of rejecting the hypothesis "true model is the system of two simpler models (5)" on the significance level α. Data simulated from model (2), used statistics T_{12}, T_{21} for a known $\mathbf{\Sigma}$.

Parameter matrices	α	$\mathbf{\Sigma}_1$	$\mathbf{\Sigma}_2$	$\mathbf{\Sigma}_3$	$\mathbf{\Sigma}_4$	$\mathbf{\Sigma}_5$	$\mathbf{\Sigma}_6$
(13), (14)	5	100	64.96	100	65.51	100	30.04
(13), (14)	1	100	43.76	100	43.60	100	14.38
$100\cdot(13)$, $100\cdot(14)$	5	100	100	100	100	100	100
$100\cdot(13)$, $100\cdot(14)$	1	100	100	100	100	100	100
$100\cdot(13)$, (14)	5	100	65.17	100	65.00	100	30.12
$100\cdot(13)$, (14)	1	100	43.37	100	43.86	100	13.92
(13), $100\cdot(14)$	5	100	100	100	100	100	100
(13), $100\cdot(14)$	1	100	100	100	100	100	100

Table 4. Empirical probabilities (in %) of rejecting the hypothesis "true model is the system of two simpler models (5)" on the significance level α. Data simulated from model (2), used statistics F_{12}, F_{21} for an estimated $\mathbf{\Sigma}$.

Parameter matrices	α	$\mathbf{\Sigma}_1$	$\mathbf{\Sigma}_2$	$\mathbf{\Sigma}_3$	$\mathbf{\Sigma}_4$	$\mathbf{\Sigma}_5$	$\mathbf{\Sigma}_6$
(13), (14)	5	100	51.24	100	51.93	100	19.32
(13), (14)	1	100	25.69	100	25.67	100	6.04
$100\cdot(13)$, $100\cdot(14)$	5	100	100	100	100	100	100
$100\cdot(13)$, $100\cdot(14)$	1	100	100	100	100	100	100
$100\cdot(13)$, (14)	5	100	51.78	100	51.29	100	19.47
$100\cdot(13)$, (14)	1	100	25.52	100	25.59	100	5.86
(13), $100\cdot(14)$	5	100	100	100	100	100	100
(13), $100\cdot(14)$	1	100	100	100	100	100	100

6. Conclusions

The proposed tests and tolerance domain for negligible parameters seem to be proper methods for decomposition of multivariate models. Both methods are valid only for independent multiresponses with the same covariance matrix. The methodology for different forms of covariance matrix or different types of multivariate models is similar; however, the explicit formulas for estimators and their characteristics require tedious and complicated computations. Therefore, we leave these topics for future research.

Acknowledgements. The authors are grateful to the referee for the comments, which helped to greatly improve the paper.

REFERENCES

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

[2] HUMAK, K. M. S.: *Statistische Methoden der Modellbildung, Band 1*. Akademie-Verlag, Berlin, 1977.

[3] KSHIRSAGAR, A. M.: *Multivariate Analysis*. Marcel Dekker, Inc., New York, 1972.

[4] KUBÁČEK, L.: *Multivariate Statistical Models Revisited*. Palacký University, Olomouc, 2008.

[5] KUBÁČEK, L.: *Seemingly Unrelated Regression Models*, Appl. Math. (to appear).

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

[7] RAO, C. R.: *Linear Statistical Inference and Its Applications*. John Wiley & Sons, Inc., New York, 1965.

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

[9] SEBER, G. A. F.: *Multivariate Observations*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2004.

[10] ZELLNER, A.: *An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias*, J. Amer. Statist. Assoc. **57** (1962), 348–368.

Received October 31, 2011

Department of Mathematical Analysis and Applications of Mathematics
Faculty of Science
Palacký University
17. listopadu 12
CZ–771-46 Olomouc
CZECH REPUBLIC

E-mail: eva.fiserova@upol.cz
lubomir.kubacek@upol.cz