ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
2008

ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
MATHEMATICA 47 (2008)

MATHEMATICA 47

PALACKÝ UNIVERSITY OLOMOUC
OLOMOUC 2008

ISBN 978-80-244-2121-6
ISSN 0231-9721

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008)

CONTENTS

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 7–22

Insensitivity Region for Variance Components in General Linear Model

Hana BOHÁČOVÁ

Department of Mathematics
Faculty of Economics and Administration, University of Pardubice
Studentská 84, 532 10 Pardubice, Czech Republic
e-mail: bohacova@upce.cz

(Received January 29, 2008)

Abstract

In linear regression models the estimator of variance components needs a suitable choice of a starting point for an iterative procedure for a determination of the estimate. The aim of this paper is to find a criterion for a decision whether a linear regression model enables to determine the estimate reasonably and whether it is possible to do so when using the given data.

Key words: Linear regression model; variance components; insensitivity region.

2000 Mathematics Subject Classification: 62J05, 62F10

1 Notation

$\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$	the n-dimensional random vector $\mathbf{Y}$ possesses the normal distribution with the mean value $\mathbf{X}\boldsymbol{\beta}$ and variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$
$\{\mathbf{A}\}_{i,j}$	the component of matrix $\mathbf{A}$ on its (i,j)-th position
$r(\mathbf{A})$	the rank of the matrix $\mathbf{A}$
$\mathrm{Tr}(\mathbf{A})$	the trace of the square matrix $\mathbf{A}$, $\mathrm{Tr}(\mathbf{A}) = \sum_i \{\mathbf{A}\}_{i,i}$
$\mathbf{A}^+$	the Moore–Penrose generalized matrix inverse (see [4] for more details)
$\mathbf{M}_{\mathbf{A}}$	the projection matrix on an orthogonal complement (in Euclidean sense) of the column space of the matrix $\mathbf{A}$
$\left.\frac{\partial \mathbf{A}}{\partial \mathbf{t}}\right\vert_{\mathbf{t}=\mathbf{t}_0}$	the value of the partial derivative of the matrix $\mathbf{A}$ according to $\mathbf{t}$ for $\mathbf{t} = \mathbf{t}_0$

2 Introduction

The main aim of this paper is to describe how the variance components estimates in a linear regression model depend on small input prior variance components values changes. We need some input prior values of the variance components when computing their minimum norm quadratic unbiased estimators (MINQUE). The question is how to get these prior values and whether the choice is suitable. We can figure out the variances of the estimates based on the given prior values and then investigate how these variances change when using different prior values. Having the estimates variances not too high seems to be a comprehensible requirement. So the task now is to find a set of admissible changes of the input variance components values (for the given variance components prior values), it means a set of such changes of the input values which cause ε-multiple increase of the estimates variances at the most.

3 General linear regression model

Let's consider following regression model (according to [5], page 62):

$$\mathbf{Y} \sim N_n\left(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}\right). \tag{1}$$

Suppose that $n \times k$ matrix $\mathbf{X}$ is known and of full column rank $r(\mathbf{X}) = k$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)'$ is a vector of unknown fixed effects parameters and the variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ satisfies

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \sum_{i=1}^{r} \theta_i \mathbf{V}_i. \tag{2}$$

$\theta_1, \theta_2, \ldots, \theta_r$ in (2) are unknown variance components (the object of our interest) and $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_r$ are known symmetrical matrices. We suppose $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is positive definite. No restrictions such as $\theta_i \geq 0$ or $\mathbf{V}_i$ positive semidefinite need hold.

4 Variance components insensitivity region

4.1 Variance components estimator

According to [4], page 101 the $\boldsymbol{\theta}_0$-MINQUE (it means the minimum norm quadratic unbiased estimator with prior variance components values $\boldsymbol{\theta}_0$) of variance components $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_r)'$ in model (1) is

$$\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) = \mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+} \begin{pmatrix} \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_1(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \\ \vdots \\ \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_r(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \end{pmatrix} \tag{3}$$

and the variance-covariance matrix of the variance components estimates $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$ is

$$\mathrm{Var}_{\boldsymbol{\theta}_0}\left(\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)\right) = 2\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+} \tag{4}$$

where (see [4], page 171)

$$(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}\mathbf{M_X})^{+} = {\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}}^{-1} - {\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}}^{-1}\mathbf{X}(\mathbf{X}'{\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}}^{-1}\mathbf{X})^{-1}\mathbf{X}'{\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}}^{-1} \tag{5}$$

and $\mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}\mathbf{M_X})^{+}}$ is matrix with

$$\left\{\mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}\mathbf{M_X})^{+}}\right\}_{i,j} = \mathrm{Tr}\left[\mathbf{V}_i(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}\mathbf{M_X})^{+}\mathbf{V}_j(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta_0}}\mathbf{M_X})^{+}\right]$$

on its (i,j)-th position.

In practice in the first step the value $\boldsymbol{\theta}_0$ in (3) can be chosen arbitrarily. In the second step the value of $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$ is chosen instead of $\boldsymbol{\theta}_0$. In this way we can procede and stop the iterative procedure after a suitable number of the steps. The problem is to recognize whether the choice of the starting value $\boldsymbol{\theta}_0$ is sufficient for deriving a reasonable estimate of $\boldsymbol{\theta}$ and whether it is sufficient for stopping the iterative procedure already after the first step. This problem can be solved as follows.

It seems to be comprehensible to have the variances of the θ_i estimators not too high. Let's use some given linear combination of the components of vector $\boldsymbol{\theta}$. Suppose that the coeficients of this linear combination are the components of vector $\mathbf{g}$. We will investigate the variance of the estimator of $\mathbf{g}'\boldsymbol{\theta}$ instead of variances of all the variance components separately.

Remark 4.1 We achieve the equality $\mathbf{g}'\boldsymbol{\theta} = \theta_i$ when using the i-th unit vector for $\mathbf{g}$. It means we still have the possiblity to take the variance of the estimator of each of the variance components under control and moreover we can monitor the variances of different linear combinations of the variance components.

As we know the $\boldsymbol{\theta}$ estimator $\widehat{\boldsymbol{\theta}}$ depends on the prior input value $\boldsymbol{\theta}_0$, it is $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$. Next we find out the difference between the variance of $\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$ and the variance of $\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta})$. A set $\mathcal{N}_{\mathbf{g},\boldsymbol{\theta}_0}$ can be found such that $\boldsymbol{\theta}_0+\delta\boldsymbol{\theta} \in \mathcal{N}_{\mathbf{g},\boldsymbol{\theta}_0}$ leads to the inequality

$$\sqrt{\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta})\right]} \leq (1+\varepsilon)\sqrt{\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)\right]}, \tag{6}$$

where $\varepsilon > 0$ is a sufficiently small real number.

We are looking after a set of $\delta\boldsymbol{\theta}$—small changes of input variance components values—holding (6) for a given r-dimensional vector $\mathbf{g}$ and given $\boldsymbol{\theta}_0$ in what follows.

In order to find such a set we need to express $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta})$. We can approximate it like this

$$\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}) \approx \widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) + \sum_{i=1}^{r} \left.\frac{\partial\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \cdot \delta\theta_i. \tag{7}$$

The appropriate linear combination $\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta})$ of variance components estimator fulfils

$$\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}) \approx \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) + \sum_{i=1}^{r} \left.\frac{\partial\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \cdot \delta\theta_i. \tag{8}$$

Thus we need to know the partial derivative $\frac{\partial\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}$.

At first let us find the first derivative of $\left\{\mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{M_X})^+}\right\}_{i,j}$ according to θ_k:

$$\left\{\frac{\partial\mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{M_X})^+}}{\partial\theta_k}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right\}_{i,j} = \frac{\partial}{\partial\theta_k}\operatorname{Tr}\left[\mathbf{V}_i(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{M_X})^+\mathbf{V}_j(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{M_X})^+\right]\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$= \operatorname{Tr}\Big[\mathbf{V}_i(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{M_X}\mathbf{V}_k\mathbf{M_X}(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_j(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+$$

$$+ \mathbf{V}_i(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_j(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{M_X}\mathbf{V}_k\mathbf{M_X}(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\Big]$$

$$= 2\operatorname{Tr}\left[\mathbf{V}_i(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{M_X}\mathbf{V}_k\mathbf{M_X}(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_j(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\right].$$

If we denote matrix having

$$\operatorname{Tr}(\mathbf{V}_i\mathbf{A}\mathbf{V}_j\mathbf{B}), \tag{9}$$

on its (i,j)-th position with $\mathbf{C}_{\mathbf{A},\mathbf{B}}$ (for arbitrary matrices $\mathbf{A}$, $\mathbf{B}$ of $n\times n$ dimension), we can continue as follows

$$\begin{aligned}\frac{\partial\mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{M_X})^+}}{\partial\theta_k}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} &= 2\mathbf{C}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{M_X}\mathbf{V}_k\mathbf{M_X}(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+,(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\\ &= 2\mathbf{C}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+,(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}.\end{aligned}$$

Let's introduce following notation

$$\boldsymbol{\gamma} = \begin{pmatrix}\mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_1(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y}\\ \vdots\\ \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_r(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y}\end{pmatrix}.$$

Now we can write

$$\frac{\partial\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial\theta_k}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -2\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

$$\times\,\mathbf{C}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+,(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\boldsymbol{\gamma}$$

$$+\,\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

$$\times\begin{pmatrix}2\mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_1(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y}\\ \vdots\\ 2\mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_r(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y}\end{pmatrix}. \tag{10}$$

According to (10) we have

$$\frac{\partial\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial\theta_k}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -2\mathbf{g}'\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

$$\times \mathbf{C}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+,(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\boldsymbol{\gamma}$$

$$+ 2\mathbf{g}'\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

$$\times \begin{pmatrix} \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_1(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \\ \vdots \\ \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_r(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \end{pmatrix}. \tag{11}$$

If we denote

$$\mathbf{a}'_k = \mathbf{g}'\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

$$\times \mathbf{C}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+,(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+},$$

$$\mathbf{b}'_{\mathbf{g}} = \mathbf{g}'\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

and

$$\boldsymbol{\zeta}_k = \begin{pmatrix} \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_1(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \\ \vdots \\ \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_r(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \end{pmatrix},$$

we can write

$$\left.\frac{\partial \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -2\begin{pmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_r \end{pmatrix}\boldsymbol{\gamma} + 2\begin{pmatrix} \mathbf{b}'_{\mathbf{g}}\boldsymbol{\zeta}_1 \\ \vdots \\ \mathbf{b}'_{\mathbf{g}}\boldsymbol{\zeta}_r \end{pmatrix}. \tag{12}$$

Using (8) together with (12) we get

$$\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta}) \approx \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) + (\delta\boldsymbol{\theta})'\left[-2\begin{pmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_r \end{pmatrix}\boldsymbol{\gamma} + 2\begin{pmatrix} \mathbf{b}'_{\mathbf{g}}\boldsymbol{\zeta}_1 \\ \vdots \\ \mathbf{b}'_{\mathbf{g}}\boldsymbol{\zeta}_r \end{pmatrix}\right]. \tag{13}$$

4.2 $\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta})\right]$ derivation

In order to find a set of $\delta\boldsymbol{\theta}$ which is described in (6) we have to derive

$$\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta})\right].$$

According to (4)

$$\mathrm{Var}_{\boldsymbol{\theta}_0}\left[\mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)\right] = 2\mathbf{g}'\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\mathbf{g}. \tag{14}$$

What we need to know further is $\mathrm{Var}_{\boldsymbol{\theta}_0} \frac{\partial \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$.

$$\mathrm{Var}_{\boldsymbol{\theta}_0} \frac{\partial \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathrm{Var}_{\boldsymbol{\theta}_0} \left[-2 \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix} \boldsymbol{\gamma} + 2 \begin{pmatrix} \mathbf{b_g}'\boldsymbol{\zeta}_1 \\ \vdots \\ \mathbf{b_g}'\boldsymbol{\zeta}_r \end{pmatrix} \right]$$

$$= 4 \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix} \mathrm{Var}_{\boldsymbol{\theta}_0}(\boldsymbol{\gamma})\,(\mathbf{a}_1, \ldots, \mathbf{a}_r) - 4 \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix} \mathrm{cov}_{\boldsymbol{\theta}_0} \left[\boldsymbol{\gamma}, \begin{pmatrix} \mathbf{b_g}'\boldsymbol{\zeta}_1 \\ \vdots \\ \mathbf{b_g}'\boldsymbol{\zeta}_r \end{pmatrix} \right]$$

$$- 4\,\mathrm{cov}_{\boldsymbol{\theta}_0} \left[\begin{pmatrix} \mathbf{b_g}'\boldsymbol{\zeta}_1 \\ \vdots \\ \mathbf{b_g}'\boldsymbol{\zeta}_r \end{pmatrix}, \boldsymbol{\gamma} \right] (\mathbf{a}_1, \ldots, \mathbf{a}_r) + 4\,\mathrm{Var}_{\boldsymbol{\theta}_0} \begin{pmatrix} \mathbf{b_g}'\boldsymbol{\zeta}_1 \\ \vdots \\ \mathbf{b_g}'\boldsymbol{\zeta}_r \end{pmatrix}. \tag{15}$$

In view of (4) and of the definition of $\boldsymbol{\gamma}$

$$\mathrm{Var}_{\boldsymbol{\theta}_0}(\boldsymbol{\gamma}) = 2\mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}. \tag{16}$$

Concerning $\mathrm{Var}_{\boldsymbol{\theta}_0} \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix}$:

$$\left\{ \mathrm{Var}_{\boldsymbol{\theta}_0} \begin{pmatrix} \mathbf{b_g}'\boldsymbol{\zeta}_1 \\ \vdots \\ \mathbf{b_g}'\boldsymbol{\zeta}_r \end{pmatrix} \right\}_{k,l} = \mathrm{cov}_{\boldsymbol{\theta}_0} \big(\mathbf{b_g}'\boldsymbol{\zeta}_k, \mathbf{b_g}'\boldsymbol{\zeta}_l\big) = \mathbf{b_g}'\,\mathrm{cov}_{\boldsymbol{\theta}_0}\big(\boldsymbol{\zeta}_k, \boldsymbol{\zeta}_l\big)\mathbf{b_g}$$

$$= \mathbf{b_g}'\,\mathrm{cov}_{\boldsymbol{\theta}_0} \left[\begin{pmatrix} \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_1(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \\ \vdots \\ \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_k(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_r(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \end{pmatrix}, \right.$$

$$\left. \begin{pmatrix} \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_l(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_1(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \\ \vdots \\ \mathbf{Y}'(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_l(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V}_r(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{Y} \end{pmatrix} \right] \mathbf{b_g}$$

$$= \mathbf{b_g}'\mathbf{D}_{\boldsymbol{\zeta}_k, \boldsymbol{\zeta}_l}\mathbf{b_g}, \tag{17}$$

where

$$\left\{\mathbf{D}_{\boldsymbol{\zeta}_k,\boldsymbol{\zeta}_l}\right\}_{s,t} = \mathrm{cov}_{\boldsymbol{\theta}_0}\left[\mathbf{Y}'(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_k(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_s(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{Y},\right.$$

$$\left.\mathbf{Y}'(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_l(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_t(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{Y}\right]$$

$$= 2\,\mathrm{Tr}\left[(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_k(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_s(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\right.$$

$$\left.\times\,(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_l(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_t(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\right]$$

$$= 2\,\mathrm{Tr}\left[\mathbf{V}_s(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_l(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\right.$$

$$\left.\times\,\mathbf{V}_t(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_k(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\right].$$

Next

$$\mathrm{cov}_{\boldsymbol{\theta}_0}\left[-2\begin{pmatrix}\mathbf{a}_1'\\ \vdots\\ \mathbf{a}_r'\end{pmatrix}\boldsymbol{\gamma}, 2\begin{pmatrix}\mathbf{b_g}'\boldsymbol{\zeta}_1\\ \vdots\\ \mathbf{b_g}'\boldsymbol{\zeta}_r\end{pmatrix}\right] =$$

$$= -4\begin{pmatrix}\mathbf{a}_1'\\ \vdots\\ \mathbf{a}_r'\end{pmatrix}\mathrm{cov}_{\boldsymbol{\theta}_0}\left[\boldsymbol{\gamma}, \begin{pmatrix}\mathbf{b_g}'\boldsymbol{\zeta}_1\\ \vdots\\ \mathbf{b_g}'\boldsymbol{\zeta}_r\end{pmatrix}\right] = -4\begin{pmatrix}\mathbf{a}_1'\\ \vdots\\ \mathbf{a}_r'\end{pmatrix}\mathbf{D}_{\boldsymbol{\gamma},\boldsymbol{\zeta}}, \tag{18}$$

where

$$\left\{\mathbf{D}_{\boldsymbol{\gamma},\boldsymbol{\zeta}}\right\}_{i,j} = \mathrm{cov}_{\boldsymbol{\theta}_0}\left(\gamma_i, \mathbf{b_g}'\boldsymbol{\zeta}_j\right)$$

$$= \mathrm{cov}_{\boldsymbol{\theta}_0}\left[\mathbf{Y}'(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_i(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{Y}, \sum_{u=1}^{r}\mathbf{b}_{\mathbf{g}_u}\mathbf{Y}'(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_j\right.$$

$$\left.\times\,(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_u(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{Y}\right] = \sum_{u=1}^{r} 2\,\mathrm{Tr}\left[(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_i\right.$$

$$\times\,(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{b}_{\mathbf{g}_u}(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_j(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+$$

$$\left.\times\,\mathbf{V}_u(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\right]$$

$$= 2\,\mathrm{Tr}\left[\mathbf{V}_i(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_j(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\sum_{u=1}^{r}\mathbf{b}_{\mathbf{g}_u}\mathbf{V}_u(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\right]$$

$$= 2\,\mathrm{Tr}\left[\mathbf{V}_i(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_j(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\mathbf{V}_\mathbf{g}(\mathbf{M}_\mathbf{X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M}_\mathbf{X})^+\right].$$

In the previous text following important fact was used (together with equality (5)).

Lemma 4.1 *(see [4], page 101) Let $n \times n$ matrices $\mathbf{A}$, $\mathbf{B}$ be symmetrical. Let $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$, where $\mathbf{AX} = \mathbf{BX} = \mathbf{0}$. Then*

$$\mathrm{cov}\left(\mathbf{Y}'\mathbf{A}\mathbf{Y}, \mathbf{Y}'\mathbf{B}\mathbf{Y}\right) = 2\,\mathrm{Tr}\left(\mathbf{A}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\mathbf{B}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\right).$$

Following notation was used as well:

$\boldsymbol{\gamma}_i$ … i-th component of $\boldsymbol{\gamma}$, $\mathbf{b}_{\mathbf{g}u}$ … u-th component of $\mathbf{b}_{\mathbf{g}}$, $\mathbf{V}_{\mathbf{g}} = \sum_{u=1}^{r} \mathbf{b}_{\mathbf{g}u}\mathbf{V}_u$.

According to (15), (16), (17) and (18) we have

$$\mathbf{W}_{\mathbf{g}} = \mathrm{Var}_{\boldsymbol{\theta}_0} \frac{\partial \mathbf{g}'\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 8 \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix} \mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+} (\mathbf{a}_1, \dots, \mathbf{a}_r)$$

$$+ 4\mathbf{b}_{\mathbf{g}}' \begin{pmatrix} \mathbf{D}_{\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_1} & \dots & \mathbf{D}_{\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_r} \\ \vdots & \vdots & \vdots \\ \mathbf{D}_{\boldsymbol{\zeta}_r,\boldsymbol{\zeta}_1} & \dots & \mathbf{D}_{\boldsymbol{\zeta}_r,\boldsymbol{\zeta}_r} \end{pmatrix} \mathbf{b}_{\mathbf{g}} - 4 \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix} \mathbf{D}_{\boldsymbol{\gamma},\boldsymbol{\zeta}} - 4\mathbf{D}_{\boldsymbol{\gamma},\boldsymbol{\zeta}}' (\mathbf{a}_1, \dots, \mathbf{a}_r)$$

$$= 8 \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix} \mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+} (\mathbf{a}_1, \dots, \mathbf{a}_r) + 8\mathbf{S}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V_g}(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

$$- 8 \begin{pmatrix} \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_r' \end{pmatrix} \mathbf{C}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+,(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V_g}(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}$$

$$- 8\mathbf{C}'_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+,(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+\mathbf{V_g}(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+} (\mathbf{a}_1, \dots, \mathbf{a}_r). \quad (19)$$

Notation defined in (9) and denoting of matrix having on its (i, j)-th position $\mathrm{Tr}\,(\mathbf{AV}_i\mathbf{AV}_j)$ with $S_{\mathbf{A}}$ was used.

4.2.1 Insenstitivity region formulation

If we determine the insensitivity region for the variance components as a set of all $\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta}$ with $\delta\boldsymbol{\theta}$ satisfying (6), we get

$$\mathcal{N}_{\mathbf{g},\boldsymbol{\theta}_0} = \left\{\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \delta\boldsymbol{\theta}'\mathbf{W}_{\mathbf{g}}\delta\boldsymbol{\theta} \le 2\varepsilon\,\mathrm{Var}_{\boldsymbol{\theta}_0}\left(\mathbf{g}'\widehat{\boldsymbol{\theta}}\right)\right\}$$

$$= \left\{\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \delta\boldsymbol{\theta}'\mathbf{W}_{\mathbf{g}}\delta\boldsymbol{\theta} \le 4\varepsilon\mathbf{g}'\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+}\mathbf{g}\right\}. \quad (20)$$

Remark 4.2 More precise form of $\mathcal{N}_{\mathbf{g},\boldsymbol{\theta}_0}$ is (see also [4] and [1])

$$\mathcal{N}_{\mathbf{g},\boldsymbol{\theta}_0} = \left\{\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \delta\boldsymbol{\theta}'\mathbf{W}_{\mathbf{g}}\delta\boldsymbol{\theta} \le (2\varepsilon + \varepsilon^2)\,\mathrm{Var}\left(\mathbf{g}'\widehat{\boldsymbol{\theta}}\right)\right\}.$$

Because ε is choiced to be a small positive number we usually can use 2ε instead of $(2\varepsilon + \varepsilon^2)$ as used in (20).

5 Numerical study

Let the Michaelis–Menten regression function

$$f(x) = \frac{\gamma_1 x}{\gamma_2 + x}$$

be considered. We can measure values of this function for $x_1 = 0.5$; $x_2 = 1.5$ with dispersion of $\sigma_1^2 = 0.04$ and for $x_3 = 7$; $x_4 = 9$ with dispersion of $\sigma_2^2 = 0.36$. Let's suppose the true values of γ_1 and γ_2 are $\gamma_1 = 10$ and $\gamma_2 = 5$. The values of σ_1^2, σ_2^2, γ_1 and γ_2 are apriori unknown for us. We have two measurements for each point x_1, x_2, x_3, x_4. Let's include the measured data into an observation vector

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_8 \end{pmatrix}.$$

The normal distribution of the random vector $\mathbf{Y}$ is assumed. We need to describe such a nonlinear situation with a linear model. For $f(x_i) = \frac{\gamma_1 x_i}{\gamma_2 + x_i}$, $i = 1, 2, 3, 4$ we have

$$\frac{\partial f(x_i)}{\partial \gamma_1} = \frac{x_i}{\gamma_2 + x_i}, \qquad \frac{\partial f(x_i)}{\partial \gamma_2} = -\frac{\gamma_1 x_i}{(\gamma_2 + x_i)^2}.$$

Since

$$f(x_i, \gamma_1, \gamma_2) \approx f(x_i, \gamma_1^0, \gamma_2^0) + \frac{\partial f}{\partial \gamma_1}(\gamma_1 - \gamma_1^0) + \frac{\partial f}{\partial \gamma_2}(\gamma_2 - \gamma_2^0),$$

we can write:

$$\mathbf{Y} - \mathbf{f}_{\gamma_1^0, \gamma_2^0} \sim N_8 \left[\mathbf{X} \begin{pmatrix} \delta\gamma_1 \\ \delta\gamma_2 \end{pmatrix}, \mathbf{\Sigma}_{\boldsymbol{\theta}} \right]. \tag{21}$$

Here

$$\mathbf{f}_{\gamma_1^0, \gamma_2^0} = \begin{pmatrix} \frac{\gamma_1^0 x_1}{\gamma_2^0 + x_1} \\ \frac{\gamma_1^0 x_1}{\gamma_2^0 + x_1} \\ \frac{\gamma_1^0 x_2}{\gamma_2^0 + x_2} \\ \frac{\gamma_1^0 x_2}{\gamma_2^0 + x_2} \\ \frac{\gamma_1^0 x_3}{\gamma_2^0 + x_3} \\ \frac{\gamma_1^0 x_3}{\gamma_2^0 + x_3} \\ \frac{\gamma_1^0 x_4}{\gamma_2^0 + x_4} \\ \frac{\gamma_1^0 x_4}{\gamma_2^0 + x_4} \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \frac{x_1}{\gamma_2^0 + x_1} & -\frac{\gamma_1^0 x_1}{(\gamma_2^0 + x_1)^2} \\ \frac{x_1}{\gamma_2^0 + x_1} & -\frac{\gamma_1^0 x_1}{(\gamma_2^0 + x_1)^2} \\ \frac{x_2}{\gamma_2^0 + x_2} & -\frac{\gamma_1^0 x_2}{(\gamma_2^0 + x_2)^2} \\ \frac{x_2}{\gamma_2^0 + x_2} & -\frac{\gamma_1^0 x_2}{(\gamma_2^0 + x_2)^2} \\ \frac{x_3}{\gamma_2^0 + x_3} & -\frac{\gamma_1^0 x_3}{(\gamma_2^0 + x_3)^2} \\ \frac{x_3}{\gamma_2^0 + x_3} & -\frac{\gamma_1^0 x_3}{(\gamma_2^0 + x_3)^2} \\ \frac{x_4}{\gamma_2^0 + x_4} & -\frac{\gamma_1^0 x_4}{(\gamma_2^0 + x_4)^2} \\ \frac{x_4}{\gamma_2^0 + x_4} & -\frac{\gamma_1^0 x_4}{(\gamma_2^0 + x_4)^2} \end{pmatrix}, \quad \begin{pmatrix} \delta\gamma_1 \\ \delta\gamma_2 \end{pmatrix} = \begin{pmatrix} \gamma_1 - \gamma_1^0 \\ \gamma_2 - \gamma_2^0 \end{pmatrix}$$

and

$$\mathbf{\Sigma}_{\boldsymbol{\theta}} = \theta_1 \mathbf{V}_1 + \theta_2 \mathbf{V}_2$$

with $\theta_1 = \sigma_1^2,\ \theta_2 = \sigma_2^2$,

$$\mathbf{V}_1 = \begin{pmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{pmatrix}, \qquad \mathbf{V}_2 = \begin{pmatrix} 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1 \end{pmatrix}.$$

We consider a special case of a general linear regression model described in section 3—a mixed linear model, the variance components have to be non-negative and matrices $\mathbf{V}_1, \mathbf{V}_2$ are evidently positive semidefinite. We have to take this fact into account when determining the insensitivity regions.

We will work with following measured data

$$\mathbf{Y} = \begin{pmatrix} 1.245\\ 0.779\\ 2.264\\ 2.258\\ 6.612\\ 5.909\\ 6.827\\ 6.301 \end{pmatrix}$$

Components of this observation vector are the result of the data simulation from the normal distribution with mean equals to the real value of f for appropriate x_i and standard deviation $\sigma_1 = 0.2$, $\sigma_2 = 0.6$ respectively.

The next task is to decide whether we are able to get reasonable estimates of the variance components θ_1 and θ_2 when having this one observation vector only. We find some starting values θ_1^0 and θ_2^0 and establish the insensitivity region according to (20). Next we compute the variance components estimates based on these starting values and confidence region for the variance components. If the confidence region is inbedded into the insensitivity region, then the choice of starting θ_1^0 and θ_2^0 is good enough to determine the estimates based on them.

What we can do is to get a rough estimate $\theta_{1,1}^0$ of θ_1 based on y_1 and y_2—the first two components of observation vector:

$$\theta_{1,1}^0 = (y_1 - \overline{y}_{12})^2 + (y_2 - \overline{y}_{12})^2$$

and $\theta_{1,2}^0$ based on y_3 and y_4

$$\theta_{1,2}^0 = (y_3 - \overline{y}_{34})^2 + (y_4 - \overline{y}_{34})^2,$$

where $\overline{y}_{12}$ ($\overline{y}_{34}$) denotes arithmetic average of y_1 and y_2 (y_3 and y_4 respectively). Since all the four values $y_1, \ldots, y_4$ were simulated with dispersion θ_1, the starting value θ_1^0 can be counted as an arithmetic average of $\theta_{1,1}^0$ and $\theta_{1,2}^0$:

$$\theta_1^0 = \frac{\theta_{1,1}^0 + \theta_{1,2}^0}{2} = 0.0543.$$

Similarly

$$\theta_{2,1}^0 = (y_5 - \overline{y}_{56})^2 + (y_6 - \overline{y}_{56})^2, \qquad \theta_{2,2}^0 = (y_7 - \overline{y}_{78})^2 + (y_8 - \overline{y}_{78})^2$$

and

$$\theta_2^0 = \frac{\theta_{2,1}^0 + \theta_{2,2}^0}{2} = 0.1927.$$

Next we need some starting values γ_1^0 and γ_2^0. Lineweaver–Burke transformation (see [2] for more details) is frequently used to get the starting values of γ_1 and γ_2. However following approach can be useful sometimes. Since

$$y = \frac{\gamma_1 x}{\gamma_2 + x},$$

$$\gamma_1 x - \gamma_2 y = xy. \tag{22}$$

According to this we can put together four systems of two linear equations:

$$\begin{matrix}\gamma_1 x_1 - \gamma_2 y_1 = x_1 y_1 \\ \gamma_1 x_2 - \gamma_2 y_3 = x_2 y_3\end{matrix}, \qquad \begin{matrix}\gamma_1 x_1 - \gamma_2 y_2 = x_1 y_2 \\ \gamma_1 x_2 - \gamma_2 y_4 = x_2 y_4\end{matrix},$$

$$\begin{matrix}\gamma_1 x_3 - \gamma_2 y_5 = x_3 y_5 \\ \gamma_1 x_4 - \gamma_2 y_7 = x_4 y_7\end{matrix}, \qquad \begin{matrix}\gamma_1 x_3 - \gamma_2 y_6 = x_3 y_6 \\ \gamma_1 x_4 - \gamma_2 y_8 = x_4 y_8\end{matrix}.$$

When we denote the solutions of these systems with $(\gamma_{1,1}^0, \gamma_{2,1}^0), \ldots, (\gamma_{1,4}^0, \gamma_{2,4}^0)$, we can get the starting values γ_1^0, γ_2^0 as the arithmetic averages again:

$$\gamma_1^0 = \frac{\gamma_{1,1}^0 + \gamma_{1,2}^0 + \gamma_{1,3}^0 + \gamma_{1,4}^0}{4} \quad \text{and} \quad \gamma_2^0 = \frac{\gamma_{2,1}^0 + \gamma_{2,2}^0 + \gamma_{2,3}^0 + \gamma_{2,4}^0}{4}.$$

The results of this procedure are:

$$\gamma_1^0 = 16.068, \qquad \gamma_2^0 = 8.250.$$

Now we are ready to determine the variance components insensitivity region for

$$\boldsymbol{\theta} = \boldsymbol{\theta}_0 = \begin{pmatrix} \theta_1^0 \\ \theta_2^0 \end{pmatrix}.$$

Let's consider $\mathbf{g}_1 = (1, 0)'$ at first. In this case we have $\mathbf{g}_1'\boldsymbol{\theta} = \theta_1$.

Using (20), $\varepsilon = 0.1$ and taking into account fact that the negative input variance components values does not make sense in our case we get:

$$\mathcal{N}_{\mathbf{g}_1, \boldsymbol{\theta}_0} = \Big\{ \boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} \geq 0 \\ \wedge\, \delta\boldsymbol{\theta}' \begin{pmatrix} 0.000806 & -0.000227 \\ -0.000227 & 0.000064 \end{pmatrix} \delta\boldsymbol{\theta} \leq 0.000393 \Big\}. \tag{23}$$

For $\varepsilon = 0.1$ we get a set of all $\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}$ which don't increase the standard deviation of $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta})$ more than by 10 %.

Now the same once more for $\mathbf{g}_2 = (0,1)'$ and $\varepsilon = 0.1$

$$\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}_0} = \{\boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} : \boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} \geq 0$$

$$\delta\boldsymbol{\theta}' \begin{pmatrix} 0.0107 & -0.00300 \\ -0.00300 & 0.000846 \end{pmatrix} \delta\boldsymbol{\theta} \leq 0.00508 \Big\} . \qquad (24)$$

According to (3) we get (for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$) variance components estimate:

$$\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0) = \begin{pmatrix} 0.0474 \\ 0.165 \end{pmatrix} .$$

(Let's denote the real values of the variance components with $\boldsymbol{\theta}^*$, we have $\boldsymbol{\theta}^* = \binom{0.04}{0.36}$. The difference $(\boldsymbol{\theta}^* - \widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0))$ is compatible with the variance of the data.)

Let's determine another set—a rectangle with center $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$ which covers the real value of $\boldsymbol{\theta}$ with probability of $(1-\alpha)$—$\boldsymbol{\theta}$ confidence region. According to Chebyshev (see [6]) we have

$$P\left\{ |\widehat{\theta}_1(\boldsymbol{\theta}_0) - \theta_1| \leq k\sqrt{\operatorname{Var}_{\boldsymbol{\theta}_0}\left[\widehat{\theta}_1(\boldsymbol{\theta}_0)\right]} \right\} \geq 1 - \frac{1}{k^2}$$

and

$$P\left\{ |\widehat{\theta}_2(\boldsymbol{\theta}_0) - \theta_2| \leq k\sqrt{\operatorname{Var}_{\boldsymbol{\theta}_0}\left[\widehat{\theta}_2(\boldsymbol{\theta}_0)\right]} \right\} \geq 1 - \frac{1}{k^2}.$$

According to Bonferroni (see [3])

$$P\left\{ |\widehat{\theta}_1(\boldsymbol{\theta}_0) - \theta_1| \leq k\sqrt{\operatorname{Var}_{\boldsymbol{\theta}_0}\left[\widehat{\theta}_1(\boldsymbol{\theta}_0)\right]} \wedge |\widehat{\theta}_2(\boldsymbol{\theta}_0) - \theta_2| \leq k\sqrt{\operatorname{Var}_{\boldsymbol{\theta}_0}\left[\widehat{\theta}_2(\boldsymbol{\theta}_0)\right]} \right\}$$

$$\geq 1 - \frac{2}{k^2}. \qquad (25)$$

In (25) we get a rectangle with center $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$. We determine this rectangle to include the variance components with probality of $(1-\alpha)$. This means we need to have $(1 - \frac{2}{k^2}) = (1-\alpha)$, so $k = \sqrt{\frac{2}{\alpha}}$. As mentioned above the variance components cannot be negative in this case as they are the variances in fact. We have to involve this into our consideration. The $\boldsymbol{\theta}$ confidence region $\mathcal{E}_{\boldsymbol{\theta}}$ is a set of non-negative θ_1, θ_2 satisfying (25):

$$\mathcal{E}_{\boldsymbol{\theta}} = \left\{ \boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} : |\widehat{\theta}_1(\boldsymbol{\theta}_0) - \theta_1| \leq \sqrt{\frac{2}{\alpha}\operatorname{Var}_{\boldsymbol{\theta}_0}\left[\widehat{\theta}_1(\boldsymbol{\theta}_0)\right]} \right.$$

$$\left. \wedge\, |\widehat{\theta}_2(\boldsymbol{\theta}_0) - \theta_2| \leq \sqrt{\frac{2}{\alpha}\operatorname{Var}_{\boldsymbol{\theta}_0}\left[\widehat{\theta}_2(\boldsymbol{\theta}_0)\right]} \wedge \theta_1 \geq 0 \wedge \theta_2 \geq 0 \right\}$$

According to (4) the variance-covariance matrix of variance components estimates $\widehat{\boldsymbol{\theta}}$ is

$$\mathrm{Var}_{\boldsymbol{\theta}_0}\left(\widehat{\boldsymbol{\theta}}\right) = 2\mathbf{S}^{-1}_{(\mathbf{M_X}\boldsymbol{\Sigma}_{\boldsymbol{\theta}_0}\mathbf{M_X})^+} = \begin{pmatrix} 0.00197 & -0.0000894 \\ -0.0000894 & 0.0254 \end{pmatrix}.$$

For $(1-\alpha) = 0.95$ level of confidence we have

$$\mathcal{E}_{\boldsymbol{\theta}} = \langle 0; 0.328\rangle \times \langle 0; 1.173\rangle \tag{26}$$

This confidence region $\mathcal{E}_{\boldsymbol{\theta}}$ is a subset of both $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}_0}$ as shown in figures 2 and 4 given bellow.

It means the starting values θ_1^0, θ_2^0 are sufficient not only for deriving a reasonable estimate of $\boldsymbol{\theta}$ based on them—it is enough to stop the iterative procedure after the first step already.

Since we know the real values of θ_1, θ_2 we can repeat the same routine with these real values instead of the starting ones. Denote:

$$\boldsymbol{\theta}^* = \begin{pmatrix} \theta_1^* \\ \theta_2^* \end{pmatrix} = \begin{pmatrix} 0,04 \\ 0,36 \end{pmatrix}.$$

The insensitivity region for $\mathbf{g}_1 = \binom{1}{0}$ and $\varepsilon = 0.1$ is

$$\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}^*} = \{\boldsymbol{\theta}^* + \delta\boldsymbol{\theta} : \boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} \geq 0$$

$$\delta\boldsymbol{\theta}' \begin{pmatrix} 0.000143 & -0.0000159 \\ -0.0000159 & 0.00000177 \end{pmatrix} \delta\boldsymbol{\theta} \leq 0.000214 \Big\}. \tag{27}$$

For $\mathbf{g}_2 = \binom{0}{1}$ and $\varepsilon = 0.1$

$$\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}^*} = \{\boldsymbol{\theta}^* + \delta\boldsymbol{\theta} : \boldsymbol{\theta}_0 + \delta\boldsymbol{\theta} \geq 0$$

$$\delta\boldsymbol{\theta}' \begin{pmatrix} 0.0118 & -0.00131 \\ -0.00131 & 0.000145 \end{pmatrix} \delta\boldsymbol{\theta} \leq 0.0175 \Big\}. \tag{28}$$

$\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}^*)$ is according to (3):

$$\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}^*) = \begin{pmatrix} 0.0476 \\ 0.164 \end{pmatrix}.$$

The $\boldsymbol{\theta}$ confidence region with center $\boldsymbol{\theta}^*$ for $(1-\alpha) = 0.95$ level of confidence is

$$\mathcal{E}_{\boldsymbol{\theta}} = \langle 0; 0.254\rangle \times \langle 0; 2.032\rangle \tag{29}$$

In figures number 2 and 4 given bellow we can see the situation for $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ is quite similar to that for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. The confidence region is again a subset of both the insensitivity regions. This is exactly what we could expect and it shows that model (21) enables to determine reasonable variance components estimates.

The insensitivity regions (23) and (27) relating to $\mathbf{g}_1 = \binom{1}{0}$ and confidence regions (26) and (29) are visible in Fig. 1.

The relative position of the insenstivity regions and confidence regions is not obvious in Fig. 1. Fig. 2 gives a more detailed view.

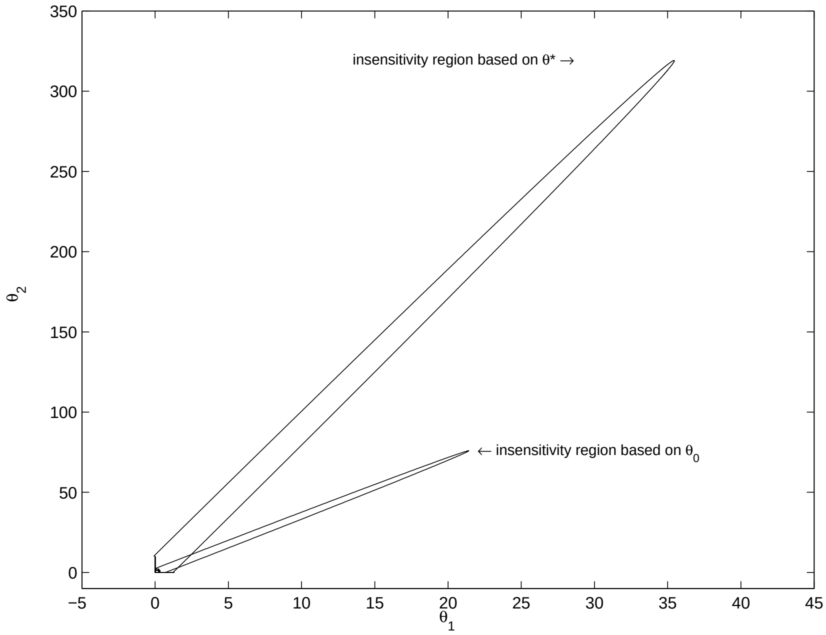


Figure 1: Insensitivity regions $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}^*}$ for $\varepsilon = 0.1$ and appropriate confidence regions for $\alpha = 0.05$.

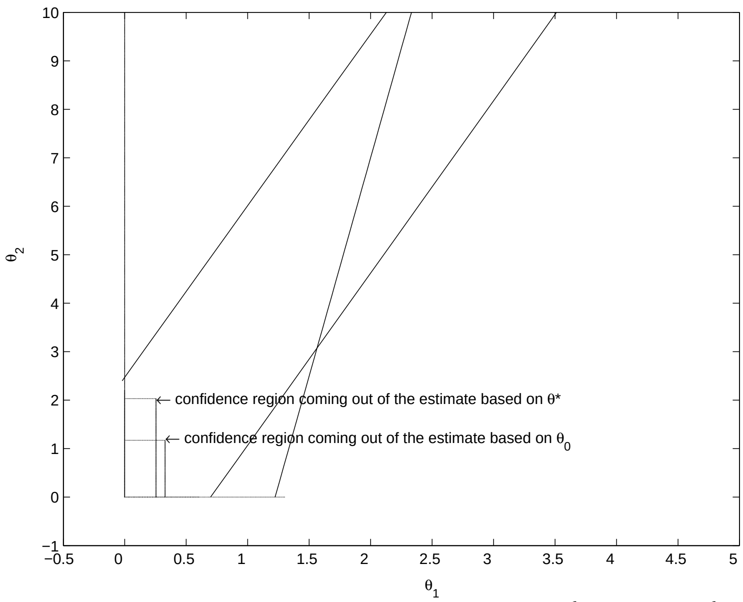


Figure 2. Relative position of insensitivity regions $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}^*}$ and corresponding confidence regions—detail.

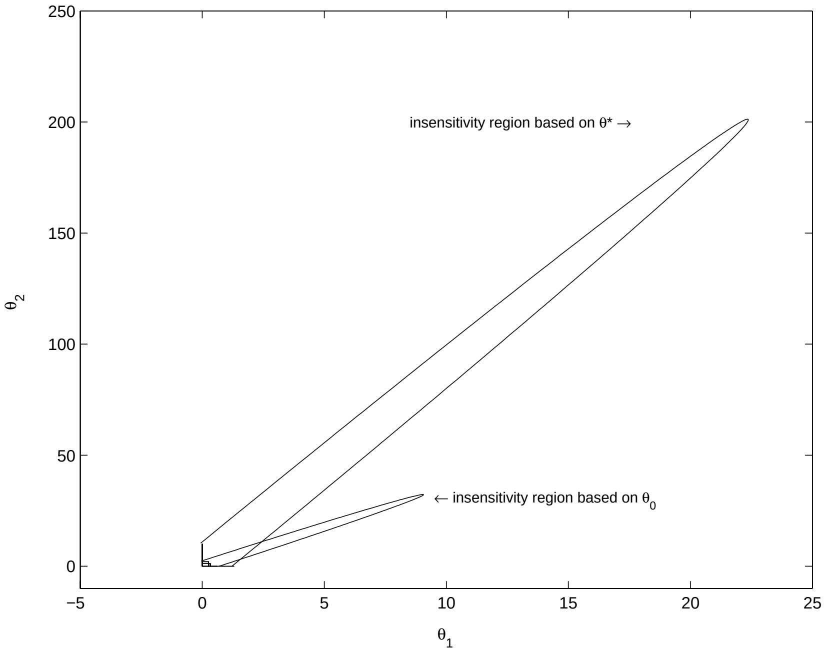


Figure 3. Insensitivity regions $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}^*}$ for $\varepsilon = 0.1$ and appropriate confidence regions for $\alpha = 0.05$.

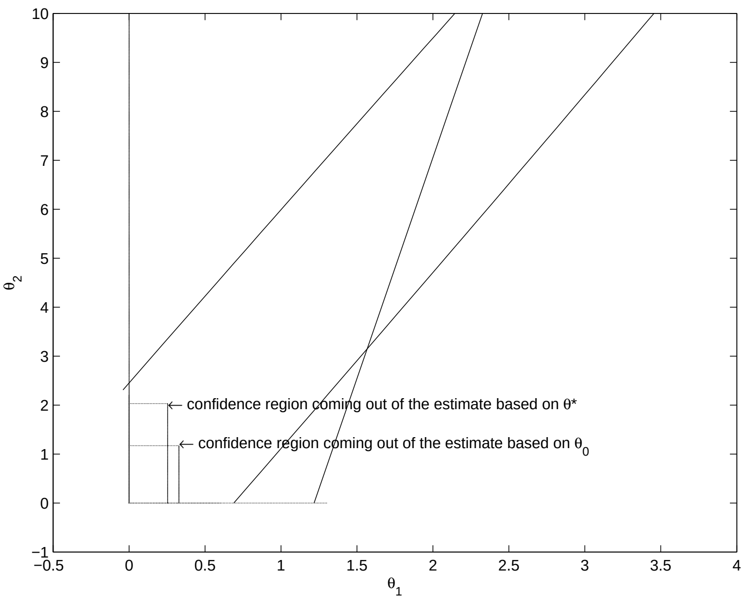


Figure 4. Relative position of insensitivity regions $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}^*}$ and corresponding confidence regions—detail.

The insensitivity regions (24) and (28) relating to $\mathbf{g}_2 = \binom{0}{1}$ and confidence regions (26) and (29) are visible in Fig. 3.

The relative position of the insenstivity regions and confidence regions is again shown in Fig. 4 in detail.

The conclusion is that we can feel free to use the $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$ as a $\boldsymbol{\theta}$ estimate without apprehension of the variance of the estimates being too large because both the insensitivity regions for $\boldsymbol{\theta}_0$ $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}_0}$ unambiguously cover the $\boldsymbol{\theta}$ confidence region based on $\widehat{\boldsymbol{\theta}}(\boldsymbol{\theta}_0)$.

The comparison of $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}^*}$ is visible in Fig. 1. Fig. 3 contains similar comparison of $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}^*}$. When we compare the size of the insensitivity regions $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_1,\boldsymbol{\theta}^*}$ ($\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}_0}$ and $\mathcal{N}_{\mathbf{g}_2,\boldsymbol{\theta}^*}$ respectively) we can see the insensitivity regions based on the real values $\boldsymbol{\theta}^*$ are much larger then those based on the prior values $\boldsymbol{\theta}_0$. This means the variance increase is slower when we base the estimates on the real values of the variance components compared to the estimates based on the prior values generated from the observation vector.

References

[1] Boháčová, H., Heckenbergerová, J.: *Insensitivity regions for fixed effects parameters in mixed linear model with type I constraints and calculation problems linked with them.* In: Forum Statisticum Slovacum 6/2007, ISSN 1336-7420, (in Czech).

[2] Ebel, S., Alert, D., Schaefer, U.: *Calibration in TLC/HPTLC using the Michaelis-Menten function.* Chromatographia 1/1984.

[3] Humak, K. M. S.: Statistischen Methoden der Modellbildung Band I, Statistische Inferenz für lineare Parameter. *Akademie–Verlag, Berlin*, 1977.

[4] Kubáček, L., Kubáčková, L.: Statistika a metrologie. *Vyd. University Palackáho, Olomouc*, 2000.

[5] Rao, C. R., Kleffe, J.: Estimation of Variance Components and Application. *North-Holland, Amsterdam–New York–Oxford–Tokyo*, 1988.

[6] Rényi, A.: Probability theory. *Akadémiai Kiadó, Budapest*, 1970.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 23–25

Discriminator Order Algebras*

Ivan CHAJDA

*Department of Algebra and Geometry, Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: chajda@inf.upol.cz*

(Received January 24, 2008)

Abstract

We prove that an order algebra assigned to a bounded poset with involution is a discriminator algebra.

Key words: Order algebra; ordered set; involution; ternary discriminator.

2000 Mathematics Subject Classification: 06A06, 06A11, 08A40

In accordance with [1], by an order algebra we mean an algebra defined on an ordered set whose operations are derived by means of the order relation and, conversely, the partial order is determined by these operations.

Let $\mathcal{P} = (P; \leq, 1)$ be an ordered set with the greatest element 1. The following two operations are introduced in [1]:

$$x \to y = \begin{cases} 1 \text{ if } x \leq y \\ \\ y \text{ otherwise} \end{cases} \qquad \text{and} \qquad x \circ y = \begin{cases} y \text{ if } x \leq y \\ \\ 1 \text{ otherwise.} \end{cases}$$

Let us mention that these operations are not independent:

Observation 1 For any ordered set $\mathcal{P} = (P; \leq, 1)$ we have $x \circ y = (x \to y) \to y$.

Moreover, we have $x \to y \in \{y, 1\}$ and hence for any interval $[y, 1]$ of $\mathcal{P}$ it holds $x \to y \in [y, 1]$. Thus, having $a \in [y, 1]$, we can define a unary operation on the interval $[y, 1]$ assigning to a the element $a^y = a \to y$. Evidently, $y^y = 1$ and

*Supported by the Research Project MSM 6198959214.

$1^y = y$ thus this operation interchanges the endpoints of the interval $[y, 1]$ and hence it is called a *section switching mapping*. Hence, the operation $\rightarrow$ determines not only the the order $\leq$ but also the family $({}^p)_{p\in P}$ of section switching mappings, i.e. the extended structure $\mathcal{P} = (P; \leq, 1, ({}^p)_{p\in P})$.

We can ask if also conversely the operation $\circ$ can determine the operation $\rightarrow$. Since also $x \circ y \in [y, 1]$, the operation $(x \circ y)^y$ is defined correctly whenever ^{y} denotes the section switching mapping on the interval $[y, 1]$. Hence, we can state

Observation 2 Let $\mathcal{P} = (P; \leq, 1, ({}^p)_{p\in P})$ be an ordered set with 1 and with section switching mappings. Then $x \rightarrow y = (x \circ y)^y$ for the above mention operations $\rightarrow$ and $\circ$.

An ordered set $\mathcal{P}$ is *bounded* if it has a least element 0 and a greatest element 1. This will be expressed by the notation $\mathcal{P} = (P; \leq, 0, 1)$.

By an involution on a set P is meant a mapping of P into itself denoted by $x \mapsto x'$ satisfying $x'' = x$. Every bounded poset $\mathcal{P} = (P; \leq, 0, 1)$ admits some involutions. Let us pick up one of them which satisfies $0' = 1$. Hence, $x'' = x$ gets immediately $1' = 0$ and thus this involution is a switching mapping. Then we can enlarge the type of $\mathcal{P}$ and we will write $\mathcal{P} = (P; \leq, 0, 1,')$ to express the fact that this involution is considered as a basic operation of $\mathcal{P}$. From now on, $\mathcal{P} = (P; \leq, 0, 1,')$ will be called a *poset with involution*.

Let $\mathcal{P} = (P; \leq, 0, 1)$ be a bounded poset. By a *globalization* (frequently called also a Baaz operation named by M. Baaz) is meant a unary operation $\triangle$ on P defined by

$$\triangle(1) = 1 \quad \text{and} \quad \triangle(x) = 0 \quad \text{for } x \neq 1.$$

Observation 3 In every poset with involution $\mathcal{P} = (P; \leq, 0, 1,')$, we can define a globalization $\triangle$ by means of $\rightarrow$, $'$ and 0 as follows

$$\triangle(x) = x' \rightarrow 0.$$

Another binary operation defined on an ordered set $(P; \leq)$ is mentioned in [1]:

$$x \sqcap y = \begin{cases} x & \text{if } x \leq y \\ \\ y & \text{otherwise.} \end{cases}$$

Now, let $\mathcal{P} = (P; \leq, 0, 1,')$ be a poset with involution. Define the assigned order algebra $\mathcal{A}(P) = (P; \rightarrow, \sqcap,', 0)$ of type $(2, 2, 1, 0)$ where $\rightarrow$ and $\sqcap$ are the above mentioned operations and $'$ is the involution of P.

As stated by Observations 1 and 3, the globalization $\triangle$ and the operation $\circ$ (as well as the constant 1) are term operations of $\mathcal{A}(P)$. We can state our main result:

Theorem 1 *Let* $\mathcal{P} = (P; \leq, 0, 1,')$ *be a poset with involution and* $\mathcal{A}(P) = (P; \rightarrow, \sqcap,', 0)$ *the assigned algebra. Then* $\mathcal{A}(P)$ *is a discriminator algebra whose ternary discriminator is*

$$t(x, y, z) = ((\triangle((x \rightarrow y)' \circ (y \rightarrow x)')') \rightarrow z) \sqcap (\triangle(((x \rightarrow y)' \circ (y \rightarrow x)')' \rightarrow 0) \rightarrow x).$$

Proof If card $P = 1$, the proof is trivial. Suppose card $P > 1$, i.e. $0 \neq 1$. It is an easy observation that $\sqcap$ satisfies

$$x \sqcap 1 = x = 1 \sqcap x. \tag{1}$$

For the sake of brevity, denote by

$$e(x, y) = ((x \to y)' \circ (y \to x)')'.$$

Due to the previous Observations 1 and 3, $e(x, y)$ is a term operation of $\mathcal{A}(P)$.

Clearly $e(x, x) = (1' \circ 1')' = (0 \circ 0)' = 0' = 1$. Suppose $x \neq y$.

(a) If $x < y$ then $x \neq 1$ and $x \to y = 1$, $y \to x = x$ and hence

$$e(x, y) = (1' \circ x')' = (0 \circ x') = x'' = x \neq 1.$$

(b) If $y < x$ then $y \neq 1$, i.e. $y' \neq 0$ and $x \to y = y$, $y \to x = 1$ thus

$$e(x, y) = (y' \circ 1')' = (y' \circ 0)' = 1' = 0 \neq 1.$$

(c) If $x \parallel y$ then $x \to y = y$, $y \to x = x$ and

$$e(x, y) = (y' \circ x')' = \begin{cases} 1' = 0 & \text{for } y' \leq x' \\ x'' = x & \text{for } y' \not\leq x'. \end{cases}$$

Since $x \parallel y$ we have $x \neq 1$ thus $e(x, y) \neq 1$ for $x \neq y$ in all the cases.

The term $t(x, y, z)$ can be clearly rewritten as follows

$$t(x, y, z) = (\triangle(e(x, y)) \to z) \sqcap (\triangle(e(x, y) \to 0) \to x).$$

Using of (1), we compute

$$t(x, x, z) = (\triangle(1) \to z) \sqcap (\triangle(1 \to 0) \to x) = (1 \to z) \sqcap (0 \to x) = z \sqcap 1 = z$$

and for $x \neq y$

$$t(x, y, z) = (0 \to z) \sqcap (\triangle(0 \to 0) \to x) = 1 \sqcap (\triangle(1) \to x) = 1 \sqcap x = x.$$

Hence, $t(x, y, z)$ is a term function of $\mathcal{A}(P)$ which is the ternary discriminator. $\square$

References

[1] Berman J., Blok W. J.: *Algebras Defined from Ordered Sets and the Varieties they Generate.* Order **23** (2006), 65–88.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 27–36

Monadic Basic Algebras*

Ivan CHAJDA [1], Miroslav KOLAŘÍK [2]

[1] *Department of Algebra and Geometry, Faculty of Science, Palacký University*
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: chajda@inf.upol.cz

[2] *Department of Computer Science, Faculty of Science, Palacký University*
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kolarik@inf.upol.cz

(Received October 30, 2007)

Abstract

The concept of monadic MV-algebra was recently introduced by A. Di Nola and R. Grigolia as an algebraic formalization of the many-valued predicate calculus described formerly by J. D. Rutledge [9]. This was also genaralized by J. Rachůnek and F. Švrček for commutative residuated ℓ-monoids since MV-algebras form a particular case of this structure. Basic algebras serve as a tool for the investigations of much more wide class of non-classical logics (including MV-algebras, orthomodular lattices and their generalizations). This motivates us to introduce the monadic basic algebra as a common generalization of the mentioned structures.

Key words: Basic algebra; monadic basic algebra; existential quantifier; universal quantifier; lattice with section antitone involution.

2000 Mathematics Subject Classification: 06D35, 03G25

Having an MV-algebra $\mathcal{A} = (A; \oplus, \neg, 0)$, one can derive the structure of bounded distributive lattice $\mathcal{L}(A) = (A; \vee, \wedge, 0, 1)$ where $1 = \neg 0$, $x \vee y = \neg(\neg x \oplus y) \oplus y$ and $x \wedge y = \neg(\neg x \vee \neg y)$. Moreover, to any element $a \in A$ one can assign an antitone involution $x \mapsto x^a$ on the interval $[a, 1]$ in $\mathcal{L}(A)$ given by $x^a = \neg x \oplus a$ (for $x \in [a, 1]$). Hence, $\mathcal{L}(A)$ is a lattice equipped by a set $(^a)_{a \in A}$ of partial unary operations defined on the so-called *sections* where for each $x \in [a, 1]$ we have $x^{aa} = x$ and for $x, y \in [a, 1]$ with $x \leq y$ we have $y^a \leq x^a$ (see e.g. [3] for

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

details). Such an enriched lattice (not necessarily distributive) is denoted by $\mathcal{L} = (L; \vee, \wedge, ({}^a)_{a\in L}, 0, 1)$ and is called a *lattice with section antitone involutions.*

Although this structure plays a crucial role in some formalizations of non-classical logics, it can be difficult to deal with since it is not a total algebra and, moreover, its similarity type depends on the cardinality of its elements. To improve this discrepancy, the following concept was introduced. Let us only note that the following axiom system (BA1)–(BA4) was recently involved in [6] as a simplification of the previous one (see e.g. [1, 2, 5]).

Definition 1 By a *basic algebra* is meant an algebra $\mathcal{A} = (A; \oplus, \neg, 0)$ of type $(2, 1, 0)$ satisfying the following axioms

(BA1) $x \oplus 0 = x$;

(BA2) $\neg\neg x = x$ (double negation);

(BA3) $\neg(\neg x \oplus y) \oplus y = \neg(\neg y \oplus x) \oplus x$ (Łukasiewicz axiom);

(BA4) $\neg(\neg(\neg(x \oplus y) \oplus y) \oplus z) \oplus (x \oplus z) = \neg 0$.

In what follows we will denote $\neg 0$ by 1 (as it is usual for MV-algebras). It is plain to show that every basic algebra satisfies also the identities $\neg 1 = 0$, $0 \oplus x = x$ and $\neg x \oplus x = 1$, see e.g. [4, 6].

As promised above, we can get the mutual relationship between lattices with section antitone involutions and basic algebras. For the proof, see e.g. [1] or [5].

Proposition 1 *(a) Let $\mathcal{L} = (L; \vee, \wedge, ({}^a)_{a\in L}, 0, 1)$ be a lattice with section antitone involutions. Then the assigned algebra $\mathcal{A}(L) = (L; \oplus, \neg, 0)$, where*

$$x \oplus y = (x^0 \vee y)^y \quad \text{and} \quad \neg x = x^0$$

is a basic algebra.

(b) Conversely, given a basic algebra $\mathcal{A} = (A; \oplus, \neg, 0)$, we can assign a bounded lattice with section antitone involutions $\mathcal{L}(A) = (A; \vee, \wedge, ({}^a)_{a\in A}, 0, 1)$, where $1 = \neg 0$,

$$x \vee y = \neg(\neg x \oplus y) \oplus y, \qquad x \wedge y = \neg(\neg x \vee \neg y)$$

and for each $a \in A$, the mapping $x \mapsto x^a = \neg x \oplus a$ is an antitone involution on the principal filter $[a, 1]$, where the order is given by

$$x \leq y \quad \text{if and only if} \quad \neg x \oplus y = 1.$$

(c) The assignments are in a one-to-one correspondence, i.e. $\mathcal{A}(\mathcal{L}(A)) = \mathcal{A}$ and $\mathcal{L}(\mathcal{A}(L)) = \mathcal{L}$.

Hence, when investigating basic algebras, we can switch to lattices with section antitone involutions whenever it is useful.

The lattice $\mathcal{L}(A) = (A; \vee, \wedge, ({}^a)_{a\in A}, 0, 1)$ will be referred as an *assigned lattice* of a basic algebra $\mathcal{A} = (A; \oplus, \neg, 0)$ and the order $\leq$ of $\mathcal{L}(A)$ as the *induced order* of $\mathcal{A}$.

Definition 2 By a *monadic basic algebra* is meant an algebra $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ of type $(2, 1, 1, 0)$ where $(A; \oplus, \neg, 0)$ is a basic algebra and the unary operation $\exists$ satisfies the following identities

(E1) $x \leq \exists x$;

(E2) $\exists(x \vee y) = \exists x \vee \exists y$;

(E3) $\exists(\neg \exists x) = \neg \exists x$;

(E4) $\exists(\exists x \oplus \exists y) = \exists x \oplus \exists y$.

The mapping $\exists \colon A \to A$ is called an *existential quantifier* on $\mathcal{A}$. By a *strict monadic basic algebra* will be called a monadic basic algebra satisfying the identity

(E5) $\exists(x \oplus x) = \exists x \oplus \exists x$.

Lemma 1 *Let* $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ *be a monadic basic algebra. Then the following conditions are satisfied:*

(i) $\exists 1 = 1$*;*

(ii) $\exists 0 = 0$*;*

(iii) $\exists\exists x = \exists x$*;*

(iv) $x \leq \exists y$ *if and only if* $\exists x \leq \exists y$*;*

(v) *if* $x \leq y$ *then* $\exists x \leq \exists y$*;*

(vi) $\neg \exists x \leq \exists(\neg x)$*;*

Proof Let x, y be arbitrary elements of $\mathcal{A}$.
(i): By (E1), $1 \leq \exists 1$, thus $\exists 1 = 1$ as 1 is the greatest element of $\mathcal{A}$.
(ii): By (i) and (E3), $0 = \neg 1 = \neg \exists 1 = \exists(\neg \exists 1) = \exists(\neg 1) = \exists 0$.
(iii): By (ii) and (E4), $\exists\exists x = \exists(\exists x \oplus 0) = \exists(\exists x \oplus \exists 0) = \exists x \oplus \exists 0 = \exists x \oplus 0 = \exists x$.
(iv): If $\exists x \leq \exists y$ then by (E1) also $x \leq \exists y$. On the other hand using (iii) and (E2), if $x \leq \exists y$ then $\exists y = \exists\exists y = \exists(x \vee \exists y) = \exists x \vee \exists\exists y = \exists x \vee \exists y$. Thus $\exists y = \exists x \vee \exists y$, and therefore $\exists x \leq \exists y$.
(v): Let $x \leq y$. Then, by (E2), we obtain $\exists y = \exists(x \vee y) = \exists x \vee \exists y$, and hence $\exists x \leq \exists y$.
(vi): Since $x \leq \exists x$ and hence $\neg x \geq \neg \exists x$, we conclude $\exists(\neg x) \geq \neg x \geq \neg \exists x$. □

In what follows let $\exists$ be a fixed existential quantifier defined on a basic algebra $\mathcal{A} = (A; \oplus, \neg, 0)$. By means of $\exists$, a unary operation $\forall$ can be defined on $\mathcal{A}$ by the rule

$$\forall x := \neg(\exists \neg x). \tag{1}$$

Lemma 2 *Let* $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ *be a monadic basic algebra and* $\forall$ *is defined by (R). Then the following conditions are satisfied*

(A1) $\forall x \leq x$*;*

(A2) $\forall(x \wedge y) = \forall x \wedge \forall y$*;*

(A3) $\forall(\neg\forall x) = \neg\forall x$;

(A4) $\forall(\forall x \odot \forall y) = \forall x \odot \forall y$, *where* $x \odot y = \neg(\neg x \oplus \neg y)$.

If, moreover, $\mathcal{A}$ *is a strict monadic basic algebra, then it satisfies also*

(A5) $\forall(x \odot x) = \forall x \odot \forall x$.

Proof By (E1), $\neg x \leq \exists\neg x$ thus $x = \neg\neg x \geq \neg(\exists\neg x) = \forall x$ proving (A1). To prove (A2), we use (E2) and the De Morgan laws:

$$\forall(x \wedge y) = \forall(\neg(\neg x \vee \neg y)) = \neg\exists(\neg x \vee \neg y)$$

$$= \neg((\exists\neg x) \vee (\exists\neg y)) = \neg(\exists\neg x) \wedge \neg(\exists\neg y) = \forall x \wedge \forall y.$$

Prove (A3): $\forall(\neg\forall x) = \neg\exists(\neg\neg\forall x) = \neg\exists(\neg(\exists\neg x)) = \neg\neg(\exists\neg x) = \neg\forall x$ by (E3). For (A4) we compute by (E4)

$$\forall(\forall x \odot \forall y) = \neg(\exists\neg(\neg(\exists\neg x) \odot \neg(\exists\neg y))) = \neg\exists(\exists\neg x \oplus \exists\neg y)$$

$$= \neg(\exists\neg x \oplus \exists\neg y) = \neg(\neg\neg(\exists\neg x) \oplus \neg\neg(\exists\neg y)) = \forall x \odot \forall y.$$

Assume that $\exists$ satisfies also (E5). Then

$$\forall(x \odot x) = \neg(\exists\neg(x \odot x)) = \neg(\exists\neg(\neg\neg x \odot \neg\neg x))$$

$$\neg(\exists(\neg x \oplus \neg x)) = \neg(\exists\neg x \oplus \exists\neg x) = (\neg(\exists\neg x)) \odot (\neg(\exists\neg x)) = \forall x \odot \forall x.$$

□

A unary operation $\forall\colon A \to A$ on a basic algebra $\mathcal{A} = (A; \oplus, \neg, 0)$ satisfying (A1)–(A4) will be called a *universal quantifier.*

It is a routine way to prove also the converse:

Lemma 3 *Let* $\mathcal{A} = (A; \oplus, \neg, 0)$ *be a basic algebra and* $\forall$ *be a universal quantifier on* $\mathcal{A}$. *Define*

$$\exists x := \neg(\forall\neg x).$$

Then $\mathcal{A}_\exists = (A; \oplus, \neg, \exists, 0)$ *is a monadic basic algebra. Moreover, if it satisfies also (A5) then* $\mathcal{A}_\exists$ *is a strict monadic basic algebra.*

Remark 1 Let $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ be a monadic basic algebra. Then $\exists$ is a closure operator and $\forall$ is an interior operator on the poset $(A; \leq)$, where the relation $\leq$ is the induced order on $\mathcal{A}$.

In what follows, we are going to prove a connection between monadic basic algebras and enriched lattices with section antitone involutions similarly as it was done for basic algebras in the Proposition. For this, let us recall some concepts.

For an algebra $\mathcal{A} = (A; F)$, by a *retraction* is meant an idempotent endomorphism h of $\mathcal{A}$, i.e. an endomorphism satisfying $h(h(x)) = h(x)$ for every $x \in A$. It is well-known that if h is a retraction of $\mathcal{A}$ then its image $\mathcal{A}_0 = h(A)$ is a subalgebra of $\mathcal{A}$, the so-called *retract of* $\mathcal{A}$.

In particular, if $\mathcal{S} = (S; \vee, 0)$ is a join-semilattice with 0, by a *retraction* is meant a self-mapping e of S satisfying (e1) $e(x \vee y) = e(x) \vee e(y), \quad e(0) = 0,$

(e2) $e(e(x)) = e(x).$

This retraction is called *extensive* if it satisfies also

(e3) $x \leq e(x).$

Example 1 Consider the bounded join-semilattice $\mathcal{S} = (A; \vee, 0, 1)$, where $A = \{0, a, b, 1\}$, depicted in Fig. 1.

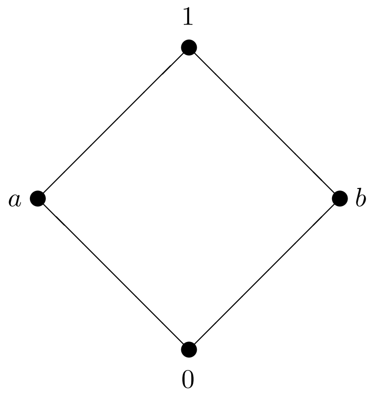


Fig. 1

Define $e\colon A \to A$ as follows

$$e(0) = 0, \quad e(a) = 1, \quad e(b) = b, \quad e(1) = 1.$$

Then e is an extensive retraction of $\mathcal{S}$ and the retract $S_0 = e(S)$ is the chain $\{0, b, 1\}$. Remark that the semilattice $\mathcal{S}$ can be considered also as a lattice but this e is not a lattice retraction since

$$e(a \wedge b) = e(0) = 0 \neq b = 1 \wedge b = e(a) \wedge e(b).$$

Now, let $\mathcal{L} = (L; \vee, \wedge, (^a)_{a \in L}, 0, 1)$ be a lattice with section antitone involutions. A mapping $e\colon L \to L$ will be called an **e**-*retraction* if it is an extensive retraction of the join-semilattice reduct $(L; \vee, 0)$ satisfying one more condition

(e4) $e(e(x)^{e(y)}) = e(x)^{e(y)}$ for every pair $y \leq x.$

Let us note that $y \leq x$ implies $e(y) \leq e(x)$ just by (e1).

If e is an e-retraction on a lattice $\mathcal{L} = (L; \vee, \wedge, (^a)_{a \in L}, 0, 1)$ with section antitone involutions then the enriched structure $\mathcal{L}_e = (L; \vee, \wedge, (^a)_{a \in L}, e, 0, 1)$ will be called a *monadic lattice*.

We are going to prove

Theorem 1 *Let $\mathcal{L}_e = (L; \vee, \wedge, (^a)_{a \in L}, e, 0, 1)$ be a monadic lattice and $\mathcal{A}_e(L) = (L; \oplus, \neg, e, 0)$ an algebra such that $\mathcal{A}(L) = (L; \oplus, \neg, 0)$ is a basic algebra assigned to the reduct $\mathcal{L} = (L; \vee, \wedge, (^a)_{a \in L}, 0, 1)$. Then $\mathcal{A}_e(L)$ is a monadic basic algebra.*

Proof We need only to show that $\mathcal{A}_e(L)$ satisfies the conditions (E3) and (E4) from Definition 2. To prove (E3) we compute:

$$e(\neg e(x)) = e(e(x)^0) \overset{(e1)}{=} e(e(x)^{e(0)}) \overset{(e4)}{=} e(x)^{e(0)} \overset{(e1)}{=} e(x)^0 = \neg e(x)$$

Further, we check the following identity

$$e(\neg e(x) \vee e(y)) = \neg e(x) \vee e(y). \tag{2}$$

For this, we compute

$$e(\neg e(x) \vee e(y)) \overset{(E3)}{=} e(e(\neg e(x)) \vee e(y))$$

$$\overset{(e1),(e2)}{=} e(\neg e(x)) \vee e(y) \overset{(E3)}{=} \neg e(x) \vee e(y).$$

Now, we are ready to prove (E4):

$$e(x) \oplus e(y) = (\neg e(x) \vee e(y))^{e(y)} \overset{(A)}{=} (e(\neg e(x) \vee e(y)))^{e(y)}$$

$$\overset{(e4)}{=} e((e(\neg e(x) \vee e(y)))^{e(y)}) \overset{(A)}{=} e((\neg e(x) \vee e(y))^{e(y)}) = e(e(x) \oplus e(y)).$$

□

We can prove also the converse.

Theorem 2 *Let $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ be a monadic basic algebra, let $\mathcal{L}(A) = (A; \vee, \wedge, (^a)_{a \in A}, 0, 1)$ be the assigned lattice of the reduct $(A; \oplus, \neg, 0)$. Then $\mathcal{L}_\exists(A) = (A; \vee, \wedge, (^a)_{a \in A}, \exists, 0, 1)$ is a monadic lattice.*

Proof We prove that the mapping $e\colon x \to \exists x$ is an e-retraction of $\mathcal{L}_\exists(A)$. Trivially, we have: $e(x \vee y) = \exists(x \vee y) = \exists x \vee \exists y = e(x) \vee e(y)$ and $e(0) = \exists 0 = 0$. Further, $e(e(x)) = \exists\exists x = \exists x = e(x)$ by (iii) of Lemma 1 and $x \leq e(x) = \exists x$ by (E1). We prove (e4): Since $x^y = \neg x \oplus y$ (for $x \in [y, 1]$), we have

$$e(e(x)^{e(y)}) = \exists((\exists x)^{(\exists y)}) = \exists((\neg \exists x) \oplus \exists y) \overset{(E3)}{=} \exists((\exists(\neg \exists x)) \oplus \exists y)$$

$$\overset{(E4)}{=} (\exists(\neg \exists x)) \oplus \exists y \overset{(E3)}{=} (\neg \exists x) \oplus \exists y = (\exists x)^{(\exists y)} = e(x)^{e(y)}.$$

□

Remark 2 If $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ is a strict monadic basic algebra then the assigned monadic lattice $\mathcal{L}(A)$ satisfies the condition

(e5) $\quad e((x^0 \vee x)^x) = (e(x)^0 \vee e(x))^{e(x)}$

(where $e(x)$ stands for $\exists x$ in $\mathcal{A}$) and vice versa, if a monadic lattice $\mathcal{L}$ satisfies (e5) then the assigned monadic basic algebra $\mathcal{A}(L)$ is strict.

Let $\mathcal{A} = (A; \oplus, \neg, 0)$ be a basic algebra. It is plain to check that the identity mapping $id(x) = x$ is an existential quantifier on $\mathcal{A}$. Moreover, define a mapping $j\colon A \to A$ as follows

$$j(0) = 0 \quad \text{and} \quad j(x) = 1 \quad \text{for } x \neq 0.$$

Then also j is an existential quantifier on $\mathcal{A}$. Hence, by Theorem 2, id and j are e-retractions on the assigned lattice $\mathcal{L}(A)$.

Example 2 For a basic algebra $\mathcal{H} = (H; \oplus, \neg, 0)$ with $H = \{0, a, b, 1\}$, where $\neg 0 = 1$, $\neg a = a$, $\neg b = b$, $\neg 1 = 0$, the assigned lattice is depicted in Fig. 2 (the antitone involutions in at most two-elements sections are determined uniquely).

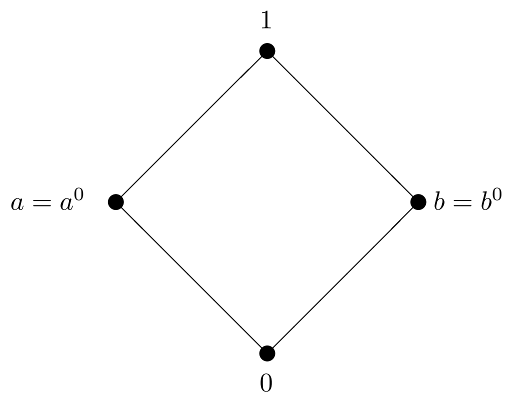


Fig. 2

There are four of e-retractions, namely id, j and h_1, h_2 defined by

$$h_1(0) = 0, \quad h_1(1) = 1, \quad h_1(a) = a, \quad h_1(b) = 1$$

and

$$h_2(0) = 0, \quad h_2(1) = 1, \quad h_2(a) = 1, \quad h_2(b) = b.$$

In what follows, we can borrow the following concept of relatively complete subalgebra, defined for MV-algebras in [7] and for residuated ℓ-monoids in [8]:

Definition 3 A subalgebra B of a basic algebra $\mathcal{A} = (A; \oplus, \neg, 0)$ is called *relatively complete* if for every $a \in A$ the set $\{b \in B;\, a \leq b\}$ has the least element. Further, a relative complete subalgebra B is called *m-relatively complete* if

$$\begin{aligned} &\text{for all } a \in A \text{ for all } b \in B\colon b \geq a \oplus a \text{ implies} \\ &\quad \text{that there exists } v \in B\colon v \geq a \text{ and } b \geq v \oplus v. \end{aligned} \tag{3}$$

Theorem 3 *Let $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ be a monadic basic algebra and $A_0 = \{\exists x;\, x \in A\}$. Then A_0 is a relatively complete subalgebra of $\mathcal{A}$. If, moreover, $\mathcal{A}$ is a strict monadic basic algebra then $\mathcal{A}_0$ is an m-relatively complete subalgebra of $\mathcal{A}$.*

Proof Due to (E3), (E4) and (ii), (iii) of Lemma 1, $\mathcal{A}_0$ is a subalgebra of $\mathcal{A}$. Let $a \in A$ and $B_a = \{b \in A_0;\ a \leq b\}$. Then $\exists a \in B_a$ and for any $b \in B_a$ we have $b = \exists d$ for some $d \in A$. Hence, $\exists a \leq \exists\exists d = \exists d = b$, thus $\exists a$ is the least element of B_a. Hence, $\mathcal{A}_0$ is a relatively complete subalgebra of $\mathcal{A}$. Assume that $\mathcal{A}$ is a strict monadic basic algebra. Let $a \in A$, $b \in A_0$ and $b \geq a \oplus a$. Then for $v = \exists a$ we have $v \geq a$ due to (E1) and, due to (E5), $b = \exists b \geq \exists(a \oplus a) = \exists a \oplus \exists a = v \oplus v$ proving (C). □

We have shown that any existential quantifier $\exists$ on a basic algebra $\mathcal{A} = (A; \oplus, \neg, 0)$ induces a relatively complete subalgebra $\mathcal{A}_0 = \exists A$ of $\mathcal{A}$. Also conversely, every relatively complete subalgebra of $\mathcal{A}$ gives rise to an existential quantifier.

We say that a basic algebra $\mathcal{A}$ is $\oplus$-*monotonous* if $x \geq y$ implies $x \oplus x \geq y \oplus y$. Let us note that e.g. every MV-algebra or an effect algebra satisfies this condition.

Theorem 4 *Let $\mathcal{A} = (A; \oplus, \neg, 0)$ be a basic algebra and $\mathcal{A}_0$ its relatively complete subalgebra. For any $a \in A$, define $\exists a = \inf\{b \in A_0;\ a \leq b\}$. Then $\mathcal{A}_\exists = (A; \oplus, \neg, \exists, 0)$ is a monadic basic algebra. If, moreover, $\mathcal{A}$ is $\oplus$-monotonous and $\mathcal{A}_0$ is an m-relatively complete subalgebra of $\mathcal{A}$ then $\mathcal{A}_\exists = (A; \oplus, \neg, \exists, 0)$ is a strict monadic basic algebra.*

Proof It is evident that $x \leq \inf\{b \in A_0;\ x \leq b\} = \exists x$ and that $x \leq y$ implies $\exists x \leq \exists y$, i.e. also $\exists(x \vee y) \geq \exists x \vee \exists y$. Since $\mathcal{A}_0$ is a subalgebra of $\mathcal{A}$ and $\exists x, \exists y \in A_0$, also $\exists x \vee \exists y \in A_0$ and $x \leq \exists x$, $y \leq \exists y$ thus also $x \vee y \leq \exists x \vee \exists y$. Hence, $\exists x \vee \exists y \in \{b \in A_0;\ x \vee y \leq b\} = B_{x\vee y}$, i.e. $\exists(x \vee y) = \inf B_{x \vee y} \leq \exists x \vee \exists y$.

Evidently, $\exists x = x$ for any $x \in A_0$. Since $\exists x \in A_0$ for each $x \in A$ and $\mathcal{A}_0$ is a subalgebra of $\mathcal{A}$, it yields also $\neg\exists x \in A_0$ and hence $\exists(\neg\exists x) = \neg\exists x$. We obtain $\exists(\exists x \oplus \exists y) = \exists x \oplus \exists y$ in a similar way.

Altogether, $\mathcal{A} = (A; \oplus, \neg, \exists, 0)$ is a monadic basic algebra.

Assume now that $\mathcal{A}_0$ is an m-relatively complete subalgebra of $\mathcal{A}$. Let $x \in A$ and denote by $D = \{b \in A_0;\ x \leq b\}$. Then $\exists(x \oplus x) \geq x \oplus x$ as shown above and, by the condition (C), there exists a $v \in D$ with $\exists(x \oplus x) \geq v \oplus v$. Since $v \in D$ and $\exists x = \inf D$, thus $\exists(x \oplus x) \geq v \oplus v \geq \exists x \oplus \exists x$ by $\oplus$-monotonicity. Conversely, $\exists x \geq x$ yields $\exists x \oplus \exists x \geq x \oplus x$ by $\oplus$-monotonicity of $\mathcal{A}$ and, by (E4),

$$\exists x \oplus \exists x = \exists(\exists x \oplus \exists x) \geq \exists(x \oplus x).$$

□

Let $\mathcal{L}_i$ $(i \in I)$ be bounded lattices (or semilattices). By a *horizontal sum* is meant a lattice (semilattice) $\mathcal{L}$ which is a union of $\mathcal{L}_i$ $(i \in I)$ such that

$$\mathcal{L}_i \cap \mathcal{L}_j = \{0, 1\} \quad \text{for} \quad i \neq j.$$

Let $\mathcal{A}_i$ $(i \in I)$ be basic algebras. By a horizontal sum of $\mathcal{A}_i$ is meant a basic algebra $\mathcal{A}$ assigned to the lattice $\mathcal{L}$ which is the horizontal sum of the assigned lattices $\mathcal{L}(\mathcal{A}_i)$, $i \in I$.

Theorem 5 *Let a basic algebra* $\mathcal{A} = (A; \oplus, \neg, 0)$ *be a horizontal sum of basic algebras* $\mathcal{A}_i$ $(i \in I)$*. Let* $\exists_i$ *be an existential quantifier on* $\mathcal{A}_i$*, i.e. every* $(A_i; \oplus, \neg, \exists_i, 0)$ *is a monadic basic algebra. Let* $\exists\colon A \to A$ *be a mapping whose restriction on each* A_i *is equal to* $\exists_i$*. Then* $\mathcal{A}_\exists = (A; \oplus, \neg, \exists, 0)$ *is a monadic basic algebra.*

Proof We must check the axioms (e1) – (e4) for $\exists$ on the assigned lattice $\mathcal{L}(A)$. Trivially, we have $\exists 0 = 0$, $\exists(\exists x) = \exists x$ and $x \leq \exists x$. For $x, y \in A_i$ we have $\exists(x \vee y) = \exists x \vee \exists y$ by the definition. If $x \in A_i$, $y \in A_j$ for $i \neq j$ then $x \vee y = 1$ but also $\exists x \vee \exists y = 1$ thus, by (i) of Lemma 1, $1 = \exists(x \vee y) = \exists x \vee \exists y$. To check the condition (e4) is almost trivial since $x \leq y$ only if $x, y \in A_i$ and, inside A_i, it holds by the definition. □

Example 3 Consider the basic algebra $\mathcal{A} = (A; \oplus, \neg, 0)$, where $A = \{0, a, b, c, 1\}$, whose assigned lattice $\mathcal{L}(A)$ is in Fig. 3.

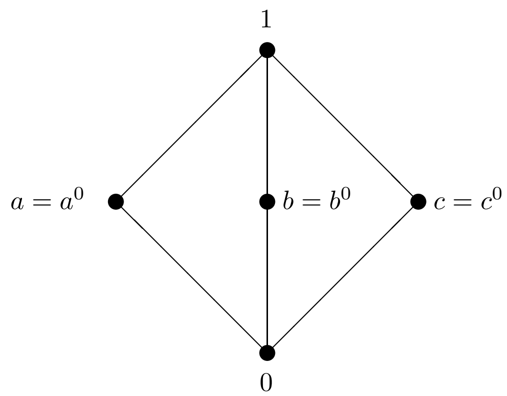


Fig. 3

Clearly, $\mathcal{A}$ is a horizontal sum of $\mathcal{H}$ (see Example 2) and the three element chain MV-algebra $\{0, c, 1\}$. Let us note that also $\mathcal{H}$ is a horizontal sum of two three element chain MV-algebras. One can easily verify that the mapping h defined by

$$h(a) = a, \quad h(b) = b, \quad h(c) = 1, \quad h(0) = 0, \quad h(1) = 1$$

is an e-retraction on $\mathcal{L}(A)$. In fact, h is composed by the e-retraction id on $\mathcal{H}$ and j on $\{0, c, 1\}$. The subalgebra $\mathcal{A}_0 = h(A)$ is clearly $\{0, a, b, 1\}$ (which is isomorphic to $\mathcal{H}$). It is an m-relatively complete subalgebra of $\mathcal{A}$.

Example 4 Consider again the basic algebra from Example 3. Let $B = \{0, c, 1\}$. It is a routine way to check that B is a relatively complete subalgebra of $\mathcal{A}$ and, by Theorem 4, it induces an existential quantifier. Of course, $\exists 0 = 0$, $\exists 1 = 1$ and we can easily compute

$$\exists a = \inf\{x \in B;\ a \leq x\} = 1,$$
$$\exists b = \inf\{x \in B;\ b \leq x\} = 1,$$
$$\exists c = \inf\{x \in B;\ c \leq x\} = c.$$

References

[1] Chajda, I., Emanovský, P.: *Bounded lattices with antitone involutions and properties of MV-algebras.* Discuss. Math., Gen. Algebra Appl. **24** (2004), 31–42.

[2] Chajda, I., Halaš, R.: *A basic algebra is an MV-algebra if and only if it is a BCC-algebra.* Intern. J. Theor. Phys., to appear.

[3] Chajda, I., Halaš, R., Kühr, J.: *Distributive lattices with sectionally antitone involutions.* Acta Sci. Math. (Szeged) **71** (2005), 19–33.

[4] Chajda, I., Halaš, R., Kühr, J.: *Many-valued quantum algebras.* Algebra Universalis, to appear.

[5] Chajda, I., Halaš, R., Kühr, J.: Semilattice Structures. *Heldermann Verlag, Lemgo, Germany*, 2007.

[6] Chajda, I., Kolařík, M.: *Independence of axiom system of basic algebras.* Soft Computing, to appear, DOI 10.1007/s00500-008-0291-2.

[7] Di Nola, A., Grigolia, R.: *On monadic MV-algebras.* Ann. Pure Appl. Logic **128** (2006), 212–218.

[8] Rachůnek, J., Švrček, F.: *Monadic bounded commutative residuated ℓ-monoids.* Order, to appear.

[9] Rutledge, J. D.: *On the definition of an infinitely-many-valued predicate calculus.* J. Symbolic Logic **25** (1960), 212–216.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 37–46

On Structure Space of Γ-Semigroups

S. CHATTOPADHYAY [1], S. KAR [2]

Department of Pure Mathematics, University of Calcutta
35, Ballygunge Circular Road, Kolkata-700019, India
[1] *e-mail: chatterjees04@yahoo.co.in*
[2] *e-mail: karsukhendu@yahoo.co.in*

(Received August 20, 2007)

Abstract

In this paper we introduce the notion of the structure space of Γ-semigroups formed by the class of uniformly strongly prime ideals. We also study separation axioms and compactness property in this structure space.

Key words: Γ-semigroup; uniformly strongly prime ideal; Noetherian Γ-semigroup, hull-kernel topology, structure space.

2000 Mathematics Subject Classification: 20M17

1 Introduction

In [4], L. Gillman studied "Rings with Hausdorff structure space" and in [7], C. W. Kohls studied "The space of prime ideals of a ring". In [1], M. R. Adhikari and M. K. Das studied 'Structure spaces of semirings'.

In [9], M. K. Sen and N. K. Saha introduced the notion of Γ-Semigroup. Some works on Γ-Semigroups may be found in [10], [8], [5], [6], [2] and [3].

In this paper we introduce and study the structure space of Γ-Semigroups. For this we consider the collection $\mathcal{A}$ of all proper uniformly strongly prime ideals of a Γ-Semigroup S and we give a topology $\tau_{\mathcal{A}}$ on $\mathcal{A}$ by means of closure operator defined in terms of intersection and inclusion relation among these ideals of the Γ-Semigroup S. We call the topological space $(\mathcal{A}, \tau_{\mathcal{A}})$—the structure space of the Γ-Semigroup S. We study separation axioms, compactness and connectedness in this structure space.

2 Preliminaries

Definition 2.1 Let $S = \{a, b, c, \dots\}$ and $\Gamma = \{\alpha, \beta, \gamma, \dots\}$ be two nonempty sets. S is called a Γ-semigroup if

(i) $a\alpha b \in S$, for all $\alpha \in \Gamma$ and $a, b \in S$ and

(ii) $(a\alpha b)\beta c = a\alpha(b\beta c)$, for all $a, b, c \in S$ and for all $\alpha, \beta \in \Gamma$.

S is said to be Γ-semigroup with zero if there exists an element $0 \in S$ such that $0\alpha a = a\alpha 0 = 0$ for all $\alpha \in \Gamma$.

Example 2.2 Let S be a set of all negative rational numbers. Obviously S is not a semigroup under usual product of rational numbers. Let

$$\Gamma = \{-\tfrac{1}{p} : p \text{ is prime}\}.$$

Let $a, b, c \in S$ and $\alpha \in \Gamma$. Now if $a\alpha b$ is equal to the usual product of rational numbers a, α, b, then $a\alpha b \in S$ and $(a\alpha b)\beta c = a\alpha(b\beta c)$. Hence S is a Γ-semigroup.

Definition 2.3 Let S be a Γ-semigroup and $\alpha \in \Gamma$. Then $e \in S$ is said to be an α-idempotent if $e\alpha e = e$. The set of all α-idempotents is denoted by E_α and we denote $\bigcup_{\alpha\in\Gamma} E_\alpha$ by $E(S)$. The elements of $E(S)$ are called idempotent element of S.

Definition 2.4 A nonempty subset I of a Γ-semigroup S is called an ideal if $I\Gamma S \subseteq I$ and $S\Gamma I \subseteq I$ where for subsets U, V of S and Δ of Γ, $U\Delta V = \{u\alpha v : u \in U, v \in V, \alpha \in \Delta\}$.

Definition 2.5 A nonempty subset I of a Γ-semigroup S is called an ideal if $I\Gamma S \subseteq I$ and $S\Gamma I \subseteq I$ where for subsets U, V of S and Δ of Γ, $U\Delta V = \{u\alpha v : u \in U, v \in V, \alpha \in \Delta\}$. An ideal I of S is called a proper ideal if $I \neq S$.

Definition 2.6 A proper ideal P of a Γ-Semigroup S is called a prime ideal of S if $A\Gamma B \subseteq P$ implies $A \subseteq P$ or $B \subseteq P$ for any two ideals A, B of S.

Definition 2.7 An ideal I of a Γ-semigroup S is said to be full if $E(S) \subseteq I$.

An ideal I of a Γ-semigroup S is said to be a prime full ideal if it is both prime and full.

Theorem 2.8 *Let S be a Γ-semigroup. For an ideal P of S, the following are equivalent.*

(i) *If A and B are ideals of S such that $A\Gamma B \subseteq P$ then either $A \subseteq P$ or $B \subseteq P$.*

(ii) *If $a\Gamma S\Gamma b \subseteq P$ then either $a \in P$ or $b \in P (a, b \in S)$*

(iii) *If I_1 and I_2 are two right ideals of S such that $I_1\Gamma I_2 \subseteq P$ then either $I_1 \subseteq P$ or $I_2 \subseteq P$.*

(iv) *If J_1 and J_2 are two left ideals of S such that $J_1\Gamma J_2 \subseteq P$ then either $J_1 \subseteq P$ or $J_2 \subseteq P$.*

Proof (i) $\Rightarrow$ (ii): Suppose $a\Gamma S\Gamma b \subseteq P$. Then $<a>\Gamma<a>\Gamma<b>\Gamma<b> \subseteq P$. Since $<a>\Gamma<a>$, $<b>\Gamma<b>$ are ideals of S, so by (i) we have either $<a>\Gamma<a> \subseteq P$ or $<b>\Gamma<b> \subseteq P$. By repeated uses of (i) we get $a \in <a> \subseteq P$ or $b \in <b> \subseteq P$.

(ii) $\Rightarrow$ (iii): Let $I_1\Gamma I_2 \subseteq P$. Let $I_1 \nsubseteq P$. Then there exists an element $a_1 \in I_1$ such that $a_1 \notin P$. Then for every $a_2 \in I_2$ we have $a_1\Gamma S\Gamma a_2 \subseteq I_1\Gamma I_2 \subseteq P$. Hence from (ii) $a_2 \in P$. Thus $I_2 \subseteq P$. Similarly (ii) implies (iv).

The proof is completed by observing that (i) is implied obviously either by (iii) or by (iv). □

Definition 2.9 An ideal P of a Γ-Semigroup S is called a uniformly strongly prime ideal(usp ideal) if S and Γ contain finite subsets F and Δ respectively such that $x\Delta F\Delta y \subseteq P$ implies that $x \in P$ or $y \in P$ for all $x, y \in S$.

Theorem 2.10 *Let S be a Γ-semigroup. Then every uniformly strongly prime ideal is a prime ideal.*

Proof Let P be a uniformly strongly prime ideal of S. Then S and Γ contain finite subsets F and Δ respectively such that $x\Delta F\Delta y \subseteq P$ implies that $x \in P$ or $y \in P$ for all $x, y \in S$. Now let $a\Gamma S\Gamma b \subseteq P$. Thus we have $a\Delta F\Delta b \subseteq a\Gamma S\Gamma b \subseteq P$ and hence we have $a \in P$ or $b \in P$. Hence P is prime ideal by Theorem 2.8. □

Throughout this paper S will always denote a Γ-Semigroup with zero and unless otherwise stated a Γ-Semigroup means a Γ-Semigroup with zero.

3 Structure space of Γ-semigroups

Suppose $\mathcal{A}$ is the collection of all uniformly strongly prime ideals of a Γ-Semigroup S. For any subset A of $\mathcal{A}$, we define

$$\overline{A} = \{I \in \mathcal{A} \colon \bigcap_{I_\alpha \in A} I_\alpha \subseteq I\}.$$

It is easy to see that $\overline{\emptyset} = \emptyset$.

Theorem 3.1 *Let A, B be any two subsets of $\mathcal{A}$. Then*

(i) $A \subseteq \overline{A}$

(ii) $\overline{\overline{A}} = \overline{A}$

(iii) $A \subseteq B \Longrightarrow \overline{A} \subseteq \overline{B}$

(iv) $\overline{A \cup B} = \overline{A} \cup \overline{B}$

Proof (i): Clearly, $\bigcap_{I_\alpha \in A} I_\alpha \subseteq I_\alpha$ for each α and hence $A \subseteq \overline{A}$.

(ii): By (i), we have $\overline{A} \subseteq \overline{\overline{A}}$. For converse part, let $I_\beta \in \overline{\overline{A}}$. Then $\bigcap_{I_\alpha \in \overline{A}} I_\alpha \subseteq I_\beta$. Now $I_\alpha \in \overline{A}$ implies that $\bigcap_{I_\gamma \in A} I_\gamma \subseteq I_\alpha$ for all $\alpha \in \Lambda$. Thus

$$\bigcap_{I_\gamma \in A} I_\gamma \subseteq \bigcap_{I_\alpha \in \overline{A}} I_\alpha \subseteq I_\beta \quad \text{i.e.} \quad \bigcap_{I_\gamma \in A} I_\gamma \subseteq I_\beta.$$

So $I_\beta \in \overline{A}$ and hence $\overline{\overline{A}} \subseteq \overline{A}$. Consequently, $\overline{\overline{A}} = \overline{A}$.

(iii): Suppose that $A \subseteq B$. Let $I_\alpha \in \overline{A}$. Then $\bigcap_{I_\beta \in A} I_\beta \subseteq I_\alpha$. Since $A \subseteq B$, it follows that

$$\bigcap_{I_\beta \in B} I_\beta \subseteq \bigcap_{I_\beta \in A} I_\beta \subseteq I_\alpha.$$

This implies that $I_\alpha \in \overline{B}$ and hence $\overline{A} \subseteq \overline{B}$.

(iv): Clearly, $\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$.

For the reverse part, let $I_\alpha \in \overline{A \cup B}$. Then $\bigcap_{I_\beta \in A \cup B} I_\beta \subseteq I_\alpha$.
It is easy to see that

$$\bigcap_{I_\beta \in A \cup B} I_\beta = \Big(\bigcap_{I_\beta \in A} I_\beta\Big) \cap \Big(\bigcap_{I_\beta \in B} I_\beta\Big).$$

Since $\bigcap_{I_\beta \in A} I_\beta$ and $\bigcap_{I_\beta \in B} I_\beta$ are ideals of S, we have

$$\Big(\bigcap_{I_\beta \in A} I_\beta\Big)\Gamma\Big(\bigcap_{I_\beta \in B} I_\beta\Big) \subseteq \Big(\bigcap_{I_\beta \in A} I_\beta\Big) \cap \Big(\bigcap_{I_\beta \in B} I_\beta\Big) = \bigcap_{I_\beta \in A \cup B} I_\beta \subseteq I_\alpha$$

Since every uniformly strongly prime ideal is prime, I_α is a prime ideal of S and hence either $\bigcap_{I_\beta \in A} I_\beta \subseteq I_\alpha$ or $\bigcap_{I_\beta \in B} I_\beta \subseteq I_\alpha$ i.e. either $I_\alpha \in \overline{A}$ or $I_\alpha \in \overline{B}$ i.e. $I_\alpha \in \overline{A} \cup \overline{B}$. Consequently, $\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$ and hence $\overline{A \cup B} = \overline{A} \cup \overline{B}$. □

Definition 3.2 The closure operator $A \longrightarrow \overline{A}$ gives a topology $\tau_{\mathcal{A}}$ on $\mathcal{A}$. This topology $\tau_{\mathcal{A}}$ is called the hull-kernel topology and the topological space $(\mathcal{A}, \tau_{\mathcal{A}})$ is called the structure space of the Γ-Semigroup S.

Let I be a ideal of a Γ-Semigroup S. We define

$$\Delta(I) = \{I' \in \mathcal{A} \colon I \subseteq I'\} \quad \text{and} \quad C\Delta(I) = \mathcal{A} \setminus \Delta(I) = \{I' \in \mathcal{A} \colon I \nsubseteq I'\}.$$

Now we have the following result:

Proposition 3.3 *Any closed set in $\mathcal{A}$ is of the form $\Delta(I)$, where I is a ideal of a Γ-Semigroup S.*

Proof Let $\overline{A}$ be any closed set in $\mathcal{A}$, where $A \subseteq \mathcal{A}$. Let $A = \{I_\alpha \colon \alpha \in \Lambda\}$ and $I = \bigcap_{I_\alpha \in A} I_\alpha$. Then I is a ideal of S. Let $I' \in \overline{A}$. Then $\bigcap_{I_\alpha \in A} I_\alpha \subseteq I'$. This implies that $I \subseteq I'$. Consequently, $I' \in \Delta(I)$. So $\overline{A} \subseteq \Delta(I)$.

Conversely, let $I' \in \Delta(I)$. Then $I \subseteq I'$ i.e. $\bigcap_{I_\alpha \in A} I_\alpha \subseteq I'$. Consequently, $I' \in \overline{A}$ and hence $\Delta(I) \subseteq \overline{A}$. Thus $\overline{A} = \Delta(I)$. □

Corollary 3.4 *Any open set in $\mathcal{A}$ is of the form $C\Delta(I)$, where I is an ideal of S.*

Let S be a Γ-Semigroup and $a \in S$. We define

$$\Delta(a) = \{I \in \mathcal{A} \colon a \in I\} \quad \text{and} \quad C\Delta(a) = \mathcal{A} \setminus \Delta(a) = \{I \in \mathcal{A} \colon a \notin I\}.$$

Then we have the following result:

Proposition 3.5 *$\{C\Delta(a) \colon a \in S\}$ forms an open base for the hull-kernel topology $\tau_{\mathcal{A}}$ on $\mathcal{A}$.*

Proof Let $U \in \tau_{\mathcal{A}}$. Then $U = C\Delta(I)$, where I is an ideal of S. Let $J \in U = C\Delta(I)$. Then $I \not\subseteq J$. This implies that there exists $a \in I$ such that $a \notin J$. Thus $J \in C\Delta(a)$. Now it remains to show that $C\Delta(a) \subset U$. Let $K \in C\Delta(a)$. Then $a \notin K$. This implies that $I \not\subseteq K$. Consequently, $K \in U$ and hence $C\Delta(a) \subset U$. So we find that $J \in C\Delta(a) \subset U$. Thus $\{C\Delta(a) \colon a \in S\}$ is an open base for the hull-kernel topology $\tau_{\mathcal{A}}$ on $\mathcal{A}$.

Theorem 3.6 *The structure space $(\mathcal{A}, \tau_{\mathcal{A}})$ is a T_0-space.*

Proof Let I_1 and I_2 be two distinct elements of $\mathcal{A}$. Then there is an element a either in $I_1 \setminus I_2$ or in $I_2 \setminus I_1$. Suppose that $a \in I_1 \setminus I_2$. Then $C\Delta(a)$ is a neighbourhood of I_2 not containing I_1. Hence $(\mathcal{A}, \tau_{\mathcal{A}})$ is a T_0-space. □

Theorem 3.7 *$(\mathcal{A}, \tau_{\mathcal{A}})$ is a T_1-space if and only if no element of $\mathcal{A}$ is contained in any other element of $\mathcal{A}$.*

Proof Let $(\mathcal{A}, \tau_{\mathcal{A}})$ be a T_1-space. Suppose that I_1 and I_2 be any two distinct elements of $\mathcal{A}$. Then each of I_1 and I_2 has a neighbourhood not containing the other. Since I_1 and I_2 are arbitrary elements of $\mathcal{A}$, it follows that no element of $\mathcal{A}$ is contained in any other element of $\mathcal{A}$.

Conversely, suppose that no element of $\mathcal{A}$ is contained in any other element of $\mathcal{A}$. Let I_1 and I_2 be any two distinct elements of $\mathcal{A}$. Then by hypothesis, $I_1 \not\subset I_2$ and $I_2 \not\subset I_1$. This implies that there exist $a, b \in S$ such that $a \in I_1$ but $a \notin I_2$ and $b \in I_2$ but $b \notin I_1$. Consequently, we have $I_1 \in C\Delta(b)$ but $I_1 \notin C\Delta(a)$ and $I_2 \in C\Delta(a)$ but $I_2 \notin C\Delta(b)$ i.e. each of I_1 and I_2 has a neighbourhood not containing the other. Hence $(\mathcal{A}, \tau_{\mathcal{A}})$ is a T_1-space. □

Corollary 3.8 *Let $\mathcal{M}$ be the set of all proper maximal ideals of a Γ-Semigroup S with unities. Then $(\mathcal{M}, \tau_{\mathcal{M}})$ is a T_1-space, where $\tau_{\mathcal{M}}$ is the induced topology on $\mathcal{M}$ from $(\mathcal{A}, \tau_{\mathcal{A}})$.*

Theorem 3.9 *$(\mathcal{A}, \tau_{\mathcal{A}})$ is a Hausdorff space if and only if for any two distinct pair of elements I, J of $\mathcal{A}$, there exist $a, b \in S$ such that $a \notin I$, $b \notin J$ and there does not exist any element K of $\mathcal{A}$ such that $a \notin K$ and $b \notin K$.*

Proof Let $(\mathcal{A}, \tau_{\mathcal{A}})$ be a Hausdorff space. Then for any two distinct elements I, J of $\mathcal{A}$, there exist basic open sets $C\Delta(a)$ and $C\Delta(b)$ such that $I \in C\Delta(a)$, $J \in C\Delta(b)$ and $C\Delta(a) \cap C\Delta(b) = \emptyset$. Now $I \in C\Delta(a)$ and $J \in C\Delta(b)$ imply that $a \notin I$ and $b \notin J$. If possible, let K be any element of $\mathcal{A}$ such that $a \notin K$ and $b \notin K$. Then $K \in C\Delta(a)$, $K \in C\Delta(b)$ and hence $K \in C\Delta(a) \cap C\Delta(b)$, a contradiction, since $C\Delta(a) \cap C\Delta(b) = \emptyset$. Thus there does not exist any element K of $\mathcal{A}$ such that $a \notin K$ and $b \notin K$.

Conversely, suppose that the given condition holds and $I, J \in \mathcal{A}$ such that $I \neq J$. Let $a, b \in S$ be such that $a \notin I$, $b \notin J$ and there does not exist any K of $\mathcal{A}$ such that $a \notin K$ and $b \notin K$. Then $I \in C\Delta(a)$, $J \in C\Delta(b)$ and $C\Delta(a) \cap C\Delta(b) = \emptyset$. This implies that $(\mathcal{A}, \tau_{\mathcal{A}})$ is a Hausdorff space. □

Corollary 3.10 *If $(\mathcal{A}, \tau_{\mathcal{A}})$ is a Hausdorff space, then no proper uniformly strongly prime ideal contains any other proper uniformly strongly prime ideal. If $(\mathcal{A}, \tau_{\mathcal{A}})$ contains more than one element, then there exist $a, b \in S$ such that $\mathcal{A} = C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$, where I is the ideal generated by a, b.*

Proof Suppose that $(\mathcal{A}, \tau_{\mathcal{A}})$ is a Hausdorff space. Since every Hausdorff space is a T_1-space, $(\mathcal{A}, \tau_{\mathcal{A}})$ is a T_1-space. Hence by Theorem 3.7, it follows that no proper uniformly strongly prime ideal contains any other proper uniformly strongly prime ideal. Now let $J, K \in \mathcal{A}$ be such that $J \neq K$. Since $(\mathcal{A}, \tau_{\mathcal{A}})$ is a Hausdorff space, there exist basic open sets $C\Delta(a)$ and $C\Delta(b)$ such that $J \in C\Delta(a)$, $K \in C\Delta(b)$ and $C\Delta(a) \cap C\Delta(b) = \emptyset$. Let I be the ideal generated by a, b. Then I is the smallest ideal containing a and b. Let $K \in \mathcal{A}$. Then either $a \in K$, $b \notin K$ or $a \notin K$, $b \in K$ or $a, b \in K$. The case $a \notin K$, $b \notin K$ is not possible, since $C\Delta(a) \cap C\Delta(b) = \emptyset$. Now in the first case, $K \in C\Delta(b)$ and hence $\mathcal{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. In the second case, $K \in C\Delta(a)$ and hence $\mathcal{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. In the third case, $K \in \Delta(I)$ and hence $\mathcal{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. So we find that $\mathcal{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. Again, clearly $C\Delta(a) \cup C\Delta(b) \cup \Delta(I) \subseteq \mathcal{A}$. Hence $\mathcal{A} = C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. □

Theorem 3.11 *$(\mathcal{A}, \tau_{\mathcal{A}})$ is a regular space if and only if for any $I \in \mathcal{A}$ and $a \notin I$, $a \in S$, there exist an ideal J of S and $b \in S$ such that $I \in C\Delta(b) \subseteq \Delta(J) \subseteq C\Delta(a)$.*

Proof Let $(\mathcal{A}, \tau_{\mathcal{A}})$ be a regular space. Let $I \in \mathcal{A}$ and $a \notin I$. Then $I \in C\Delta(a)$ and $\mathcal{A} \setminus C\Delta(a)$ is a closed set not containing I. Since $(\mathcal{A}, \tau_{\mathcal{A}})$ is a regular space, there exist disjoints open sets U and V such that $I \in U$ and $\mathcal{A} \setminus C\Delta(a) \subseteq V$. This implies that $\mathcal{A} \setminus V \subseteq C\Delta(a)$. Since V is open, $\mathcal{A} \setminus V$ is closed and hence there exists an ideal J of S such that $\mathcal{A} \setminus V = \Delta(J)$, by Proposition 3.3. So we find that $\Delta(J) \subseteq C\Delta(a)$. Again, since $U \cap V = \emptyset$, we have $V \subseteq \mathcal{A} \setminus U$. Since U is open, $\mathcal{A} \setminus U$ is closed and hence there exists an ideal K of S such that $\mathcal{A} \setminus U = \Delta(K)$ i.e. $V \subseteq \Delta(K)$. Since $I \in U$, $I \notin \mathcal{A} \setminus U = \Delta(K)$. This implies that $K \nsubseteq I$. Thus there exists $b \in K (\subset S)$ such that $b \notin I$. So $I \in C\Delta(b)$. Now we show that $V \subseteq \Delta(b)$. Let $M \in V \subseteq \Delta(K)$. Then $K \subseteq M$. Since $b \in K$, it follows that $b \in M$ and hence $M \in \Delta(b)$. Consequently, $V \subseteq \Delta(b)$. This implies that $\mathcal{A} \setminus \Delta(b) \subseteq \mathcal{A} \setminus V = \Delta(J) \Longrightarrow C\Delta(b) \subseteq \Delta(J)$. Thus we find that $I \in C\Delta(b) \subseteq \Delta(J) \subseteq C\Delta(a)$.

Conversely, suppose that the given condition holds. Let $I \in \mathcal{A}$ and $\Delta(K)$ be any closed set not containing I. Since $I \notin \Delta(K)$, we have $K \not\subset I$. This implies that there exists $a \in K$ such that $a \notin I$. Now by the given condition, there exists an ideal J of S and $b \in S$ such that $I \in C\Delta(b) \subseteq \Delta(J) \subseteq C\Delta(a)$. Since $a \in K$, $C\Delta(a) \cap \Delta(K) = \emptyset$. This implies that $\Delta(K) \subseteq \mathcal{A} \setminus C\Delta(a) \subseteq \mathcal{A} \setminus \Delta(J)$. Since $\Delta(J)$ is a closed set, $\mathcal{A} \setminus \Delta(J)$ is an open set containing the closed set $\Delta(K)$. Clearly, $C\Delta(b) \cap (\mathcal{A} \setminus \Delta(J)) = \emptyset$. So we find that $C\Delta(b)$ and $\mathcal{A} \setminus \Delta(J)$ are two disjoints open sets containing I and $\Delta(K)$ respectively. Consequently, $(\mathcal{A}, \tau_{\mathcal{A}})$ is a regular space. □

Theorem 3.12 *$(\mathcal{A}, \tau_{\mathcal{A}})$ is a compact space if and only if for any collection $\{a_\alpha\}_{\alpha \in \Lambda} \subset S$ there exists a finite subcollection $\{a_i \colon i = 1, 2, \ldots, n\}$ in S such that for any $I \in \mathcal{A}$, there exists a_i such that $a_i \notin I$.*

Proof Let $(\mathcal{A}, \tau_{\mathcal{A}})$ be a compact space. Then the open cover $\{C\Delta(a_\alpha) \colon a_\alpha \in S\}$ of $(\mathcal{A}, \tau_{\mathcal{A}})$ has a finite subcover $\{C\Delta(a_i) \colon i = 1, 2, \ldots, n\}$. Let I be any element of $\mathcal{A}$. Then $I \in C\Delta(a_i)$ for some $a_i \in S$. This implies that $a_i \notin I$. Hence $\{a_i \colon i = 1, 2, \ldots, n\}$ is the required finite subcollection of elements of S such that for any $I \in \mathcal{A}$, there exists a_i such that $a_i \notin I$.

Conversely, suppose that the given condition holds. Let $\{C\Delta(a_\alpha) \colon a_\alpha \in S\}$ be an open cover of $\mathcal{A}$. Suppose to the contrary that no finite subcollection of $\{C\Delta(a_\alpha) \colon a_\alpha \in S\}$ covers $\mathcal{A}$. This means that for any finite set $\{a_1, a_2, \ldots, a_n\}$ of elements of S,

$$
\begin{array}{ll}
 & C\Delta(a_1) \cup C\Delta(a_2) \cup \ldots \cup C\Delta(a_n) \neq \mathcal{A} \\
\Longrightarrow & \Delta(a_1) \cap \Delta(a_2) \cap \ldots \cap \Delta(a_n) \neq \emptyset \\
\Longrightarrow & \text{there exists } I \in \mathcal{A} \text{ such that } I \in \Delta(a_1) \cap \Delta(a_2) \cap \ldots \cap \Delta(a_n) \\
\Longrightarrow & a_1, a_2, \ldots, a_n \in I, \text{ which contradicts our hypothesis .}
\end{array}
$$

So the open cover $\{C\Delta(a_\alpha) \colon a_\alpha \in S\}$ has a finite subcover and hence $(\mathcal{A}, \tau_{\mathcal{A}})$ is compact.

Corollary 3.13 *If S is finitely generated, then $(\mathcal{A}, \tau_{\mathcal{A}})$ is a compact space.*

Proof Let $\{a_i \colon i = 1, 2, \ldots, n\}$ be a finite set of generators of S. Then for any $I \in \mathcal{A}$, there exists a_i such that $a_i \notin I$, since I is a proper uniformly strongly prime ideal of S. Hence by Theorem 3.12, $(\mathcal{A}, \tau_{\mathcal{A}})$ is a compact space. □

Definition 3.14 A Γ-Semigroup S is called a Noetherian Γ-Semigroup if it satisfies the ascending chain condition on ideals of S i.e. if $I_1 \subseteq I_2 \subseteq \ldots \subseteq I_n \subseteq \ldots$ is an ascending chain of ideals of S, then there exists a positive integer m such that $I_n = I_m$ for all $n \geq m$.

Theorem 3.15 *If S is a Noetherian Γ-Semigroup, then $(\mathcal{A}, \tau_{\mathcal{A}})$ is countably compact.*

Proof Let $\{\Delta(I_n)\}_{n=1}^{\infty}$ be a countable collection of closed sets in $\mathcal{A}$ with finite intersection property (FIP). Let us consider the following ascending chain of prime ideals of S: $<I_1> \subseteq <I_1 \cup I_2> \subseteq <I_1 \cup I_2 \cup I_3> \subseteq \ldots$

Since S is a Noetherian Γ-Semigroup, there exists a positive integer m such that $<I_1 \cup I_2 \cup \ldots \cup I_m> = <I_1 \cup I_2 \cup \ldots \cup I_{m+1}> = \ldots$

Thus it follows that $<I_1 \cup I_2 \cup \ldots \cup I_m> \in \bigcap_{n=1}^{\infty} \Delta(I_n)$. Consequently, $\bigcap_{n=1}^{\infty} \Delta(I_n) \neq \emptyset$ and hence $(\mathcal{A}, \tau_{\mathcal{A}})$ is countably compact. □

Corollary 3.16 *If S is a Noetherian Γ-Semigroup and $(\mathcal{A}, \tau_{\mathcal{A}})$ is second countable, then $(\mathcal{A}, \tau_{\mathcal{A}})$ is compact.*

Proof Proof follows from Theorem 3.15 and the fact that a second countable space is compact if it is countably compact. □

Remark 3.17 Let $\{I_\alpha\}$ be a collection of prime ideals of a Γ-semigroup S. Then $\bigcap I_\alpha$ is an ideal of S but it may not be a prime ideal of S, in general.

However; in particular, we have the following result:

Proposition 3.18 *Let $\{I_\alpha\}$ be a collection of prime ideals of a Γ-semigroup S such that $\{I_\alpha\}$ forms a chain. Then $\bigcap I_\alpha$ is a prime ideal of S.*

Proof Clearly, $\bigcap I_\alpha$ is an ideal of S. Let $A\Gamma B \subseteq \bigcap I_\alpha$ for any two ideals A, B of S. If possible, let $A, B \nsubseteq \bigcap I_\alpha$. Then there exist α and β such that $A \nsubseteq I_\alpha$ and $B \nsubseteq I_\beta$. Since I_α is a chain, let $I_\alpha \subseteq I_\beta$. This implies that $B \nsubseteq I_\alpha$. Since $A\Gamma B \subseteq I_\alpha$ and I_α is prime, we must have either $A \subseteq I_\alpha$ or $B \subseteq I_\alpha$, a contradiction. Therefore, either $A \subseteq \bigcap I_\alpha$ or $B \subseteq \bigcap I_\alpha$. Consequently, $\bigcap I_\alpha$ is a prime ideal of S. □

Definition 3.19 The structure space $(\mathcal{A}, \tau_{\mathcal{A}})$ is called irreducible if for any decomposition $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$, where $\mathcal{A}_1$ and $\mathcal{A}_2$ are closed subsets of $\mathcal{A}$, we have either $\mathcal{A} = \mathcal{A}_1$ or $\mathcal{A} = \mathcal{A}_2$.

Theorem 3.20 *Let A be a closed subset of $\mathcal{A}$. Then A is irreducible if and only if $\bigcap_{I_\alpha \in A} I_\alpha$ is a prime ideal of S.*

Proof Let A be irreducible. Let P and Q be two ideals of S such that $P\Gamma Q \subseteq \bigcap_{I_\alpha \in A} I_\alpha$. Then $P\Gamma Q \subseteq I_\alpha$ for all α. Since I_α is prime, either $P \subseteq I_\alpha$ or $Q \subseteq I_\alpha$ which implies for $I_\alpha \in A$ either $I_\alpha \in \{\overline{P}\}$ or $I_\alpha \in \{\overline{Q}\}$. Hence $A = (A \cap \overline{P}) \cup (A \cap \overline{Q})$. Since A is irreducible and $(A \cap \overline{P})$, $(A \cap \overline{Q})$ are closed, it follows that $A = A \cap \overline{P}$ or $A = A \cap \overline{Q}$ and hence $A \subseteq \overline{P}$ or $A \subseteq \overline{Q}$. This implies that $P \subseteq \bigcap_{I_\alpha \in A} I_\alpha$ or $Q \subseteq \bigcap_{I_\alpha \in A} I_\alpha$. Consequently, $\bigcap_{I_\alpha \in A} I_\alpha$ is a prime ideal of S.

Conversely, suppose that $\bigcap_{I_\alpha \in A} I_\alpha$ is a prime ideal of S. Let $A = A_1 \cup A_2$, where A_1 and A_2 are closed subsets of A. Then $\bigcap_{I_\alpha \in A} I_\alpha \subseteq \bigcap_{I_\alpha \in A_1} I_\alpha$ and $\bigcap_{I_\alpha \in A} I_\alpha \subseteq \bigcap_{I_\alpha \in A_2} I_\alpha$. Also

$$\bigcap_{I_\alpha \in A} I_\alpha = \bigcap_{I_\alpha \in A_1 \cup A_2} I_\alpha = \Big(\bigcap_{I_\alpha \in A_1} I_\alpha\Big) \cap \Big(\bigcap_{I_\alpha \in A_2} I_\alpha\Big).$$

Now

$$\Big(\bigcap_{I_\alpha \in A_1} I_\alpha\Big)\Gamma\Big(\bigcap_{I_\alpha \in A_2} I_\alpha\Big) \subseteq \Big(\bigcap_{I_\alpha \in A_1} I_\alpha\Big) \quad \text{and} \quad \Big(\bigcap_{I_\alpha \in A_1} I_\alpha\Big)\Gamma\Big(\bigcap_{I_\alpha \in A_2} I_\alpha\Big) \subseteq \Big(\bigcap_{I_\alpha \in A_2} I_\alpha\Big).$$

Thus we have

$$\Big(\bigcap_{I_\alpha \in A_1} I_\alpha\Big)\Gamma\Big(\bigcap_{I_\alpha \in A_2} I_\alpha\Big) \subseteq \Big(\bigcap_{I_\alpha \in A_1} I_\alpha\Big) \cap \Big(\bigcap_{I_\alpha \in A_2} I_\alpha\Big).$$

Since $\bigcap_{I_\alpha \in A} I_\alpha$ is prime, it follows that either

$$\bigcap_{I_\alpha \in A_1} I_\alpha \subseteq \bigcap_{I_\alpha \in A} I_\alpha \quad \text{or} \quad \bigcap_{I_\alpha \in A_2} I_\alpha \subseteq \bigcap_{I_\alpha \in A} I_\alpha.$$

So we find that

$$\bigcap_{I_\alpha \in A} I_\alpha = \bigcap_{I_\alpha \in A_1} I_\alpha \quad \text{or} \quad \bigcap_{I_\alpha \in A} I_\alpha = \bigcap_{I_\alpha \in A_2} I_\alpha.$$

Let $I_\beta \in A$. Then we have

$$\bigcap_{I_\alpha \in A_1} I_\alpha \subseteq I_\beta \quad \text{or} \quad \bigcap_{I_\alpha \in A_2} I_\alpha \subseteq I_\alpha.$$

Since $A_1, A_2 \subseteq A$, so either $I_\alpha \subseteq I_\beta$ for all $I_\alpha \in A_1$ or $I_\alpha \subseteq I_\beta$ for all $I_\alpha \in A_2$. Thus $I_\beta \in \overline{A_1} = A_1$ or $I_\beta \in \overline{A_2} = A_2$, since A_1 and A_2 are closed. i.e. $A = A_1$ or $A = A_2$. □

Let $\mathcal{C}$ be the collection of all uniformly strongly prime full ideals of a Γ-semigroup S. Then we see that $\mathcal{C}$ is a subset of $\mathcal{A}$ and hence $(\mathcal{C}, \tau_{\mathcal{C}})$ is a topological space, where $\tau_{\mathcal{C}}$ is the subspace topology.

In general, $(\mathcal{A}, \tau_{\mathcal{A}})$ is not compact and connected. But in particular, for the topological space $(\mathcal{C}, \tau_{\mathcal{C}})$, we have the following results:

Theorem 3.21 *$(\mathcal{C}, \tau_{\mathcal{C}})$ is a compact space.*

Proof Let $\{\Delta(I_\alpha) \colon \alpha \in \Lambda\}$ be any collection of closed sets in $\mathcal{C}$ with finite intersection property. Let I be the uniformly strongly prime full ideal generated by $E(S)$. Since any uniformly strongly prime full ideal J contains $E(S)$, J contains I. Hence $I \in \bigcap_{\alpha \in \Lambda} \Delta(I_\alpha) \neq \emptyset$. Consequently, $(\mathcal{C}, \tau_{\mathcal{C}})$ is a compact space. □

Theorem 3.22 *$(\mathcal{C}, \tau_{\mathcal{C}})$ is a connected space.*

Proof Let I be the uniformly strongly prime ideal generated by $E(S)$. Since any uniformly strongly prime full ideal J contains $E(S)$, J contains I. Hence I belongs to any closed set $\Delta(I')$ of $\mathcal{C}$. Consequently, any two closed sets of $\mathcal{C}$ are not disjoint. Hence $(\mathcal{C}, \tau_{\mathcal{C}})$ is a connected space. □

References

[1] Adhikari, M. R., Das, M. K.: *Structure Spaces of Semirings.* Bull. Cal. Math. Soc. **86** (1994), 313–317.

[2] Dutta, T. K., Chattopadhyay, S.: *On uniformly strongly prime* Γ*-semigroup.* Analele Stiintifice Ale Universitatii "AL. I. CUZA" IASI Tomul LII, s.I, Math., 2006, f.2, 325–335.

[3] Dutta, T. K., Chattopadhyay, S.: *On Uniformly strongly prime* Γ*-semigroup (2).* Accepted.

[4] Gillman, L.: *Rings with Hausdorff structure space.* Fund. Math. **45** (1957), 1–16.

[5] Chattopadhyay, S.: *Right Orthodox* Γ*-semigroup.* Southeast Asian Bull. of Mathematics **29** (2005), 23–30.

[6] Chattopadhyay, S.: *Right inverse* Γ*-semigroup.* Bull. Cal. Math. Soc **93**, 6 (2001), 435–442.

[7] Kohls, C. W.: *The space of prime ideals of a ring.* Fund. Math. **45** (1957), 17–27.

[8] Saha, N. K.: *On* Γ*-semigroup III.* Bull. Cal. Math. Soc. **80** (1988), 1–12.

[9] Sen, M. K., Saha, N. K.: *On* Γ*-semigroup I.* Bull. Cal. Math. Soc. **78** (1986), 180–186.

[10] Sen, M. K., Chattopadhyay, S.: *Semidirect Product of a Monoid and a* Γ*-semigroup.* East-West J. of Math. **6**, 2 (2004), 131–138.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 47–53

Frankl's Conjecture for Large Semimodular and Planar Semimodular Lattices*

Gábor CZÉDLI [1], E. Tamás SCHMIDT [2]

[1] *Bolyai Institute, University of Szeged,*
Aradi vértanúk tere 1, Szeged, Hungary
e-mail: czedli@math.u-szeged.hu

[2] *Mathematical Institute of the Budapest University of Technology and Economics*
Műegyetem rkp. 3, H-1521 Budapest, Hungary
e-mail: schmidt@math.bme.hu

(Received April 11, 2008)

Abstract

A lattice L is said to satisfy (the lattice theoretic version of) Frankl's conjecture if there is a join-irreducible element $f \in L$ such that at most half of the elements x of L satisfy $f \leq x$. Frankl's conjecture, also called as union-closed sets conjecture, is well-known in combinatorics, and it is equivalent to the statement that every finite lattice satisfies Frankl's conjecture.

Let m denote the number of nonzero join-irreducible elements of L. It is well-known that L consists of at most 2^m elements. Let us say that L is large if it has more than $5 \cdot 2^{m-3}$ elements. It is shown that every large semimodular lattice satisfies Frankl's conjecture. The second result states that every finite semimodular planar lattice L satisfies Frankl's conjecture. If, in addition, L has at least four elements and its largest element is join-reducible then there are at least two choices for the above-mentioned f.

Key words: Union-closed sets; Frankl's conjecture; lattice, semimodularity; planar lattice.

2000 Mathematics Subject Classification: 05A05, sec.: 06E99

*This research was partially supported by the NFSR of Hungary (OTKA), grant no. T 049433 and K 60148.

Given an m-element finite set $A = \{a_1, \ldots, a_m\}$, $m \geq 3$, a *family* (or, in other words, a set) $\mathcal{F}$ of at least two subsets of A, i.e. $\mathcal{F} \subseteq P(A)$, is called a *union-closed family* (over A) if $X \cup Y \in \mathcal{F}$ whenever $X, Y \in \mathcal{F}$. It was Peter Frankl in 1979 (cf. Frankl [9]) who formulated the following conjecture, now called as *Frankl's conjecture* or *the union-closed sets conjecture*: if $\mathcal{F}$ is as above then there exists an element of A which is contained in at least half of the members of $\mathcal{F}$. In spite of at least three dozen papers, cf. the bibliography given in [8], this conjecture is still open.

Now let L be a finite lattice. As usual, the set of its nonzero join-irreducible elements will be denoted by $J(L)$. We say that L satisfies (the lattice theoretic version of) Frankl's conjecture if $|L| = 1$ or there is an $f \in J(L)$ such that for the principal filter $\uparrow f = \{x \in L : f \leq x\}$ we have $|\uparrow f| \leq |L|/2$. Stanley [17] and Poonen [14] or Abe and Nakano [3] have shown that (the original) Frankl's conjecture is true if and only if all finite lattices satisfy (the lattice theoretic) Frankl's conjecture. (For details one can also see [6].) This fact has initiated a series of lattice theoretical results given by Abe and Nakano [1], [2], [3], [4], Herrmann and Langsdorf [13], and Reinhold [15], and two combinatorial results achieved by means of lattices, cf. [6] and [8]. In particular, lower semimodular lattices satisfy Frankl's conjecture by [15], and the method of [15] makes it clear that the situation for (upper) semimodular lattices is much harder. In fact, it is (and it remains) unknown if semimodular lattices satisfy Frankl's conjecture. The goal of the present paper is to present two subclasses of the class of finite semimodular lattices such that every lattice L in these subclasses satisfies Frankl's conjecture; in fact, L usually satisfies the conjecture in a bit stronger form.

For elements x and y of a lattice L, let $x \preceq y$ denote the "covers or equals" relation. That is, $x \preceq y$ iff $x \leq y$ and there is no $z \in L$ with $x < z < y$. Recall that L is called (upper) *semimodular* if, for any $a, b, c \in L$, $a \preceq b$ implies $a \vee c \preceq b \vee c$. Let $J(L)$ denote the set of non-zero join-irreducible elements of L, and let $m = |J(L)|$. Since each element of L is the join of a subset of $J(L)$, L has at most 2^m elements. Strengthening a former result of Gao and Yu [10], it is shown in [6] that L satisfies Frankl's conjecture provided $|L| \geq 2^m - 2^{m/2}$. In the semimodular case we can prove more. For simplicity, finite lattices L with more than $5 \cdot 2^{m-3} = 2^m - \frac{3}{8} \cdot 2^m$ elements will be called *large*. The *height* $h(x)$ of an element $x \in L$ is the length (number of elements minus one) of any maximal chain in the principal ideal $\downarrow x$. (This makes sense, for any two maximal chains has the same length by semimodularity.)

Theorem 1 *Let L be a finite semimodular lattice. If L is large in the sense $|L| > 5 \cdot 2^{m-3}$, where $m = |J(L)|$, then L satisfies Frankl's conjecture.*

Proof Let $A(L)$ denote the set of atoms of L.

First we show that $|J(L) \setminus A(L)| \leq 1$. By way of contradiction, assume that a_1 and a_2 are distinct elements of $J(L) \setminus A(L)$. Let $a_3, \ldots, a_m$ be the rest of nonzero join-irreducible elements, i.e., $J(L) = \{a_1, a_2, \ldots, a_m\}$. Let B_m be the boolean lattice with atoms $x_1, \ldots, x_m$, and consider both $B_m = (B_m; \vee, 0)$ and

$L = (L; \vee, 0)$ as join-semilattices with 0. Since B_m is the free join-semilattice with 0, there is a surjective homomorphism $\varphi\colon B_m \to L$, $x_i \mapsto a_i$. Let Θ denote the kernel of φ. Then, for $i = 1, 2$, the Θ-class $[x_i]$ of x_i is not a singleton, for otherwise a_i would be an atom. Since $a_i \neq 0$, we conclude that $0 \notin [x_i]$. Since Θ-classes are convex subsemilattices, there are elements $y_1 \in [x_1]$ and $y_2 \in [x_2]$ such that $y_1 \succ x_1$ and $y_2 \succ x_2$. They are distinct, for $a_1 \neq a_2$. Let $z = y_1 \wedge y_2$; it is an atom or the zero of B_m.

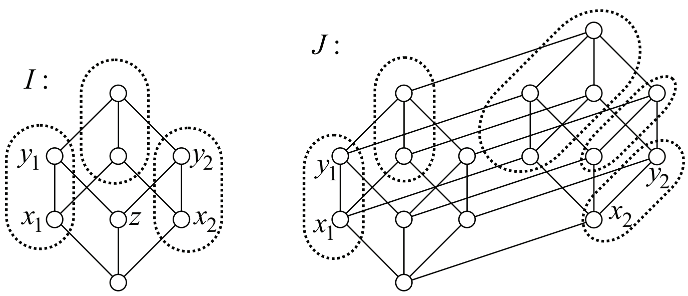


Fig. 1: Two ideals in B_m

First assume that z is an atom, and consider the ideal $I = \downarrow(y_1 \vee y_2)$ in B_m, cf. Figure 1. Let K denote the subsemilattice generated by those atoms of B_m that are not in I; K is not indicated in the figure. It follows from $(x_1, y_1), (x_2, y_2) \in \Theta$ that the restriction $\Theta|_I$ to I includes the semilattice congruence indicated in the figure. Hence Θ collapses I to five or less elements. For $u \in K$, let $u \vee I = \{u \vee t\colon t \in I\}$. If $(t_1, t_2) \in \Theta|_I$ then $(u \vee t_1, u \vee t_2) \in \Theta$. Hence Θ collapses $u \vee I$ to five or less elements. Now B_m is the union of the pairwise disjoint subsets $u \vee I$, $u \in B_m$. Therefore $L \cong B_m/\Theta$ consists of at most $5 \cdot |K| = 5 \cdot 2^{m-3}$ elements, which contradicts the assumption that L is large.

Secondly, assume that $z = 0$, and consider the ideal $J = \downarrow(y_1 \vee y_2)$, cf. Figure 1. Then the same argument as above gives $|L| \leq 9 \cdot 2^{m-4} < 5 \cdot 2^{m-3}$, a contradiction again. This proves that $|J(L) \setminus A(L)| \leq 1$.

Now, let us recall a well-known fact on semimodular lattices. An n-element subset $U = \{c_1, \dots, c_n\}$ of $A(L)$ is called *independent* if the sublattice $[U]$ generated by U is boolean with $A([U]) = U$. It is well-known, cf. e.g., Theorem IV.2.4 in Grätzer [11], that U is independent if and only if

$$(c_1 \vee \cdots \vee c_i) \wedge c_{i+1} = 0 \ \text{ for } \ i = 1, 2, \dots, n-1. \tag{1}$$

We need another, much easier version of independence: $U \subseteq J(L)$ will be called an *irredundant set* if $u \not\leq \bigvee(U \setminus \{u\})$ for every $u \in U$. In other words, $U = \{c_1, \dots, c_n\}$ is independent if no joinand can be omitted from $c_1 \vee \cdots \vee c_n$.

Now, armed with $|J(L) \setminus A(L)| \leq 1$, let us introduce some new notations. If $|J(L) \setminus A(L)| = 1$, then let a_1 be the only element of $J(L) \setminus A(L)$, let $a_2, \dots, a_k$ be the atoms in $\downarrow a_1$, and let $b_1, \dots, b_{m-k}$ be the rest of atoms. (Note that $k \geq 2$.) Otherwise, when $J(L) = A(L)$, let $k = 1$, let a_1 be an arbitrarily fixed atom, and let $b_1, \dots, b_{m-1}$ be the rest of atoms.

We claim that $|\uparrow a_1| \le |L|/2$. It suffices to show that for each $x \in \uparrow a_1$ there exists an $y = y(x) \in L \setminus \uparrow a_1$ such that $a \vee y = x$. (If there are several elements y with this property then we choose one of them.) Indeed, then the existence of the *injective* mapping $\uparrow a_1 \to L \setminus \uparrow a_1$, $x \mapsto y(x)$ will complete the proof. So, let $x \in \uparrow a_1$ be an arbitrary element. Then, clearly, there is an irredundant subset U of $J(L)$ whose join is x.

First let us assume that a_i is in U for some $1 \le i \le k$. Now we define $y = \bigvee(U \setminus \{a_i\})$. Then $x = a_i \vee y$ and $a_i \le a_1 \le x$ gives $x = a_1 \vee y$ while the irredundance of U yields $a_i \not\le y$, implying $y \notin \uparrow a_1$.

Secondly, we assume that no a_i belongs to U. Then U is a set of atoms, say $U = \{b_1, \dots, b_n\}$. Using condition (1) and the irredundance of U we conclude that U is an independent set. Define $d_i = b_1 \vee \cdots \vee b_{i-1} \vee b_{i+1} \vee \cdots b_n$. Then the d_i, $1 \le i \le n$, are the coatoms of the boolean sublattice generated by U. If $a_1 \le d_i$ for all i, then $a_1 \le \bigwedge_{i=1}^{n} d_i = 0$, a contradiction. Hence we can select an $i \in \{1, \dots, n\}$ such that $a_1 \not\le d_i$. Then $y = d_i$ does the job, for $d_i = 0 \vee d_i \prec b_i \vee d_i = x$ by semimodularity, and $d_i < a_1 \vee d_i \le x$. □

Let us recall that finite, atomistic, semimodular lattices are *geometric lattices* by definition. Using the ideas around Figure 1, it is easy to see that $(x_1, y_1) \in \Theta$ implies that at leat 2^{m-2} elements of B_m are collapsed, i.e., L has at most $2^m - 2^{m-2} = 6 \cdot 2^{m-3}$ elements. This means that $|L| > 6 \cdot 2^{m-3}$ implies $J(L) = A(L)$ and $|[x_i]| = 1$ $(i = 1, \dots, m)$, whence the above proof clearly yields the following

Corollary 1 *Let L be a finite semimodular lattice with $|L| > 6 \cdot 2^{m-3}$, where $m = |J(L)|$. Then L is a geometric lattice, and for each atom f of L, $|\uparrow f| \le |L|/2$.*

If L has a Hasse diagram whose edges cross only at vertices then L is called a *planar lattice*. Recently, Grätzer and Knapp [12] has given a useful structure theorem for finite planar semimodular lattices; this is what the present paper relies on. Although this structure theorem is now generalized to all finite semimodular lattices in [7], we have been able to treat the planar case only.

If $a \parallel b$, then $S = \{a, b, a \wedge b, a \vee b\} \subseteq L$ will be called a *square* of L. If, in addition, $a \wedge b \prec a$ and $a \wedge b \prec b$, then S is called a *covering square*. By semimodularity, $a \vee b$ covers both a and b when S is a covering square. If each covering square of L is an interval then L is said to be *slim*. A mapping is called cover-preserving if it preserves the relation $\preceq$. Let us recall

Lemma 1 *(Grätzer and Knapp [12])*

- *Each finite planar <u>slim</u> semimodular lattice is a cover-preserving join-homomorphic image of the direct product of two finite chains.*
- *Each finite planar semimodular lattice can be obtained from a slim planar semimodular lattice by inserting new, doubly irreducible elements into some of its covering squares.*

Using the connection between Frankl's original conjecture and its lattice theoretic version, Roberts [16] yields that lattices with at most forty elements satisfy Frankl's conjecture. However, to explain why $|L| \geq 4$ is assumed in our main result below, we need only the obvious observation that lattices with less than four elements satisfy Frankl's conjecture.

Theorem 2 *Let L be a finite planar semimodular lattice consisting of at least four elements. Then L satisfies Frankl's conjecture. Moreover, at least one of the following two properties hold:*

- *either $1 \in J(L)$, and therefore $|{\uparrow} f| \leq |L|/4$ for $f = 1$,*
- *or there exist two distinct elements f_1 and f_2 in $J(L)$ such that $|{\uparrow} f_i| \leq |L|/2$ for $i = 1, 2$.*

Proof Let L be a finite planar semimodular lattice with $|L| \geq 4$. We will assume that L is not a chain and $1 \notin J(L)$, for otherwise the statement is evident.

First we consider the case when L is slim. We will treat it as a join-semilattice $(L, \vee)$. In virtue of Lemma 1, there are two chains, $\{0 < 1 < \cdots < n\}$ and $\{0 < 1 < \cdots < m\}$, and a join-congruence Θ of

$$D = \{0 < 1 < \cdots < n\} \times \{0 < 1 < \cdots < m\}$$

such that, up to isomorphism, $L = (L, \vee)$ equals D/Θ. (We will not use the cover-preserving property of the canonical $L \to L/\Theta$ homomorphism.) Since L is not a chain, $n \geq 2$ and $m \geq 2$. We assume that n and m are chosen such that $m + n$ is minimal, and we prove the slim case via induction on $m + n$. The smallest case, $m = n = 2$ is evident. So we assume that $m + n > 4$. For brevity, let $u = (n, 0)$, $v = (0, m)$, $1 = (m, n)$, $h = (n - 1, m)$, cf. Figure 2.

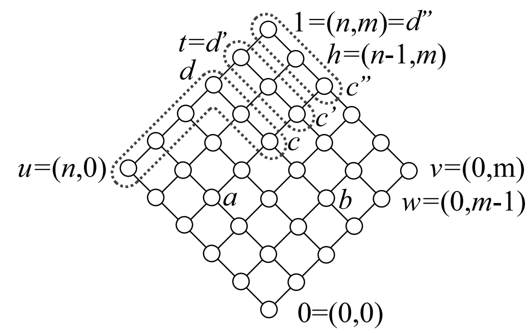


Fig. 2

Now we claim that $[u]\Theta$ and $[v]\Theta$ belong to $J(L) = J(D/\Theta)$. Their role is symmetric, so it suffices to deal with $[u]\Theta$. Suppose, by way of contradiction, that $[u]\Theta$ is not join-irreducible. Then there are $a, b \in D$ such that $[u]\Theta = [a]\Theta \vee [b]\Theta = [a \vee b]\Theta$ but $[a]\Theta < [u]\Theta$ and $[b]\Theta < [u]\Theta$, cf. Figure 2. (Although Figure 2 does not reflect the full generality, Θ is at least as large as indicated by dotted lines.) Let $c = a \vee b$. Since $[a]\Theta < [u]\Theta$ and $[b]\Theta < [u]\Theta$, we conclude that $a, b, c \in {\downarrow}h$. Let $c \prec c' \prec c'' \prec \cdots$ denote the unique maximal chain in the interval $[c, c \vee v] \subseteq {\downarrow}h$, and let $d = u \vee c$, $d' = u \vee c'$, $d'' = u \vee c'', \dots$ be the corresponding chain in the interval $[d, 1]$. Now, computing modulo Θ, for $x \in [u, d]$ we have $x = u \vee x \equiv c \vee x = d = u \vee c \equiv c \vee c = c$. Further, $d' = d \vee c' \equiv c \vee c' = c'$, $d'' = d \vee c'' \equiv c \vee c'' = c''$, etc. This means that each element of $[u, 1]$ is congruent to some element in ${\downarrow}h$ modulo Θ. Therefore, by the Third Isomorphism Theorem (cf. e.g., Thm. 6.18 in Burris and Sankappanavar [5]), $(L, \vee)$ is isomorphic to $({\downarrow}h)/\Psi$ where Ψ is the restriction of Θ to ${\downarrow}h$. However, this contradicts the minimality of $m + n$. We have seen that $[u]\Theta$ and $[v]\Theta$ are join-irreducible. $[u]\Theta = [0]\Theta$ is impossible, for otherwise L would clearly be a chain Finally, $[u]\Theta$ and $[v]\Theta$ are distinct, for otherwise $[u]\Theta = [u]\Theta \vee [v]\Theta = [u \vee v]\Theta = [1]\Theta$, a contradiction.

Now, we claim that

$$\big((0, i-1), (0, i)\big) \notin \Theta \text{ for } i = 1, \dots m. \tag{2}$$

By way of contradiction, suppose the opposite for some fixed i. Let Φ be the semilattice congruence of D whose two-element blocks are the $\{(j, i-1), (j, i)\}$, $j = 0, 1, \dots, n$, and all other blocks are singletons. Since

$$\big((j, i-1), (j, i)\big) = \big((j, 0) \vee (0, i-1), (j, 0) \vee (0, i)\big) \in \Theta,$$

we have $\Phi \subseteq \Theta$. Hence the Second Isomorphism Theorem (cf. e.g., Thm. 6.15 in [5]) gives that $(L, \vee)$ is a homomorphic image of $\{0 < 1 < \cdots < n\} \times \{0 < 1 < \cdots < m-1\}$, which contradicts the minimality of $m + n$.

Now, it follows from (2) that

$$|{\downarrow}[v]\Theta| \geq m + 1. \tag{3}$$

If, for $a \in D$, $[u]\Theta \leq [a]\Theta$ then $[a]\Theta = [u \vee a]\Theta$. This implies that

$$|{\uparrow}[u]\Theta| \leq m + 1. \tag{4}$$

We claim that

$${\uparrow}[u]\Theta \text{ is disjoint from } {\downarrow}[v]\Theta. \tag{5}$$

This comes easily, for in the opposite case we would have

$$[1]\Theta = [u \vee v]\Theta = [u]\Theta \vee [v]\Theta = [v]\Theta \in J(D/\Theta) = J(L),$$

which has been excluded previously. Now (3), (4) and (5) settle the slim case.

Finally, according to Lemma 1, the general case is obtained from the slim case via inserting new doubly irreducible elements into the interior (understood in geometrical sense in the Hasse diagram) of covering squares. Since ${\uparrow}[u]\Theta$ and ${\uparrow}[v]\Theta$ are chains, they include no covering square. Hence no new element is inserted into them. I.e., the size of ${\uparrow}[u]\Theta$ and that of ${\uparrow}[v]\Theta$ remain fixed while the size of L increases. □

References

[1] Tetsuya, A: *Excess of a lattice.* Graphs and Combinatorics **18** (2002), 395–402.

[2] Tetsuya, A: *Strong semimodular lattices and Frankl's conjecture.* Algebra Universalis **44** (2000), 379–382.

[3] Tetsuya, A., Bumpei, N.: *Frankl's conjecture is true for modular lattices, Graphs and Combinatorics.* **14** (1998), 305–311.

[4] Tetsuya, A., Bumpei, N.: *Lower semimodular types of lattices: Frankl's conjecture holds for lower quasi-semimodular lattices.* Graphs Combin. **16**, 1 (2000), 1–16.

[5] Burris, S., Sankappanavar, H. P.: *A Course in Universal Algebra.* Graduate Texts in Mathematics, 78. Springer-Verlag, New York–Berlin, 1981; The Millennium Edition, http://www.math.uwaterloo.ca/∼snburris/htdocs/ualg.html.

[6] Czédli, G.: *On averaging Frankl's conjecture for large union-closed sets*, Journal of Combinatorial Theory - Series A, to appear.

[7] Czédli, G., Schmidt, E. T.: *How to derive finite semimodular lattices from distributive lattices?* Acta Mathematica Hungarica, to appear.

[8] Czédli, G., Maróti, M, Schmidt, E. T.: *On the scope of averaging for Frankl's conjecture.* Order, submitted.

[9] Frankl, P.: *Extremal set systems. Handbook of combinatorics.* Vol. 1, 2, 1293–1329, Elsevier, Amsterdam, 1995.

[10] Weidong, G.,Hongquan, Y.: *Note on the union-closed sets conjecture.* Ars Combin. **49** (1998), 280–288.

[11] Grätzer, G.: General Lattice Theory. *Birkhäuser Verlag, Basel–Stuttgart*, 1978, sec. edi. 1998.

[12] Grätzer, G., Knapp, E.: *A note on planar semimodular lattices.* Algebra Universalis **58** (2008), 497-499.

[13] Herrmann, C., Langsdorf, R.: *Frankl's conjecture for lower semimodular lattices.* http://www.mathematik.tu-darmstadt.de:8080/∼herrmann/recherche/

[14] Poonen, B.: *Union-closed families.* J. Combinatorial Theory A **59** (1992), 253–268.

[15] Reinhold, J.: *Frankl's conjecture is true for lower semimodular lattices.* Graphs and Combinatorics **16** (2000), 115–116.

[16] Roberts, I.: Tech. Rep. No. 2/92. *School Math. Stat., Curtin Univ. Tech., Perth*, 1992.

[17] Stanley, R. P.: Enumerative Combinatorics, Vol. I. *Wadsworth & Brooks/Coole, Belmont, CA*, 1986.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 55–75

Furi–Pera Fixed Point Theorems in Banach Algebras with Applications

Smaïl DJEBALI [1], Karima HAMMACHE [2]

Department of Mathematics, E.N.S.,
Po Box 92, 16050 Kouba, Algeria
[1] *e-mail: djebali@ens-kouba.dz (corresponding author)*
[2] *e-mail: k.hammache@hotmail.com*

(Received February 17, 2008)

Abstract

In this work, we establish new Furi–Pera type fixed point theorems for the sum and the product of abstract nonlinear operators in Banach algebras; one of the operators is completely continuous and the other one is $\mathcal{D}$-Lipchitzian. The Kuratowski measure of noncompactness is used together with recent fixed point principles. Applications to solving nonlinear functional integral equations are given. Our results complement and improve recent ones in [10], [11], [17].

Key words: Banach algebra; Furi–Pera condition; fixed point theorem; measure of noncompactness; integral equations.

2000 Mathematics Subject Classification: 47H10,
secondary: 47H09, 47H30, 47J25, 46J10

1 Introduction

In many areas of natural sciences, mathematical physics, mechanics and population dynamics, problems are modeled by mathematical equations which may be reduced to perturbed nonlinear equations of the form:

$$Ay + By = y, \quad y \in M$$

where M is a closed, convex subset of a Banach space X, and A, B are two nonlinear operators. A useful tool to deal with such problems is the celebrated fixed point theorem due to Krasnozels'kĭi, 1958 (see [15, 16]):

Theorem 1.1 *Let M be a nonempty closed convex subset of a Banach space X and A, B be two maps from M to X such that*
(a) A is compact and continuous,
(b) B is a contraction,
(c) $Ax + By \in M$, for all $x, y \in M$.
Then $A + B$ has at least one fixed point in M.

We recall the

Definition 1.1 Let E be a Banach space and $f\colon E \to E$ be a mapping. Then f is said compact if $f(E)$ is compact. It is called totally bounded whenever $f(A)$ is relatively compact for any bounded subset A of E. Finally, f is completely continuous if is continuous and totally bounded.

The proof of Theorem 1.1 combines the metric Banach contraction mapping principle both with the topological Schauder's fixed point theorem [1, 7, 17, 19] and uses the fact that if E is a linear vector space, $F \subset E$ a nonempty subset and $g\colon F \to E$ a contraction, then the mapping $I - g\colon F \to (I - g)(F)$ is an homeomorphism.

In 1998, Burton [6] noticed that the Krasnozels'kĭi fixed point theorem remains valid if condition (c) is replaced by the following less restrictive one:

$$\forall y \in M,\ (x = Ay + Bx) \Longrightarrow x \in M. \tag{1}$$

However, the study of some integral equations involving the product of operators rather than the sum may be considered only in the framework of Banach algebras for which Dhage proved in 1988 the following

Theorem 1.2 [9] *Let S be a closed, convex and bounded subset of a Banach algebra X and let $A, B\colon S \to S$ be two operators such that*

(a) A is Lipschitzian with a Lipschitz constant α.

(b) $\left(\frac{I}{A}\right)^{-1}$ exists on $B(S)$, where I is an identity operator and the operator

$$\frac{I}{A}\colon X \to X \quad \textit{is defined by} \quad \left(\frac{I}{A}\right)(x) = \frac{x}{Ax}.$$

(c) B is completely continuous.

(d) $AxBy \in S$, for all $x, y \in S$.

Then the operator equation $x = AxBx$ has a solution, whenever $\alpha M < 1$, where $M := \|B(S)\| = \sup\{\|Bx\|\colon x \in S\}$.

Remark 1.1 Note that $\left(\frac{I}{A}\right)^{-1}$ exists if the operator $\frac{I}{A}$ is well defined and is one-to-one. In [10], the author improved Theorem 1.2 by removing the restrictive condition (b). The proof of the improved theorem involves the measure of noncompactness theory (see Section 2). Also, the assumption stating that A is Lipschitzian is extended to $\mathcal{D}$-Lipschitzian mappings according to the following definition. Finally, Assumption (d) is weakened to Burton's relaxed condition.

Definition 1.2 Let E be a Banach space and $f\colon E \to E$ a mapping.

(a) f is called $\mathcal{D}$-Lipschitzian with $\mathcal{D}$-function ϕ_f if there exists a continuous nondecreasing function $\phi_f\colon \mathbb{R}^+ \to \mathbb{R}^+$ such that $\phi_f(0) = 0$ and

$$\|f(x) - f(y)\| \le \phi_f(\|x - y\|),\ \forall (x, y) \in E^2.$$

(b) Moreover if $\phi_f(r) < r$, $\forall r > 0$, then f is called nonlinear contraction.

(c) In particular, if $\phi_f(r) = kr$ for some constant $0 < k < 1$, then f is a contraction.

(d) f is said non-expansive if $\phi_f(r) = r$, that is

$$\|f(x) - f(y)\| \le \|x - y\|,\ \forall (x, y) \in E^2.$$

Immediately, we have

Lemma 1.1 *Every $\mathcal{D}$-Lipschitzian mapping A is bounded, i.e. maps bounded sets into bounded sets.*

Proof Let S be a bounded subset in a Banach space E and $d = \operatorname{diam} S$ where $\operatorname{diam} S$ stands for the diameter of S. Let $s_0 \in S$ be fixed. Since ϕ_A is nondecreasing, for any $s \in S$, we have

$$\|As\| \le \|As_0\| + \|As - As_0\| \le \|As_0\| + \phi_A(\|s - s_0\|) \le \|As_0\| + \phi_A(d),$$

whence comes the result. □

Next, we state three basic existence results, important for the rest of the paper:

Theorem 1.3 ([10, Thm 2.1, p. 275]) *Let S be a closed, convex and bounded subset of a Banach algebra E and let $A, B\colon S \longrightarrow E$ be two operators such that*
(a) A is $\mathcal{D}$-Lipschitzian with $\mathcal{D}$-function ϕ_A.
(b) B is completely continuous.
(c) ($x = AxBy \Rightarrow x \in S$), for all $y \in S$.
Then the operator equation $x = AxBx$ has a solution, whenever $\phi_A(r) < r$, $\forall r > 0$ where $M := \|B(S)\|$.

The idea of extending contractions to nonlinear contractions comes from Boyd and Wong fixed point theorem which we recall hereafter for completeness; this theorem generalizes the Banach fixed point principle, dating 1922 (see e.g., [19]).

Theorem 1.4 ([5], 1969) *Let E be a Banach space and $f : E \to E$ a nonlinear contraction. Then f has a unique fixed point in E.*

In practice, condition (c) in Theorem 1.3 is not so easy to come by as it is the case in Schauder's fixed point theorem where a compact mapping is asked to map a ball into itself. In 1987, Furi and Pera introduced a new condition instead and proved the following fixed point theorem in the general framework of Fréchet spaces:

Theorem 1.5 (see [14] or [1, Thm 8.5, p. 99]) *Let E be a Fréchet space, Q a closed convex subset of E, $0 \in Q$ and let $T\colon Q \to E$ be a continuous compact mapping. Assume further that*

$$(\mathcal{FP}) \qquad \begin{cases} \text{if } \{(x_j, \lambda_j)\}_{j\geq 1} \text{ is a sequence in } \partial Q \times [0,1] \\ \text{converging to } (x, \lambda) \text{ with } x = \lambda F(x) \text{ and } 0 \leq \lambda < 1, \\ \text{then } \lambda_j F(x_j) \in Q \text{ for } j \text{ sufficiently large.} \end{cases}$$

Then T has a fixed point in Q.

Our aim in this paper is to prove new existence theorems of Dhage type with the condition (c) in Theorem 1.3 replaced by the Furi–Pera condition $(\mathcal{FP})$. More precisely, we will consider mappings of the form $F = AB + C$ where B is completely continuous and A, C are $\mathcal{D}$-Lipchitzian while F satisfies the Furi–Pera condition. This is the content of Theorems 3.1 and 3.2. In Theorem 3.3, the Furi–Pera condition is verified by another mapping, denoted N. The latter is proved to fulfill Boyd and Wong fixed point theorem (Theorem 1.4). As a consequence, we derive some known fixed point theorems obtained recently in [10, 11, 17]. The proofs are detailed in Sections 4 and 5. A further result where we relax condition (c) in Theorem 1.3 is given in Section 6. Some applications to functional nonlinear integral equations are provided in Section 7. We end the paper with some concluding remarks in Section 8. The notation := means throughout to be defined equal to. $\mathcal{B}_r(x)$ will denote the open ball in a metric space X, centered at x and of radius r and $\mathbb{R}^+$ will refer to the set of all positive real numbers. Before we present the main results of this paper, some auxiliary results are recalled hereafter.

2 Preliminaries

Definition 2.1 Let E be a Banach space and $\mathsf{B} \subset \mathcal{P}(E)$ the set of bounded subsets of E. For any subset $A \in \mathsf{B}$, define $\alpha(A) = \inf D$ where

$$D = \{\varepsilon > 0\colon A \subset \cup_{i=1}^{i=n} A_i,\ \operatorname{diam}(A_i) \leq \varepsilon,\ \forall i = 1, \ldots, n\}.$$

α is called the Kuratowski measure of noncompactness, $\alpha - MNC$ for short. Hereafter, we gather together its main properties. For more details, we refer to [3, 4, 7].

Proposition 2.1 *For any $A, B \in \mathsf{B}$, we have*
(a) $0 \leq \alpha(A) \leq \operatorname{diam}(A)$
(b) $A \subseteq B \Rightarrow \alpha(A) \leq \alpha(B)$ *(α is nondecreasing).*
(c) $\alpha(A \cup B) = \max(\alpha(A), \alpha(B))$.
(d) $\alpha(A + B) \leq \alpha(A) + \alpha(B)$ *(α is lower-additive).*
(f) $\alpha(\operatorname{Conv} A)) = \alpha(\overline{A}) = \alpha(A)$.
(h) $\alpha(A) = 0 \Rightarrow A$ *is relatively compact.*

Definition 2.2 Let E_1, E_2 be two Banach spaces and $f : E_1 \to E_2$ a continuous application which maps bounded subsets of E_1 into bounded subsets of E_2
(a) f is called α-Lipschitz if there exists some $k \geq 0$ such that

$$\alpha(f(A)) \leq k\alpha(A),$$

for any bounded subset $A \subset E_1$.
(b) f is a strict α-contraction when $k < 1$.
(c) f is said to be α-condensing whenever

$$\alpha(f(A)) < \alpha(A),$$

for any bounded subset $A \subset E_1$ with $\alpha(A) \neq 0$.

Remark 2.1 Clearly, the case $k = 0$ corresponds to f totally bounded which is of course α-condensing.

To develop further arguments, we need the following auxiliary results. They extend Theorem 1.5 to α-condensing and α-Lipschitz maps in Banach spaces, respectively (for the proofs, we refer to [17]).

Theorem 2.1 *Let E be a Banach space and Q a closed convex bounded subset of E with $0 \in Q$. In addition, assume $F\colon Q \to E$ is an α-condensing map which satisfies the Furi–Pera condition. Then F has a fixed point $x \in Q$.*

Theorem 2.2 *Let E be a Banach space and Q a closed convex bounded subset of E with $0 \in Q$. In addition, assume that $(I - F)(S)$ is a closed, $F\colon Q \to E$ is an α-Lipschitz map with $k = 1$ and satisfies the Furi–Pera condition. Then F has a fixed point $x \in Q$.*

3 Main results

We state the following main theorems of this paper.

Theorem 3.1 *Let S be a closed, convex and bounded subset of a Banach algebra X with $0 \in S$ and let $A, C\colon X \to X$ and $B\colon S \to X$ be three operators such that*

(a) A and C are $\mathcal{D}$-Lipschitzian with $\mathcal{D}$-functions ϕ_A and ϕ_C respectively.

(b) B is completely continuous.

(c) The operator $F\colon S \to X$ defined by

$$F(x) = AxBx + Cx$$

satisfies the Furi–Pera condition.
Then the abstract equation $x = AxBx + Cx$ has a solution $x \in S$ whenever

$$(\mathcal{H}_0) \qquad M\phi_A(r) + \phi_C(r) < r, \quad \forall r > 0$$

where $M = \|B(S)\|$.

Remark 3.1 More generally, one may consider n $\mathcal{D}$-Lipschitzian operators A_i $(i = 1, \cdots, n)$ with $\mathcal{D}$-functions ϕ_i and completely continuous operators B_i $(i = 1, \cdots, n)$ defined on a closed, bounded, convex subset S of a Banach algebra containing 0 and satisfying

$$\sum_{i=1}^{n} M_i \phi_i(r) < r$$

where, for each i, $M_i = \|B_i(S)\|$. If the operator

$$F(x) = \sum_{i=1}^{n} (A_i B_i)(x)$$

satisfies the Furi–Pera condition, then the abstract nonlinear equation

$$\sum_{i=1}^{n} (A_i B_i)(x) = x$$

has a solution $x \in S$.

Remark 3.2 Let $S = \mathcal{B}_R(0)$. It is well known that $F(\partial S) \subset S$ implies the Furi–Pera condition. Assume further that $A0 = C0 = 0$. Then also Assumption $(\mathcal{H}_0)$ implies the Furi–Pera condition. Indeed, let $(x_j, \lambda_j)_{j \geq 1}$ be a sequence in $\partial S \times [0, 1]$ converging to some limit (x, λ) with $x = \lambda F(x)$ and $0 \leq \lambda < 1$, then

$$\|\lambda_j F x_j\| \leq \|A x_j\| \|B x_j\| + \|C x_j\| \leq M \phi_A(\|x_j\|) + \phi_C(\|x_j\|) \leq \|x_j\| = R.$$

Hence $\lambda_j F(x_j) \in S$, $\forall j \in \mathbb{N}^*$.

Theorem 3.2 *Let S be a closed, convex and bounded subset of a Banach algebra X with $0 \in S$ and let $A, C \colon X \to X$ and $B \colon S \to X$ be three operators such that*

(a) A and C are $\mathcal{D}$-Lipschitzian with $\mathcal{D}$-functions ϕ_A and ϕ_C respectively.

(b) B is completely continuous.

(c) The operator $F \colon S \to X$ defined by $F(x) = AxBx + Cx$ satisfies the Furi–Pera condition.

Then the abstract equation $x = AxBx + Cx$ has a solution $x \in S$ provided $(I - F)(S)$ is closed and the large inequality $M\phi_A(r) + \phi_C(r) \leq r$, $\forall r > 0$ holds.

Theorem 3.3 *Let S be a closed, convex and bounded subset of a Banach algebra X with $0 \in S$ and let $A \colon X \to X$ and $B \colon S \to X$ be two operators such that*

(a) A is $\mathcal{D}$-Lipschitzian with $\mathcal{D}$-function ϕ_A.

(b) B is completely continuous.

(c) The operator $N \colon S \to X$ defined by $Nx = y$, where y is the unique solution of the operator equation $y = AyBx$, satisfies the Furi–Pera condition.

Then the operator equation $x = AxBx$ has a solution $x \in S$ provided

$$(\mathcal{H}) \qquad \begin{cases} \text{the mapping } \Phi \colon [0, +\infty) \to [0, +\infty) \\ r \mapsto \Phi(r) = r - M\phi_A(r) \text{ is increasing to infinity.} \end{cases}$$

Remark 3.3 (a) Obviously, Assumption ($\mathcal{H}$) implies that for any $r > 0$ $M\phi_A(r) < r$ and amounts to the restriction $0 < k < \frac{1}{M}$ in case ϕ_A is a k-contraction.

(b) The condition (c) in Theorem 3.3 is different from the condition (c) in Theorem 3.1 which means that $N(S) \subset S$. One can make analogy with the condition (c) in Theorem 1.1 compared with Burton's weak condition (1). Regarding these conditions, a discussion is given in the concluding remarks (see Section 7).

3.1 Some consequences

In this section, we derive four existence principles. In particular, we recover some known results: the first one is Dhage's Theorem 1.3 ([10, Thm 2.1]); the second one is a nonlinear alternative in Banach algebras ([11, Thm 2.2]), the third one is concerned with the case of the sum of a nonlinear contraction and a compact mapping ([17, Thm 2.2]) while the last one is a useful classical result.

Corollary 3.1 *Assume Assumptions (a)–(c) in Theorem 1.3 are satisfied,* $0 \in S$ *where* S *is either a ball or any subset homeomorphic to a bounded, closed convex subset and* $AB(S) \subset S$*. Then, the same conclusion of this theorem holds true provided* ($\mathcal{H}$) *is fulfilled.*

Proof We prove the corollary first in case $S = \mathcal{B}_M(0)$ and then when AB maps S into itself. In the latter case, S could be any subset homeomorphic to a closed, bounded and convex subset of X.
Step 1: $S = \mathcal{B}_M(0)$. We only have to check that condition (c) in Theorem 1.3 implies the Furi–Pera condition (c) in Theorem 3.3. For this, let $(x_j, \lambda_j)_{j\geq 1}$ be a sequence in $\partial S \times [0,1]$ converging to some limit (x, λ) with $x = \lambda N x$, $0 \leq \lambda < 1$ and show that $\lambda_j N(x_j) \in S$ for j large enough. For any $j \in \mathbb{N}^*$, $\|\lambda_j N(x_j)\| \leq \|N(x_j)\| = \|y_j\|$ where $y_j = Ay_jBx_j$. Since $x_j \in \partial S \subset S$ and condition (c) of Theorem 1.3 is satisfied, $y_j \in S$. Hence $\|y_j\| \leq M$. This implies that $\|\lambda_j N(x_j)\| \leq M$. Our claim, that is $\lambda_j N(x_j) \in S$, is then proved.
Step 2: $AB(S) \subset S$. By the Dugundji's extension theorem (see [7, 19]), let $r\colon X \to S$ be a retraction and let $\mathcal{B}$ be a ball containing S. Then consider the diagram

$$\mathcal{B} \xrightarrow{r} S \xrightarrow{AB} S.$$

From Step 1, the map $AB \circ r$ has a fixed point $x \in \mathcal{B}$, that is satisfying $Ar(x)Br(x) = x$. Since $ABr(x) \in S$, it follows that $x \in S$ and thus $r(x) = x = ABx$.
Step 3: S is homeomorphic to $\tilde{S}$, where $\tilde{S}$ is a closed, bounded and convex subset of X. Let the diagram

$$\tilde{S} \xrightarrow{h^{-1}} S \xrightarrow{AB} S \xrightarrow{h} \tilde{S},$$

where h is an homeomorphism. From Step 2, there exists some $y \in \tilde{S}$ such that $h \circ AB \circ h^{-1}(y) = y$. Then $ABx = x$, for $x = h^{-1}(y) \in S$, ending the proof of the corollary. □

Corollary 3.2 ([11, Thm 2.2, p. 272]) *Let X be a Banach algebra and let $A, B, C\colon X \longrightarrow X$ be three operators satisfying $(\mathcal{H}_0)$ together with*
(a) A and C are $\mathcal{D}$-Lipschitzian with $\mathcal{D}$-functions ϕ_A and ϕ_C respectively.
(b) B is completely continuous.
Then
(a) either $F = AB + C$ has a fixed point in X,
(b) or the set $\{x \in X,\ \lambda F(x) = x,\ 0 < \lambda < 1\}$ is unbounded.

Proof Assume Alternative (b) does not hold true. Then, there exists some positive constant R such that

$$\forall \lambda \in (0,1),\ (\lambda F(x) = x \Rightarrow \|x\| \le R). \tag{2}$$

In order to show that F satisfies the Furi–Pera condition, consider a sequence $(x_j, \lambda_j)_{j\ge 1} \in \partial S \times [0,1]$ converging to some limit (x, λ) with $x = \lambda F(x)$ and $0 < \lambda < 1$, where $S = \overline{\mathcal{B}}_{R+1}(0)$. By continuity of F, we have that

$$\|\lambda_j F(x_j)\| \le \|\lambda F(x)\| + 1, \text{ for sufficientlly large } j. \tag{3}$$

Since $x = \lambda F(x)$, (2) yields

$$\|\lambda F(x)\| = \|x\| \le R.$$

This with (3) imply that $\lambda_j F_j(x) \in S$. Our claim, namely Alternative (a), then follows from Theorem 3.1. □

The following two particular cases of Theorem 3.1 are useful in practice.

Corollary 3.3 ([17, Thm 2.2, p. 3]) *Let S be a closed, convex and bounded subset of a Banach space X with $0 \in S$ and let $F_1\colon X \to X$ and $F_2\colon S \to X$ be two operators such that*
(a) F_1 is a nonlinear contraction.
(b) F_2 is completely continuous.
(c) The sum $F = F_1 + F_2\colon S \to X$ satisfies the Furi–Pera condition $(\mathcal{FP})$.
Then F has a fixed point $x \in S$.

Proof Take $B = F_2,\ C = F_1,\ A \equiv 1$ and then $\varphi_A \equiv 0$. □

Corollary 3.4 *Let S be a closed, convex and bounded subset of a Banach algebra X with $0 \in S$ and let $A, C\colon X \to X$ and $B\colon S \to X$ be three operators such that*
(a) A and C are Lipschitzian with Lipschitz constants k_A and k_C respectively.
(b) B is completely continuous.
(c) The operator $F\colon S \to X$ defined by $F(x) = AxBx + Cx$, $x \in X$ satisfies the Furi–Pera condition $(\mathcal{FP})$.
Then the equation $x = AxBx + Cx$ has a solution $x \in S$ whenever $k_A\|B(S)\| + k_C < 1$.

4 Proofs of Theorems 3.1 and 3.2

The proofs are direct consequences of the following lemma:

Lemma 4.1 *Under Assumptions (a), (b) of Theorem 3.1 together with* $(\mathcal{H}_0)$*, the map* $F\colon S \to X$ *defined by* $F(x) = AxBx + Cx$ *is* α*-condensing.*

Proof Let $D \subset S$ be a bounded subset and $\delta > 0$. There exists a covering $(D_i)_{i=1}^n$ such that $D \subset \bigcup_{i=1}^n D_i$ and $\operatorname{diam}(D_i) \leq \alpha(D) + \delta$, for each $i = 1, \ldots, n$. For every $i \in \{1, \cdots, n\}$, let $x_1^i = x_1$, $x_2^i = x_2 \in D_i$ and $E_i = F(D_i)$. Clearly $F(D) \subset \bigcup_{i=1}^n E_i$. In addition, we have the estimates

$$\begin{aligned}
&\|F(x_1) - F(x_2)\| = \|Ax_1Bx_1 + Cx_1 - Ax_2Bx_2 - Cx_2\| \\
&\leq \|Ax_1\|\,\|Bx_1 - Bx_2\| + \|Bx_2\|\,\|Ax_1 - Ax_2\| + \|Cx_1 - Cx_2\| \\
&\leq \|Ax_1\| \operatorname{diam}\big(B(D_i)\big) + M\,\|Ax_1 - Ax_2\| + \|Cx_1 - Cx_2\| \\
&\leq \|Ax_1\|\,\alpha\big(B(D_i)\big) + M\,\phi_A(\|x_1 - x_2\|) + \phi_C(\|x_1 - x_2\|).
\end{aligned}$$

Since B is completely continuous, $\alpha\big(B(D_i)\big) = 0$, for each $i \in \{1, \cdots, n\}$ follows from Proposition 2.1(e). We infer that

$$\|F(x_1) - F(x_2)\| \leq M\,\phi_A(\|x_1 - x_2\|) + \phi_C(\|x_1 - x_2\|).$$

Since ϕ_A and ϕ_C are non decreasing, it follows that

$$\begin{aligned}
\operatorname{diam} E_i &\leq M\,\phi_A(\|x_1 - x_2\|) + \phi_C(\|x_1 - x_2\|) \\
&\leq M\phi_A(\operatorname{diam}(D_i)) + \phi_C(\operatorname{diam}(D_i)) \\
&\leq M\phi_A(\alpha(D) + \delta) + \phi_C(\alpha(D) + \delta).
\end{aligned}$$

Therefore

$$\alpha(F(D)) \leq M\phi_A(\alpha(D) + \delta) + \phi_C(\alpha(D) + \delta).$$

Since $\delta > 0$ is arbitrary, we deduce the estimate

$$\alpha(F(D)) \leq M\phi_A(\alpha(D)) + \phi_C(\alpha(D)).$$

Taking into account Assumption $(\mathcal{H}_0)$, we arrive at

$$\alpha\big(F(D)\big) < \alpha(D),$$

proving our claim. □

Remark 4.1 *It is well known (see e.g. [17, proof of Thm 2.1]) that the sum of a nonlinear contraction and a compact mapping is* α*-condensing. Lemma 4.1 extends this result to* $\mathcal{D}$*-Lipschitz mappings as well as to the product of a compact and a* $\mathcal{D}$*-Lipschitz maps.*

Proof of Theorem 3.1. By Lemma 4.1, the map $F\colon S \to X$ defined by $F(x) = AxBx + Cx$ is α-condensing. Since F satisfies the Furi–Pera condition, it follows by Theorem 2.1 that F has at least one fixed point $x \in S$, solution of the equation $x = AxBx + Cx$, ending the proof of the theorem. □

Proof of Theorem 3.2. Since the mapping $F\colon S \to X$ defined by $F(x) = AxBx + Cx$ is α-condensing by Lemma 4.1, then it is α-Lipschitz with $k = 1$. Moreover $(I-F)(S)$ is closed and F satisfies the Furi–Pera condition. Therefore, Theorem 2.2 implies that F has at least one fixed point $x \in S$, solution of the equation $x = AxBx + Cx$. □

5 Proof of Theorem 3.3

The proof follows from Theorem 2.1 once we have proved the following two technical lemmas.

Lemma 5.1 *Under the assumptions of Theorem 3.3, the operator $N\colon X \to X$ introduced in condition (c) is well defined and is bounded (on bounded subsets of X).*

Proof For any $x \in S$, let the mapping A_x be defined in X by $A_x y = AyBx$. Then, for any $y_1, y_2 \in X$,

$$\|A_x y_1 - A_x y_2\| = \|Bx\|\|Ay_1 - Ay_2\| \leq \|Bx\|\phi_A\left(\|y_1 - y_2\|\right) \leq M\phi_A\left(\|y_1 - y_2\|\right)$$

with $M\phi_A(r) < r$, $\forall r > 0$. By the Boyd and Wong fixed point theorem (see Theorem 1.4), A_x has only one fixed point $y \in X$ and so the mapping N is well defined. In addition, let $D \subset X$ be any bounded subset, $x \in D$ and $y = Nx$ where y is the unique solution of the equation $y = AyBx$. Thus

$$\|y\| = \|Bx\|\|Ay\| \leq M\|Ay\|.$$

Let $y_0 \in X$. With Assumption $(\mathcal{H})$, we have the following estimates

$$\|y\| \leq M\left(\|Ay - Ay_0\| + \|Ay_0\|\right) \leq M\phi_A\left(\|y - y_0\|\right) + M\|Ay_0\|.$$

Hence

$$\|y - y_0\| \leq \|y\| + \|y_0\| \leq M\phi_A\left(\|y - y_0\|\right) + M\|Ay_0\| + \|y_0\|.$$

It follows that

$$\Phi(\|y - y_0\|) = \|y - y_0\| - M\phi_A(\|y - y_0\|) \leq M\|Ay_0\| + \|y_0\|.$$

This in turn implies successively

$$\begin{aligned}
&\|y - y_0\| \leq \Phi^{-1}(M\|Ay_0\| + \|y_0\|)\\
&\|y\| \leq \|y - y_0\| + \|y_0\| \leq \Phi^{-1}(M\|Ay_0\| + \|y_0\|) + \|y_0\|,
\end{aligned}$$

proving our claim. □

Remark 5.1 To prove Lemma 5.1, we only need $\lim_{s\to+\infty}\Phi(s)=+\infty$ without the increasing character of Φ.

Lemma 5.2 Under the hypotheses of Theorem 3.3, the operator N introduced in the condition (c) is compact.

Proof

Claim 1. N is continuous. Let (x_n) be a sequence in S converging to some limit x. Since S is closed, $x\in S$. Moreover

$$\begin{aligned}\|Nx_n-Nx\| &= \|ANx_nBx-ANxBx\|\\ &\le \|ANx_nBx_n-ANxBx_n\|+\|ANxBx_n-ANxBx\|\\ &\le \|Bx_n\|\|ANx_n-ANx\|+\|ANx\|\|Bx_n-Bx\|\\ &\le M\phi(\|x_n-x\|)+\|ANx\|\|Bx_n-Bx\|.\end{aligned}$$

Whence

$$\limsup_{n\to\infty}\|Nx_n-Nx\|\le M\phi(\limsup_{n\to\infty}\|x_n-x\|)+\|ANx\|\limsup_{n\to\infty}\|Bx_n-Bx\|.$$

From Assumption (b), B is continuous; hence

$$\limsup_{n\to\infty}\|Nx_n-Nx\|=0,$$

yielding the continuity of N.

Claim 2. N is compact. From Lemmas 1.1 and 5.1, there exists some positive constant k_1 such that $\|ANx\|\le k_1$, $\forall x\in S$. Let $\varepsilon>0$ be given. Since S is bounded and B is completely continuous, $B(S)$ is relatively compact. Then there exists a set $\mathcal{E}=\{x_1,\dots,x_n\}\subset S$ such that

$$B(S)\subset\bigcup_{i=1}^{n}\mathcal{B}_\delta(w_i)$$

where $w_i:=B(x_i)$ and $\delta:=k_2\varepsilon$ for some constant k_2 to be selected later on. Therefore, for any $x\in S$, there exists some $x_i\in\mathcal{E}$ such that

$$0\le\|Bx-Bx_i\|\le k_2\varepsilon.$$

We have

$$\begin{aligned}\|Nx_i-Nx\| &= \|ANx_iBx_i-ANxBx\|\\ &\le \|ANx_iBx_i-ANxBx_i\|+\|ANxBx_i-ANxBx\|\\ &\le \|Bx_i\|\|ANx_i-ANx\|+\|ANx\|\|Bx_i-Bx\|\\ &\le M\phi_A(\|Nx_i-Nx\|)+k_1k_2\varepsilon. \qquad (4)\end{aligned}$$

Hence

$$\Phi(\|Nx_i-Nx\|)=\|Nx_i-Nx\|-M\phi_A(\|Nx_i-Nx\|)\le k_1k_2\varepsilon.$$

From Assumption $(\mathcal{H})$, it follows that

$$\|Nx_i - Nx\| \leq \Phi^{-1}(k_1 k_2 \varepsilon).$$

Choosing $0 < k_2 \leq \frac{\Phi(\varepsilon)}{k_1 \varepsilon}$, we obtain

$$\|Nx_i - Nx\| \leq \varepsilon.$$

We have proved that $N(S) \subset \bigcup_{i=1}^{n} \mathcal{B}_\varepsilon(Nx_i)$, showing that N is totally bounded and ending the proof of Lemma 5.2. □

Remark 5.2 In Claim 2, we correct the proof of Theorem 2.1, p. 275 of [10] where ϕ_A was taken $\phi_A(r) \leq \alpha r$, $r > 0$. To this end, Assumption $(\mathcal{H})$ is essential.

Remark 5.3 In Theorem 3.3, the condition that S is unbounded may be relaxed when B completely continuous is replaced by B compact, that is $B(S)$ relatively compact. Indeed, the proof of Lemma 5.2 remains unchanged and then we rather apply Theorem 1.5.

6 A further result

In the following, we prove that the condition (c) in Theorem 1.3 may be relaxed.

Theorem 6.1 *Let S be a closed, convex and bounded subset of a Banach algebra E such that $\operatorname{int} S \neq \emptyset$ and let $A, B\colon S \to E$ be two operators such that*
(a) A is $\mathcal{D}$-Lipschitzian with $\mathcal{D}$-function ϕ_A.
(b) B is completely continuous.
(c') $(x = AxBy \Rightarrow x \in S)$, for all $y \in \partial S$.
Then the operator equation $x = AxBx$ has a solution, whenever $M\phi_A(r) < r$, $\forall r > 0$.

Proof Let $r\colon X \to S$ be a retraction. Moreover using the Minkowski functional (see [18, Lemma 4.2.5, p. 27]), r may be chosen so that

$$\begin{cases} r(x) = x, & x \in S \\ r(x) \in \partial S, & x \notin S. \end{cases}$$

We claim that $Nr\colon X \to X$ has a fixed point (here we denote by $fg = f \circ g$). Since N is completely continuous by Lemma 5.2 and r is continuous, the composite $rN\colon S \to S$ is completely continuous. Then Schauder's fixed point theorem implies that rN has a fixed point, i.e. there exists some $x_0 \in S$ such that $rNx_0 = x_0$. As a consequence, Nr has a fixed point. Indeed, letting $y_0 = Nx_0$, we get

$$ry_0 = rNx_0 = x_0 \Longrightarrow Nry_0 = Nx_0 = y_0.$$

That is $y_0 = Ay_0Bry_0$. Since $ry_0 \in \partial S$, assumption (c') implies that $y_0 \in S$, ending the proof of the theorem. □

Remark 6.1 One may take the subset S unbounded and the operator B compact and then apply Rothe's Theorem to prove Theorem 6.1. This is the main motivation of the study of the topological structure of the subsubset $\mathcal{F}_{Nr} := \{x \in X,\, x = Nrx\} \subset S$ which is nonempty by Theorem 6.1.

(a) $\mathcal{F}_{Nr}$ *is closed.* Let $(x_n)_{n\in\mathbb{N}} \subset \mathcal{F}_{Nr}$ be a sequence such that $x_n \to x$, as $n \to +\infty$. We show that $x \in \mathcal{F}_{Nr}$. First $x_n = Nr(x_n)$ and, since N and r are continuous, $\lim_{n\to+\infty} Nr(x_n) = Nr(x)$. Then, the uniqueness of the limit implies that $x = Nr(x)$ yielding $x \in \mathcal{F}_{Nr}$.

(b) $\mathcal{F}_{Nr}$ *is compact.* Indeed, $r(\mathcal{F}_{Nr}) \subset S$ implies $N(r(\mathcal{F}_{Nr})) \subset N(S)$ and then $\alpha(N(r(\mathcal{F}_{Nr}))) \leq \alpha(N(S)) = 0$ because N is compact by Lemma 5.2. Here α is the measure of noncompactness (see Section 2). In addition, $\mathcal{F}_{Nr} \subset Nr(\mathcal{F}_{Nr})$ implies that $\alpha(\mathcal{F}_{Nr}) \leq \alpha(Nr(\mathcal{F}_{Nr})) \leq 0$; then $\alpha(\mathcal{F}_{Nr}) = 0$ and our claim follows.

7 Applications

7.1 Example 1

Let $X = C([0,1],\mathbb{R})$ be the Banach Algebra of real continuous functions defined on the interval $[0,1]$ endowed with the sup-norm

$$\|x\| = \max_{t\in[0,1]} |x(t)|.$$

For some sufficiently large positive real number R, let $S = \overline{\mathcal{B}}_R(0)$ be the closed ball centered at the origin and with radius R. Consider the nonlinear functional integral equation, for $t \in [0,1]$ and the parameter α lies in the interval $(0,1)$

$$x(t) = \left(1 + \frac{\alpha R}{R+1}|x(\mu(t))|\right)\left(q(\theta(t)) + \int_0^{\sigma(t)} g\big(s, x(\eta(s))\big)\, ds\right) \qquad (5)$$

where the functions $\mu, \theta, \sigma, \eta\colon [0,1] \to [0,1]$ are continuous. Assume that $q\colon [0,1] \to \mathbb{R}$ and $g\colon [0,1]\times\mathbb{R} \to \mathbb{R}$ are continuous functions and satisfy

$$|g\big(t,x)| \leq 1 - \|q\|_\infty, \quad \forall (t,x) \in [0,1]\times\mathbb{R}, \qquad (6)$$

where $\|q\|_\infty := \max_{t\in[0,1]} |q(t)|$. Let the mappings A and B be defined by

$$A\colon X \to X, \qquad Ax(t) = 1 + \frac{\alpha R}{R+1}|x(\mu(t))|$$

and

$$B\colon S \to X, \qquad Bx(t) = q(\theta(t)) + \int_0^{\sigma(t)} g\big(s, x(\eta(s))\big)\, ds.$$

Then the integral equation (5) is equivalent to the operator equation

$$Ax(t)Bx(t) = x(t), \quad t \in [0,1].$$

(a) *Properties of the mappings A, B.* Clearly, A is a Lipchitzian map with constant $k = \frac{R}{R+1}\cdot$ To prove B is completely continuous, let $(x_n)_{n\in\mathbb{N}} \subset X$. Since

$$|B(x_n)(t)| \leq |q(\theta(t))| + \left|\int_0^{\sigma(t)} g\big(s, x(\eta(s))\big)\, ds\right| = \|q\| + (1 - \|q\|) = 1,$$

the sequence $(B(x_n))_{n\in\mathbb{N}}$ is uniformly bounded. Moreover B is equi-continuous. To see this, let $t_1, t_2 \in [0,1]$; then

$$\begin{aligned}|B(x_n)(t_1) - B(x_n)(t_2)| &\leq |q(\theta(t_1)) - q(\theta(t_2))| + \left|\int_{\sigma(t_1)}^{\sigma(t_2)} g\big(s, x(\eta(s))\big)\, ds\right| \\ &\leq |q(\theta(t_1)) - q(\theta(t_2))| + (1 - \|q\|)|\sigma(t_1) - \sigma(t_2)|.\end{aligned}$$

The continuity of θ, σ, q on the compact interval $[0,1]$ implies that $(B(x_n))_{n\in\mathbb{N}}$ is equi-continuous and then B is completely continuous by Arzela–Ascoli Lemma.

(b) *$F = AB$ satisfies the Furi–Pera condition.* For this purpose, consider a sequence $(x_j, \lambda_j)_{j\geq 1} \in \partial S \times [0,1]$ converging to some limit (x, λ) with $x = \lambda F(x)$ and $0 \leq \lambda < 1$. Then for j sufficiently large, we have

$$\|\lambda_j F(x_j)\| \leq \lambda \|F(x)\| + 1.$$

Since $x = \lambda F(x)$, we deduce the bounds:

$$\|x\| \leq \lambda \left(1 + \frac{\alpha R}{R+1}\|x\|\right),$$

and then

$$\|x\| \leq \frac{\lambda(R+1)}{R(1-\alpha\lambda)+1}\cdot$$

Hence

$$\|\lambda_j F(x_j)\| \leq \frac{R+1}{R(1-\alpha)+1}, \quad 0 \leq \lambda_j < 1.$$

This implies that, for R large enough, namely $R \geq \frac{1}{\sqrt{1-\alpha}}$, it holds that

$$\|\lambda_j F(x_j)\| \leq R$$

and so $\lambda_j F_j(x) \in S$. Finally

$$0 < k = \frac{R}{R+1} < \frac{1}{\|B(S)\|} = 1.$$

Then all assumptions of Theorem 3.1 are met with $C = 0$ and Equation (5) has a solution in X provided (6) holds true. Notice that for this first example, Corollary 3.2 may be applied as well; this will not be the case with the next two examples.

7.2 Example 2

(a) Consider the Banach space

$$X = C_0(\mathbb{R}, \mathbb{R}) = \{x \in C(\mathbb{R}, \mathbb{R}), \lim_{|t| \to +\infty} x(t) = 0\}$$

endowed with the sup-norm

$$\|x\|_X = \sup_{t \in \mathbb{R}} \{|x(t)|\}.$$

Let a continuous function $a\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be such that

$$a \text{ is } k\text{-Lipschitz with respect to the second argument} \tag{7}$$

and then define the mapping A by

$$Ax(t) = \int_{-|t|}^{|t|} e^{-\alpha|s|} a(s, x(s))\, ds, \quad t \in \mathbb{R},$$

for some positive parameter α. Then A is $\frac{2k}{\alpha}$-Lipschitz. Indeed

$$|Ax_1(t) - Ax_2(t)| \leq \left| \int_{-|t|}^{|t|} e^{-\alpha|s|} \Big(a(s, x_1(s)) - a(s, x_2(s)) \Big) ds \right|$$

$$\leq k \int_{-|t|}^{|t|} e^{-\alpha|s|} \, |x_1(s) - x_2(s)| \; ds \leq k_1 \|x_1 - x_2\| \int_{-|t|}^{|t|} e^{-\alpha|s|}\, ds.$$

Hence

$$\|Ax_1 - Ax_2\| \leq \frac{2k}{\alpha} \|x_1 - x_2\|.$$

Thus, we assume $\frac{2k}{\alpha} < 1$. In addition

$$\|Ax\| \leq \frac{2k}{\alpha} \|x\| + \frac{2}{\alpha} \|a(.,0)\|. \tag{8}$$

(b) Let the mapping B be defined by

$$Bx(t) = \int_{-\infty}^{+\infty} G(t, s) h\big(s, x(s)\big)\, ds,$$

where the nonlinear function $h\colon \mathbb{R}^2 \to \mathbb{R}$ is continuous and verifies the growth condition:

$$|h(t, x)| \leq q(t) \Psi(|x|), \quad \forall\, t, x \in \mathbb{R} \tag{9}$$

where $q \in C_0(\mathbb{R}, \mathbb{R}^+)$ and $\Psi\colon \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous nondecreasing function. The kernel $G\colon \mathbb{R}^2 \to \mathbb{R}$ is continuous and satisfies

$$\exists \sigma, \varrho > 0, \ |G(t, s)| \leq \varrho e^{-\sigma|t-s|}, \quad \forall s, t \in \mathbb{R}. \tag{10}$$

We can show that B is completely continuous (see the proof of Theorem 2.1 in [8] for the details). Let the bounded closed and convex subset of X:

$$S = \{x \in X \colon \|x\| \leq R\},$$

where the positive constant R is to be selected later on.

(c) Assume that for any compact subset $K \subset \mathbb{R}$, there exists a positive constant $M_K > 0$ such that for any $x \in X$, $\lambda \in [0,1)$

$$x = \lambda AxBx \Longrightarrow (|x(t)| \leq M_K, \ \forall t \in K). \tag{11}$$

(d) To verify the Furi–Pera condition, let $(x_j, \lambda_j) \in \partial S \times [0,1]$ be such that, as $j \to +\infty$, $\lambda_j \to \lambda$ and $x_j \to x$ with $\lambda F(x)$ and $0 \leq \lambda < 1$. We show that $\lambda_j F(x_j) \in S$ where $F(x_j) = Ax_jBx_j$. Since Ψ is nondecreasing, we have

$$|Bx(t)| \leq \Psi(R) \int_{-\infty}^{+\infty} G(t,s)q(s)ds := \gamma(t).$$

Moreover, for each j, we have

$$\begin{aligned} &\|\lambda_i F(x_j)\| \leq \|Ax_j\| \cdot \|Bx_j\| \\ \leq &\left(\frac{2k}{\alpha}\|x_1\| + \frac{2}{\alpha}\|a(\cdot,0)\|\right)\gamma(t) \leq \left(\frac{2k}{\alpha}R + \frac{2}{\alpha}\|a(\cdot,0)\|\right)\gamma(t) \end{aligned}$$

Since $\lim_{t\to\pm\infty}\gamma(t) = 0$, there exist some t_1, t_2 and a sufficiently small positive constant M_1 such that

$$\|\lambda_i F(x_j)(t)\| \leq M_1, \quad \forall t \in (-\infty, t_1) \cup (t_2, +\infty). \tag{12}$$

In addition, for $t \in [t_1, t_2]$ and $x_j \in \partial S$, $\lim_{j\to+\infty} x_j = x = \lambda F(x)$ uniformly. Then for j large enough and $t \in [t_1, t_2]$, we have, from conditions (11) and (12), that

$$\|\lambda_i F(x_j)(t) \leq \lambda |F(x)(t)| + 1 \leq M_0 + 1. \tag{13}$$

Combining (12) and (13) and taking $R = \max(M_1, M_0 + 1)$, we get

$$\|\lambda_i F(x_j)(t)\| \leq R, \quad \forall t \in \mathbb{R}, \ \forall j \in \mathbb{N}.$$

Therefore the Furi–Pera condition is fulfilled.

As a consequence, we have proved that, under Assumptions (7), (9), (10) and (11), the nonlinear problem

$$\left(\int_{-|t|}^{|t|} e^{-\alpha|s|} a(s, x(s))\, ds\right)\left(\int_{-\infty}^{+\infty} G(t,s) h\big(s, x(s)\big)\, ds\right) = x(t), \quad t \in \mathbb{R}$$

admits, by Theorem 3.1, at least one solution $x \in \mathcal{B}_R(0) \subset C(\mathbb{R}, \mathbb{R})$.

Remark 7.1 Notice that in the particular case $a(.,0) \equiv 0$, the last condition in Theorem 3.1, namely $\|B(S)\|\frac{2k}{\alpha} < 1$, is equivalent to the Furi–Pera condition (see Remark 3.2). In such a case, we have only to find a function Ψ such that there exists some $R > 0$ such that

$$\frac{4k\varrho}{\alpha\sigma}\Psi(R)\|q\|_\infty < 1,$$

which is obviously satisfied whenever $\lim_{s\to+\infty}\Psi(s) = +\infty$.

7.3 Example 3

We will make use of the nonlinear version of Gronwall's Lemma (see [2])

Lemma 7.1 *Let $I = [a,b]$ and $u, g\colon I \to \mathbb{R}$ be positive real continuous functions. Assume there exist $c > 0$ and a continuous nondecreasing function $h\colon \mathbb{R} \to (0,+\infty)$ such that*

$$u(t) \le c + \int_a^t g(s)h(u(s))\,ds, \quad \forall t \in I.$$

Then, we have

$$u(t) \le \Psi^{-1}\left(\int_a^t g(s)\,ds\right), \quad \forall t \in I$$

with

$$\Psi(u) = \int_c^u \frac{dy}{h(y)}$$

for $u \ge c$ and Ψ^{-1} referring to the inverse of the function Ψ, provided for any $t \ge a$, $\int_a^t g(s)\,ds \in \operatorname{Dom}\Psi^{-1}$.

Let $a > 1$ and $X = C_0([a,+\infty),\mathbb{R})$ be the set of real continuous functions x defined on the interval $[a,+\infty)$ and such that $\lim_{t\to+\infty} x(t) = 0$. Equipped with the sup-norm $\|x\| = \sup_{t\ge a}|x(t)|$, it is a Banach space.

(a) On the space X, define a mapping A by

$$Ax(t) = h(t) + \int_a^t f(s,x(s))\,ds, \quad t \ge a$$

where $f\colon [a,+\infty)\times\mathbb{R} \to \mathbb{R}$ is continuous, $f(t,0) = 0$, $t \ge a$ and there exists $p \in L^1([a,+\infty),\mathbb{R}^+)$ such that the nonlinear p-Lipschitz condition is satisfied:

$$|f(t,x(t)) - f(t,y(t))| \le p(t)|x(t)-y(t)|^\delta, \quad \forall (x,y) \in X^2, \tag{14}$$

with some $0 < \delta < 1$. The function $h\colon [a,+\infty) \to \mathbb{R}$ is continuous and nonidentically zero. If we let $|p|_1 = \int_a^{+\infty} p(t)dt$, then we can see that the operator A is $|p|_1$ $\mathcal{D}$-Lipschitzian and satisfies

$$\|Ax\| \le \|h\|_\infty + |p|_1\|x\|^\delta, \quad \forall x \in X. \tag{15}$$

(b) Define a second mapping B by

$$Bx(t) = \sigma(t)\phi(x(t)) + \int_a^t G(t,s)g\big(s, x(s)\big)\, ds, \quad t \geq a$$

with continuous functions $\sigma\colon [a,+\infty) \to \mathbb{R}$ and $g\colon [a,+\infty) \times \mathbb{R} \to \mathbb{R}$ satisfying respectively $\lim_{t\to+\infty} \sigma(t) = 0$ and the growth assumption

$$|g(s,\xi)| \leq q(s)\psi(|\xi|), \quad \forall (s,x) \in [a,+\infty) \times \mathbb{R},$$

where $q \in L^1([a,+\infty),\mathbb{R}^+)$ and $\psi\colon \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous nondecreasing map. The kernel function $G\colon [a,+\infty)^2 \to \mathbb{R}$ satisfies

$$\limsup_{t\to+\infty} \int_a^{+\infty} |G(t,s)|q(s)\, ds = 0. \tag{16}$$

Finally, let

$$\overline{\sigma} = \sup_{t\geq a} |\sigma(t)| \quad \text{and} \quad \overline{\alpha} = \sup_{t\geq a} \alpha(t)$$

with

$$\alpha(t) = \int_a^{+\infty} |G(t,s)|q(s)\, ds.$$

With conditions (16), we can show, as in the proof of Theorem 2.1 in [8], that B is completely continuous. Moreover it satisfies

$$\|Bx\| \leq \overline{\sigma}\phi(\|x\|) + \overline{\alpha}\psi(\|x\|), \quad \forall\, x \in X. \tag{17}$$

(c) Assume that for any compact subset $K \subset \mathbb{R}$, there exists a positive constant $M_K > 0$ such that for any $x, y \in X$, and $\lambda \in [0,1)$,

$$y = \lambda AyBx \Longrightarrow (|y(t)| \leq M_K,\, \forall\, t \in K). \tag{18}$$

Then, it remains to check the Furi–Pera condition (c) in Theorem 3.3 for the mapping N defined by $Nx = y = AyBx$ with respect to a closed ball $S = \overline{\mathcal{B}}_R(0)$ for some positive constant R. Let $(x_j, \lambda_j) \in \partial S \times [0,1]$ be a sequence such that, as $j \to +\infty$, $\lambda_j \to \lambda$ and $x_j \to x$ with $x = \lambda N(x)$ and $0 \leq \lambda < 1$. We will show that $\lambda_j N x_j \in S$. Let $y_j = Nx_j = Ay_jBx_j$. Then $\lambda_j N x_j = \lambda_j A y_j B x_j = \lambda_j y_j$. To perform an estimate of $|y_j|$, write

$$y_j(t) = h(t)Bx_j(t) + Bx_j(t) \int_a^t f(s, y_j(s))\, ds$$

and notice that, as in (17)

$$|Bx_j(t)| \leq |\sigma(t)|\phi(R) + \alpha(t)\psi(R) := \gamma_R(t) \leq \overline{\gamma}_R, \quad t \geq a,$$

where $\overline{\gamma}_R = |\overline{\sigma}|\phi(R) + \overline{\alpha}\psi(R)$ and

$$\lim_{t\to+\infty} \gamma_R(t) = 0. \tag{19}$$

Then

$$|y_j(t)| \leq \gamma_R(t) \left(\|h\|_\infty + \int_a^t p(s)|y_j(s)|^\delta \, ds \right)$$

which yields

$$\frac{|y_j(t)|}{\gamma_R(t)} \leq \|h\|_\infty + (\overline{\gamma}_R)^\delta \int_a^t p(s) \left| \frac{y_j(s)}{\gamma_R(s)} \right|^\delta ds.$$

By Lemma 7.1, we deduce the upper bound

$$|y_j(t)| \leq \gamma_R(t)\Psi^{-1} \left((\overline{\gamma}_R)^\delta \int_a^t p(s)ds \right),$$

where

$$\Psi(u) = \int_{\|h\|_\infty}^u s^{-\delta} ds.$$

With (19), it follows that there exist $\overline{R} > 0$ and $t_1 > a$ such that

$$|y_j(t)| \leq \overline{R}. \tag{20}$$

This both with (18) enable us to distinguish between the cases t in a compact subset of $[a, +\infty)$ and t large enough and prove, as in example 2, that there exists some $R > 0$ such that $\|y_j\| \leq R$. Therefore $\lambda_j N x_j$ belong to S proving our claim follows.

Finally, Assumption $(\mathcal{H})$ in Theorem 3.3 is verified for

$$\begin{aligned} \Phi(r) &= r - \|B(S)\|\varphi_A(r) \geq r - [\overline{\sigma}\phi(R) + \overline{\alpha}\psi(R)]\, |p|_1 r^\delta \\ &= r \left(1 - [\overline{\sigma}\phi(R) + \overline{\alpha}\psi(R)]\, |p|_1 r^{\delta - 1}\right) \end{aligned} \tag{21}$$

which increases to positive infinity as $r \to +\infty$ for $0 < \delta < 1$. To sum up, we have proved that, under the hypotheses on f, g, and h the nonlinear integral equation

$$x(t) = \left(h(t) + \int_a^t f(s, x(s))\, ds \right) \left(\int_a^t g\big(s, x(s)\big)\, ds \right), \quad t \in [a, +\infty)$$

has a solution $x \in C_0([a, +\infty), \mathbb{R})$ by Theorem 3.3.

8 Concluding remarks

(a) The following functional integral equation

$$x(t) = \left(\frac{1}{1 + |x(\theta(t))|} \right) \left(q(t) + \int_0^{\sigma(t)} g\big(s, x(\eta(s))\big)\, ds \right) \tag{22}$$

is discussed in [10] and solutions are proven to exist in the unit ball $S = \mathcal{B}_1(0)$ under the assumptions of Theorem 1.3. Indeed, notice that all solutions of

Equation (22) are in S. By contrast, solutions x of Equation (5) satisfy $\|x\| \leq R+1$ and thus do not lie in $S = \mathcal{B}_R(0)$. As a consequence, Theorem 1.3 could not be used to solve Equation (5) though the Furi–Pera condition was satisfied and Theorem 3.1 was successfully applied in Example 1.

(b) If, in Example 3, we have rather applied Theorem 3.1 instead, we should be led to the following estimates regarding the verification of the Furi–Pera condition for the mapping $F = AB$:

$$\|Fx_j\| \leq \|Ax_j\|\|Bx_j\|\| \leq \left(|h\|_\infty + |p|_1\|x_j\|^\delta\right)\left(\overline{\sigma}\phi(\|x_j\|) + \overline{\alpha}\psi(\|x_j\|)\right)$$
$$\leq \left(|h\|_\infty + |p|_1 R\right)\left(\overline{\sigma}\phi(R) + \overline{\alpha}\psi(R)\right).$$

Therefore the Furi–Pera condition is satisfied whenever

$$\left(|h\|_\infty + |p|_1 R^\delta\right)\left(\overline{\sigma}\phi(R) + \overline{\alpha}\psi(R)\right) \leq R. \tag{23}$$

This inequality is somewhat restrictive and shows that the Schauder's fixed point theorem could be applied as well.

(c) As illustrated by examples 1-3, Theorems 3.1 and 3.3 show that, in practice, the Furi–Pera condition is easier to be used than the condition (c) in Theorem 1.3. Indeed, we can notice that assumption (c) in Examples 2 and 3 is a weak condition in the sense that we were able to make use of the Furi–Pera condition while neither Schauder's fixed point theorem nor Dhage's fixed point theorem could be applied. Moreover, fixed point theorems in Banach algebras are useful in applications when problems exhibit nonlinearities as the product of two integral functions. Many boundary value problem for second-order and higher-order nonlinear differential equations may be reduced to integral equations. Further to Examples 1-3, we refer for instance to the functional integral equations treated in [12, 13].

Acknowledgment The first author is grateful to Laboratory EDP&HM of École Normale Supérieure, Kouba in Algiers, for its constant support.

References

[1] Agarwal, R. A., Meehan, M., O'Regan, D.: Fixed Point Theory and Applications. Cambridge Tracts in Mathematics, Vol. 141, *Cambridge University Press*, 2001.

[2] Bainov, D., Simeonov, P.: Integral Inequalities and Applications. *Kluwer Academic Publishers*, 1992.

[3] Banas, J., Goebel, K.: Measure of Noncompactness in Banach Spaces. Math. and Appli., Vol. 57, *Marcel Dekker, New York*, 1980.

[4] Banas, J., Knap, Z.: *Measure of noncompactness and nonlinear integral equations of convolution type.* J. Math. Anal. Appl. **146** (1990), 353–362.

[5] Boyd, D. W., Wong, J. S. W.: *On nonlinear contractions.* Proc. Amer. Math. Soc. **20** (1969), 458–464.

[6] Burton, T. A.: *A fixed point theorem of Krasnozels'kĭi.* Appl. Math. Lett. **11** (1998), 85–88.

[7] Deimling, K.: Nonlinear Functional Analysis. *Springer Verlag, Berlin–Tokyo*, 1985.

[8] Djebali, S., Moussaoui, T.: *A class of second order BVPs on infinite intervals.* Elec. Jour. Qual. Theo. Diff. Eq. **4** (2006), 1–19.

[9] Dhage, B. C.: *On some variants of Schauder's fixed point principle and applications to nonlinear integral equations.* Jour. Math. Phys. Sci. **2**, 5 (1988), 603–611.

[10] Dhage, B. C.: *On a fixed point theorem in Banach algebras with applications.* Appl. Math. Lett. **18** (2005), 273–280.

[11] Dhage, B. C.: *Some nonlinear alternatives in Banach algebras with applications I.* Nonlinear Studies **12**, 3 (2005), 271–281.

[12] Dhage, B. C., Ntouyas, S. K.: *Existence results for nonlinear functional integral equations via a fixed point theorem of Krasnozels'kĭi–Schaefer type.* Nonlinear Studies **9**, 3 (2002), 307–317.

[13] Dhage, B. C., O'Regan, D.: *A fixed point theorem in Banach algebras with applications to functional integral equations.* Funct. Diff. Equ. **7**, 3-4 (2000), 259–267.

[14] Furi, M., Pera, P.: *A continuation method on locally convex spaces and applications to ODE on noncompact intervals.* Annales Polonici Mathematici **47** (1987), 331–346.

[15] Krasnozels'kĭi, M. A.: Positive Solutions of Operator Equations. *Noordhoff, Groningen*, 1964.

[16] Krasnozels'kĭi, M. A.: Integral Operators in Space of Summable Functions. *Noordhoff, Leyden*, 1976.

[17] O'Regan, D.: *Fixed-point theory for the sum of two operators.* Appl. Math. Lett. **9** (1996), 1–8.

[18] Smart, D. R.: Fixed point Theorems. *Cambridge University Press*, 1974.

[19] Zeidler, E.: Nonlinear Functional Analysis and its Applications. Vol. I: Fixed Point Theorems. *Springer Verlag, New York*, 1986.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 77–82

Remark on Properties of Bases for Additive Logratio Transformations of Compositional Data

Karel HRON

Department of Mathematical Analysis and Applications of Mathematics
Faculty of Science, Palacký University
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: hronk@seznam.cz

(Received January 9, 2008)

Abstract

The statistical analysis of compositional data, multivariate data when all its components are strictly positive real numbers that carry only relative information and having a simplex as the sample space, is in the state-of-the-art devoted to represent compositions in orthonormal bases with respect to the geometry on the simplex and thus provide an isometric transformation of the data to an usual linear space, where standard statistical methods can be used (e.g. [2], [4], [5], [9]). However, in some applications from geosciences ([14]) or statistical aspects of multicriteria evaluation theory ([13]) it seems to be convenient to use another types of bases. This paper is devoted to describe its basic properties and illustrate the results on an example.

Key words: Aitchison geometry on the simplex; bases on the simplex; additive logratio transformations.

2000 Mathematics Subject Classification: 15A03, 62H99

1 Simplicial geometry

The concept of compositional data and its geometry on the simplex (called *Aitchison geometry*) is the starting point for building up statistical models for such data. This short course follows earlier developements of compositional data ([1]) and cites the present results of the active research, as summarized in [8], [11] and [12].

Definition 1 A row vector, $\boldsymbol{x} = (x_1, \ldots, x_D)$, is called *D-parts composition* when all its components are strictly positive real numbers and they carry only relative information.

The assertion that D-parts composition (or only composition in short) carry only relative information means that all the relevant information is contained in the ratios among the parts, i.e. if c is a nonzero real number, $(x_1, \ldots, x_D)$ and $(cx_1, \ldots, cx_D)$ convey essentially the same information. A way to simplify the use of compositions is to represent them in closed form, i.e. as positive vectors with constant sum κ (usually 1 or 100 in case of percentages) of the parts. As a consequence, D-parts composition can be identified with the following vector:

Definition 2 For any composition $\boldsymbol{x}$, the *closure operation of $\boldsymbol{x}$ to the constant κ* is defined as

$$\mathcal{C}(\boldsymbol{x}) = \left(\frac{\kappa x_1}{\sum_{i=1}^{D} x_i}, \ldots, \frac{\kappa x_D}{\sum_{i=1}^{D} x_i} \right).$$

Proposition 1 *The sample space of compositional data is the simplex, defined as*

$$\mathcal{S}^D = \left\{ \boldsymbol{x} = (x_1, \ldots, x_D),\ x_i > 0,\ \sum_{i=1}^{D} x_i = \kappa \right\}.$$

The basics of Aitchison geometry on the simplex are mentioned below:

Definition 3 *Perturbation* of a composition $\boldsymbol{x} = \mathcal{C}(x_1, \ldots, x_D) \in \mathcal{S}^D$ by a composition $\boldsymbol{y} = \mathcal{C}(y_1, \ldots, y_D) \in \mathcal{S}^D$ is a composition

$$\boldsymbol{x} \oplus \boldsymbol{y} = \mathcal{C}(x_1 y_1, \ldots, x_D y_D).$$

Definition 4 *Power transformation* of a composition $\boldsymbol{x} \in \mathcal{S}^D$ by a constant $\alpha \in \mathbb{R}$ is a composition

$$\alpha \odot \boldsymbol{x} = \mathcal{C}(x_1^{\alpha}, \ldots, x_D^{\alpha}).$$

Proposition 2 *The simplex with the perturbation operation and the power transformation, $(\mathcal{S}^D, \oplus, \odot)$, is a vector space.*

The analogy between real vector space and the simplex leads to a definition of compositional (straight) line, based on operations of perturbation and power transformation, as the compositions $\boldsymbol{x}(t)$, $t \in \mathbb{R}$, satisfying

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 \oplus (t \odot \boldsymbol{u}),$$

with starting point $\boldsymbol{x}_0$ and with direction given by the composition $\boldsymbol{u}$.

Let us remark, that the neutral element is the composition $\boldsymbol{n} = \mathcal{C}(1, \ldots, 1) = (\frac{1}{D}, \ldots, \frac{1}{D})$. The vector structure of $\mathcal{S}^D$ allows us to use the concepts of linear dependence and independence.

Definition 5 A set of m compositions in $\mathcal{S}^D$, $\boldsymbol{x}_1, \ldots \boldsymbol{x}_m$, is said to be *linearly perturbation-dependent* if there exist scalars $\alpha_1, \ldots, \alpha_m$ not all zero, such that

$$(\alpha_1 \odot \boldsymbol{x}_1) \oplus \cdots \oplus (\alpha_m \odot \boldsymbol{x}_m) = \boldsymbol{n}.$$

If no such scalars exist, the set is called *linearly perturbation-independent.*

In simplex $\mathcal{S}^D$, the maximum of perturbation-independent compositions is $D-1$. Thus, $\mathcal{S}^D$ is a vector space of dimension $D-1$.

Definition 6 If compositions $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{D-1}$ are perturbation-independent, they constitute a *(simplicial) basis* of $\mathcal{S}^D$, i.e. each composition $\boldsymbol{x} \in \mathcal{S}^D$ can be expressed as

$$\boldsymbol{x} = (\alpha_1 \odot \boldsymbol{e}_1) \oplus \cdots \oplus (\alpha_{D-1} \odot \boldsymbol{e}_{D-1})$$

for some coefficients α_i, $i = 1, \ldots, D-1$, that are termed *coordinates* with respect to the basis.

For deeper investigation of the bases on the simplex, we introduce further the concepts of inner product and norm in Aitchison geometry that enable us to use concepts of orthogonality and orthonormality of the bases.

Definition 7 *Inner product* of $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}^D$,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_a = \frac{1}{D} \sum_{i=1}^{D-1} \sum_{j=i+1}^{D} \ln \frac{x_i}{x_j} \ln \frac{y_i}{y_j} = \sum_{i=1}^{D} \ln \frac{x_i}{g(\boldsymbol{x})} \ln \frac{y_i}{g(\boldsymbol{y})},$$

and *norm* of $\boldsymbol{x} \in \mathcal{S}^D$,

$$\|\boldsymbol{x}\|_a = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle_a},$$

where $g(\boldsymbol{x}) = (x_1 \ldots x_D)^{\frac{1}{D}}$ denotes the geometric mean of the parts of the compositional vector in the argument.

It is easy to see, that using orthonormal bases on the simplex, all operations and metric concepts like perturbation, power transformation, inner product and norm are translated into coordinates as ordinary vector operations (sum of two vectors and multiplication of a vector by a scalar). See [6], [7] for details.

As consequence of the mentioned concepts we obtain the following definition:

Definition 8 The cosine of the *angle* $\angle(\boldsymbol{x}, \boldsymbol{y})_a$ *between two compositions* $\boldsymbol{x}$, $\boldsymbol{y} \in \mathcal{S}^D$, satisfying $\boldsymbol{x} \neq \boldsymbol{n}$, $\boldsymbol{x} \neq \boldsymbol{y}$, is expressible as

$$\cos \angle(\boldsymbol{x}, \boldsymbol{y})_a = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle_a}{\|\boldsymbol{x}\|_a \|\boldsymbol{y}\|_a}.$$

2 Bases for additive logratio transformations

Let us have a generating system of compositions in the simplex, $\boldsymbol{w}_1, \dots, \boldsymbol{w}_D$, where $\boldsymbol{w}_i = \mathcal{C}(1, 1 \dots, \mathrm{e}, \dots, 1)$ (the number e, base of natural logarithm, is placed in the i-th column, $i = 1, \dots, D$). Then, taking any $D-1$ vectors, we obtain a basis, e.g. $\boldsymbol{w}_1, \dots, \boldsymbol{w}_{D-1}$, and any vector $\boldsymbol{x} \in \mathcal{S}^D$ can be written as

$$\boldsymbol{x} = \ln \frac{x_1}{x_D} \odot (\mathrm{e}, 1, \dots, 1, 1) \oplus \ln \frac{x_2}{x_D} \odot (1, \mathrm{e}, 1, \dots, 1) \oplus \ln \frac{x_{D-1}}{x_D} \odot (1, 1 \dots, 1, \mathrm{e}).$$

The mentioned basis has the following properties:

Theorem 1 *Let $\boldsymbol{w}_1, \dots, \boldsymbol{w}_{D-1}$ be the basis defined above. Then for $1 \le i, j \le D-1$, $i \ne j$,*

$$\langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle_a = -\frac{1}{D}, \quad \|\boldsymbol{w}_i\|_a^2 = \frac{D-1}{D}, \quad \cos\angle(\boldsymbol{w}_i, \boldsymbol{w}_j)_a = -\frac{1}{D-1}.$$

Proof We use the inner product in the form

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_a = \sum_{k=1}^{D} \ln \frac{x_k}{g(\boldsymbol{x})} \ln \frac{y_k}{g(\boldsymbol{y})}$$

for any $\boldsymbol{x}$, $\boldsymbol{y} \in \mathcal{S}^D$. Thus, in our case,

$$\langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle_a = \sum_{k=1, k \ne i,j}^{D} \ln \frac{1}{\sqrt[D]{\mathrm{e}}} \ln \frac{1}{\sqrt[D]{\mathrm{e}}} + 2 \ln \frac{\mathrm{e}}{\sqrt[D]{\mathrm{e}}} \ln \frac{1}{\sqrt[D]{\mathrm{e}}} = \frac{D-2}{D^2} - \frac{2(D-1)}{D^2} = -\frac{1}{D}.$$

Analogously

$$\|\boldsymbol{w}_i\|_a^2 = \sum_{k=1, k \ne i,j}^{D} \left(\ln \frac{1}{\sqrt[D]{\mathrm{e}}}\right)^2 + \left(\ln \frac{\mathrm{e}}{\sqrt[D]{\mathrm{e}}}\right)^2 = \frac{D-1}{D^2} + \frac{(D-1)^2}{D^2} = \frac{D-1}{D}.$$

The value for $\cos\angle(\boldsymbol{w}_i, \boldsymbol{w}_j)_a = -\frac{1}{D-1}$ is a simple consequence. □

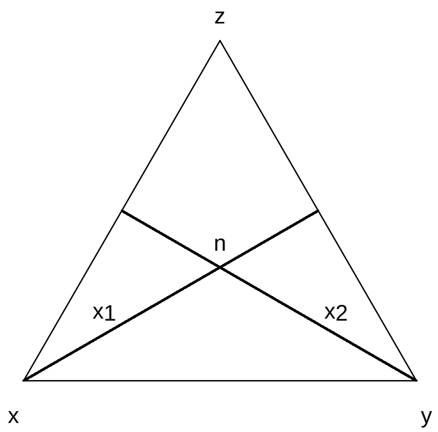
Example 1 In case of $D = 3$ we obtain $\boldsymbol{w}_1 = \mathcal{C}(e, 1, 1)$, $\boldsymbol{w}_2 = \mathcal{C}(1, e, 1)$, so thus $\|\boldsymbol{w}_1\|_a = \|\boldsymbol{w}_2\|_a = \frac{\sqrt{6}}{3}$ and $\angle(\boldsymbol{w}_1, \boldsymbol{w}_2)_a = 120°$. Compositional straight lines

$$\boldsymbol{x}_1(t) = \boldsymbol{n} \oplus (t \odot \boldsymbol{w}_1) = \mathcal{C}(\mathrm{e}^t, 1, 1),\ t \in \mathbb{R},$$

and

$$\boldsymbol{x}_2(s) = \boldsymbol{n} \oplus (s \odot \boldsymbol{w}_2) = \mathcal{C}(1, \mathrm{e}^s, 1),\ s \in \mathbb{R},$$

with neutral element $\boldsymbol{n} = \mathcal{C}(1, 1, 1)$ for starting points and directions given by $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ are displayed on Figure 1. It is clear that their common composition is just the neutral element $\boldsymbol{n}$, obtained for $t = s = 0$.



Compositional lines $\boldsymbol{x}_1(t)$ and $\boldsymbol{x}_2(s)$ with neutral element $\boldsymbol{n}$ for starting points and directions given by $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$.

The coefficients $\ln \frac{x_1}{x_D}, \ldots, \ln \frac{x_{D-1}}{x_D}$ of the above mentioned basis correspond to one member of the well known *additive logratio (alr) transformations* family, introduced by [1]. To obtain all the alr transformations, it is sufficient to choose by permutation another $D-1$ vectors from the generating system ([10]). We keep the basis chosen above, the considerations for the others are analogous. Thus, we denote by alr_D the transformation that gives the expression of a composition in additive logratio coordinates with the part x_D as ratioing part,

$$alr_D(\boldsymbol{x}) = \left(\ln \frac{x_1}{x_D}, \ln \frac{x_2}{x_D}, \ldots, \ln \frac{x_{D-1}}{x_D} \right) = \boldsymbol{y}.$$

The inverse of alr_D transformation, which gives the coordinates in the canonical basis of real space, is defined as

$$alr_D^{-1}(\boldsymbol{y}) = \mathcal{C}(\exp(y_1), \ldots, \exp(y_{D-1}), 1) = \boldsymbol{x}.$$

Let us emphasize that the alr_D (and also other transformations from the alr transformations family) is not isometric (its basis on the simplex is not orthonormal, see Theorem 1), i.e. metric concepts are not translated like as ordinary vector operations. On the other side, by many statistical methods this doesn't play a role and the remaining properties are sufficient (e.g. [1], [3], [9], for details). Moreover, the form of alr coordinates enables to use it by expert processes in multicriteria evaluation theory ([13]).

References

[1] Aitchison, J.: The statistical analysis of compositional data. *Chapman and Hall, London*, 1986.

[2] Aitchison, J., Greenacre, M.: *Biplots of compositional data.* Applied Statistics **51** (2002), 375–392.

[3] Billheimer, D.: *Compositional data in biomedical research.* In: Mateu-Figueras, G., Barceló-Vidal, C.: Compositional Data Analysis Workshop – CoDaWork'05, Proceedings, *Universitat de Girona*, 2005.

[4] Buccianti, A., Pawlowsky-Glahn, V.: *New perspectives on water chemistry and compositional data analysis.* Math. Geol. **37** (2005), 703–727.

[5] Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (eds): Compositional data analysis in the geosciences: From theory to practice. *Geological Society, London, Special Publications* **264**, 2006.

[6] Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueraz, G., Barceló-Vidal, C.: *Isometric logratio transformations for compositional data analysis.* Math. Geol. **35** (2003), 279–300.

[7] Egozcue, J. J., Pawlowsky-Glahn, V.: *Groups of parts and their balances in compositional data analysis.* Math. Geol. **37** (2005), 795–828.

[8] Egozcue, J. J., Pawlowsky-Glahn, V.: *Simplicial geometry for compositional data.* In: Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (eds.): Compositional data analysis in the geosciences: From theory to practice. *Geological Society, London, Special Publications* **264** (2006), 145–160.

[9] Filzmoser, P., Hron, K.: *Outlier detection for compositional data using robust methods.* Math. Geosci. **40** (2008), 233–248.

[10] Mateu-Figueras, G., Pawlowsky-Glahn, V., Barceló-Vidal, C.: *Distributions on the simplex.* In: Thió-Henestrosa, S., Martín-Fernández, J.A.: Compositional Data Analysis Workshop – CoDaWork'03, Proceedings, *Universitat de Girona*, 2003.

[11] Pawlowsky-Glahn, V., Egozcue, J. J.: *Geometric approach to statistical analysis on the simplex.* Stoch. Envir. Res. and Risk Ass. **15** (2001), 384–398.

[12] Pawlowsky-Glahn, V., Egozcue, J. J., Tolosana-Delgado, J.: Lecture notes on compositional data analysis. 2007, online.

[13] Talašová, J.: Fuzzy methods for multicriteria evaluation and decision making. *Publishing House of Palacký University, Olomouc*, 2006 (in Czech).

[14] Weltje, G. J.: *Ternary sandstone composition and provenance: an evaluation of the 'Dickinson model'.* In: Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V. (eds.): Compositional data analysis in the geosciences: From theory to practice. *Geological Society, London, Special Publications* **264** (2006), 79–99.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 83–100

Multivariate Models with Constraints Confidence Regions*

Lubomír KUBÁČEK

Department of Mathematical Analysis and Applications of Mathematics
Faculty of Science, Palacký University
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kubacekl@inf.upol.cz

(Received January 14, 2008)

Abstract

In multivariate linear statistical models with normally distributed observation matrix a structure of a covariance matrix plays an important role when confidence regions must be determined. In the paper it is assumed that the covariance matrix is a linear combination of known symmetric and positive semidefinite matrices and unknown parameters (variance components) which are unbiasedly estimable. Then insensitivity regions are found for them which enables us to decide whether plug-in approach can be used for confidence regions.

Key words: Multivariate model; constraints; variance components; plug-in estimator; insensitivity region.

2000 Mathematics Subject Classification: 62J05, 62H12

1 Introduction

Multivariate linear statistical models are analyzed in several monographs (cf. [1], [3], [5], etc.). Relatively small attention is given to problems of a determination of confidence regions. An attempt to contribute to a solution of the problem is the aim of the paper.

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

An apriori information on a structure of the covariance matrix in multivariate linear statistical model can be of different forms. A determination of a confidence region for the mean value parameters of the observation matrix depends essentially on this structure.

In the paper it is assumed that the covariance matrix is a linear combination of known symmetric positive semidefinite matrices and unknown, however unbiasedly estimable, coefficients (variance components). Then the plug-in approach is used for a confidence regions. For a decision whether this approach is admisible, the insensitivity regions are determined.

In the following text the models

$$\mathrm{vec}(\underline{\mathbf{Y}}) \sim N_{nm}[(\mathbf{I} \otimes \mathbf{X})\mathrm{vec}(\mathbf{B}), \mathbf{\Sigma} \otimes \mathbf{I}], \quad \mathbf{H}_1\mathbf{B}\mathbf{H}_2 + \mathbf{H}_0 = \mathbf{0} \tag{1}$$

$$\mathrm{vec}(\underline{\mathbf{Y}}) \sim N_{nm}[(\mathbf{I} \otimes \mathbf{X})\mathrm{vec}(\mathbf{B}), \mathbf{I} \otimes \mathbf{\Sigma}], \quad \mathbf{H}_1\mathbf{B}\mathbf{H}_2 + \mathbf{H}_0 = \mathbf{0} \tag{2}$$

$$\mathrm{vec}(\underline{\mathbf{Y}}) \sim N_{nr}[(\mathbf{Z}' \otimes \mathbf{X})\mathrm{vec}(\mathbf{B}), \mathbf{\Sigma} \otimes \mathbf{I}], \quad \mathbf{H}_1\mathbf{B}\mathbf{H}_2 + \mathbf{H}_0 = \mathbf{0} \tag{3}$$

$$\mathrm{vec}(\underline{\mathbf{Y}}) \sim N_{nr}[(\mathbf{Z}' \otimes \mathbf{X})\mathrm{vec}(\mathbf{B}), \mathbf{I} \otimes \mathbf{\Sigma}], \quad \mathbf{H}_1\mathbf{B}\mathbf{H}_2 + \mathbf{H}_0 = \mathbf{0}. \tag{4}$$

will be considered.

Here $\underline{\mathbf{Y}}$ is an $n \times m$ and $n \times r$, respectively, matrix (observation matrix) normally distributed, $\mathrm{vec}(\underline{\mathbf{Y}})$ is the vector composed of the columns of the matrix $\underline{\mathbf{Y}}$, $\mathbf{Z}, \mathbf{X}, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_0$ are known matrices of proper dimensions, $\mathbf{B}$ is a matrix of unknown parameters and $\mathbf{\Sigma}$ is a matrix of the structure $\mathbf{\Sigma} = \sum_{i=1}^{p} \vartheta_i \mathbf{V}_i$, $p \geq 2$. The notation $\mathrm{vec}(\underline{\mathbf{Y}}) \sim_{nm} [(\mathbf{I} \otimes \mathbf{X})\,\mathrm{vec}(\mathbf{B}), \mathbf{\Sigma} \otimes \mathbf{I}]$ means that the matrix need not be normally distributed. The matrices $\mathbf{V}_1, \ldots, \mathbf{V}_p$, are given, symmetric and positive semidefinite, $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_p)' \in \underline{\vartheta}$, is unknown vector parameter, where $\underline{\vartheta}$ is an open set in R^p (p-dimensional Euclidean space). The matrix $\mathbf{H}_0$ must satisfy the condition

$$\mathrm{vec}(\mathbf{H}_0) \in \mathcal{M}(\mathbf{H}_2' \otimes \mathbf{H}_1).$$

For the sake of simplicity either the matrix $\mathbf{H}_1$, or the matrix $\mathbf{H}_2$ will be considered to be the identity matrix $\mathbf{I}$.

Further symbols are of the following meaning.

$\mathcal{M}(\mathbf{A}_{m,n}) = \{\mathbf{Au} : \mathbf{u} \in R^n\}$ is the column subspace of the matrix $\mathbf{A}$, $\mathbf{P}_X = \mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}$, $\mathbf{A}^-$ is a generalized inverse (g-inverse) of the matrix $\mathbf{A}$, i.e. $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$, $\mathbf{M}_X = \mathbf{I} - \mathbf{P}_X$, $\mathbf{A}^+$ is the Moore–Penrose g-inverse of the matrix $\mathbf{A}$, i.e. $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$, $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$, $\mathbf{A}\mathbf{A}^+ = (\mathbf{A}\mathbf{A}^+)'$, $\mathbf{A}^+\mathbf{A} = (\mathbf{A}^+\mathbf{A})'$. Frequently used notation $(\mathbf{M}_X\mathbf{\Sigma}\mathbf{M}_X)^+$, means therefore

$$\begin{aligned}(\mathbf{M}_X\mathbf{\Sigma}\mathbf{M}_X)^+ &= \mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^-\mathbf{X}'\mathbf{\Sigma}^{-1}, \quad \mathbf{\Sigma} \ \text{ is p.d.},\\ &= \mathbf{\Sigma}^+ - \mathbf{\Sigma}^+\mathbf{X}(\mathbf{X}\mathbf{\Sigma}^-\mathbf{X})^-\mathbf{X}\mathbf{\Sigma}^+, \quad \mathcal{M}(\mathbf{X}) \subset \mathcal{M}(\mathbf{\Sigma}),\\ &= \mathbf{T}^+ - \mathbf{T}^+\mathbf{X}(\mathbf{X}'\mathbf{T}^-\mathbf{X})^-\mathbf{X}'\mathbf{T}^+, \quad \mathbf{T} = \mathbf{\Sigma} + \mathbf{X}\mathbf{X}', \quad \text{otherwise.}\end{aligned}$$

If the matrix $\mathbf{B}$ is unbiasedly estimable, then the symbol $\widehat{\widehat{\mathbf{B}}}$ denotes the best linear unbiased estimator (BLUE) of the matrix $\mathbf{B}$. ($\widehat{\widehat{\ }}$ is used in order to emphasize that the estimator respects the constraints; $\widehat{\mathbf{B}}$ is the BLUE which does not respect the constraints). If the matrix $\mathbf{B}$ is not unbiasedly estimable, however the BLUE exists for the matrix $\mathbf{XB}$, then the symbol $\widehat{\widehat{\mathbf{XB}}}$ is used. The space of all $m \times n$ matrices is $\mathcal{M}_{m,n}$.

2 Estimation of the variance components

In the following text the structure of the matrix $\mathbf{\Sigma}$ is assumed to be $\mathbf{\Sigma} = \sum_{i=1}^{p} \vartheta_i \mathbf{V}_i$. In estimation of the variance components vector $\boldsymbol{\vartheta} = (\vartheta_1, \dots, \vartheta_p)'$, the following lemma will be used.

Lemma 2.1 *Let the univariate universal linear statistical model with constraints, i.e.*

$$\mathbf{Y} \sim_n (\mathbf{X}\boldsymbol{\beta}, \sum_{i=1}^{p} \vartheta_i \mathbf{V}_i), \quad \mathbf{h} + \mathbf{H}\boldsymbol{\beta} = \mathbf{0},$$

be considered. Here the $n \times k$ matrix $\mathbf{X}$, $n \times n$ symmetric and p.s.d. matrices $\mathbf{V}_1, \dots, \mathbf{V}_p$ and the $q \times k$ matrix $\mathbf{H}$ are given. Also the q-dimensional vector $\mathbf{h}$ is given. Then the $\boldsymbol{\vartheta}_0$-MINQUE (minimum norm quadratic unbiased estimator) of the vector $\boldsymbol{\vartheta}$ is

$$\widehat{\boldsymbol{\vartheta}} = \mathbf{S}^{-1}_{(M_{XM_{H'}} \Sigma_0 M_{XM_{H'}})^+} \boldsymbol{\gamma},$$

where

$$\begin{aligned}
\boldsymbol{\gamma} &= (\gamma_1, \dots, \gamma_p)', \\
\gamma_i &= [\mathbf{Y} + \mathbf{XH}'(\mathbf{HH}')^+\mathbf{h}]' \left(\mathbf{M}_{XM_{H'}} \mathbf{\Sigma}_0 \mathbf{M}_{XM_{H'}}\right)^+ \mathbf{V}_i \\
&\quad \times \left(\mathbf{M}_{XM_{H'}} \mathbf{\Sigma}_0 \mathbf{M}_{XM_{H'}}\right)^+ [\mathbf{Y} + \mathbf{XH}'(\mathbf{HH}')^+\mathbf{h}], \quad i = 1, \dots, p, \\
\mathbf{\Sigma}_0 &= \sum_{i=1}^{p} \vartheta_{0,i} \mathbf{V}_i,
\end{aligned}$$

$$\left\{\mathbf{S}_{(M_{XM_{H'}} \Sigma_0 M_{XM_{H'}})^+}\right\}_{i,j} =$$

$$= \mathrm{Tr}\left[\mathbf{V}_i \left(\mathbf{M}_{XM_{H'}} \mathbf{\Sigma}_0 \mathbf{M}_{XM_{H'}}\right)^+ \mathbf{V}_j \left(\mathbf{M}_{XM_{H'}} \mathbf{\Sigma}_0 \mathbf{M}_{XM_{H'}}\right)^+\right], \quad i, j = 1, \dots, p,$$

and $\boldsymbol{\vartheta}_0 = (\vartheta_{0,1}, \dots, \vartheta_{0,p})'$ is an approximate value of the vector $\boldsymbol{\vartheta}$.

The $\boldsymbol{\vartheta}_0$-MINQUE of the vector $\boldsymbol{\vartheta}$ exists iff the matrix $\mathbf{S}_{(M_{XM_{H'}} \Sigma_0 M_{XM_{H'}})^+}$ is regular.

Proof cf. in [2], [4]. □

The formulae for the multivariate models with constraints can be now rewritten directly from this lemma.

Theorem 2.2 *(i) Let in the model (1) with* $\mathbf{H}_2 = \mathbf{I}$, *the matrix*

$$\mathbf{S}_{\left[M_{(I\otimes X)(I\otimes M_{H_1'})}(\Sigma_0\otimes I)M_{(I\otimes X)(I\otimes M_{H_1'})}\right]^+}$$

be regular. Then

$$\mathbf{S}_{\left[M_{(I\otimes X)(I\otimes M_{H_1'})}(\Sigma_0\otimes I)M_{(I\otimes X)(I\otimes M_{H_1'})}\right]^+} = \mathrm{Tr}\left(\mathbf{M}_{XM_{H_1'}}\right)\mathbf{S}_{\Sigma_0^+}$$

and the $\boldsymbol{\vartheta}_0$*-MINQUE is*

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}} &= \left[\mathrm{Tr}(\mathbf{M}_{XM_{H_1'}})\mathbf{S}_{\Sigma^+}\right]^{-1}\boldsymbol{\gamma},\\
\left\{\mathbf{S}_{\Sigma_0^+}\right\}_{i,j} &= \mathrm{Tr}(\mathbf{V}_i\boldsymbol{\Sigma}_0^+\mathbf{V}_j\boldsymbol{\Sigma}_0^+), \quad i,j = 1,\dots,p,\\
\boldsymbol{\gamma} &= (\gamma_1,\dots,\gamma_p)'\\
\gamma_i &= \mathrm{Tr}\left\{[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0]'\mathbf{M}_{XM_{H_1'}}[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0]\right.\\
&\quad \left.\times\, \boldsymbol{\Sigma}_0^+\mathbf{V}_i\boldsymbol{\Sigma}_0^+\right\}, \quad i = 1,\dots,p.
\end{aligned}$$

(ii) Let in the model (1) with $\mathbf{H}_1 = \mathbf{I}$, *the matrix*

$$\mathbf{S}_{\left[M_{M_{H_2}\otimes X}(\Sigma_0\otimes I)M_{M_{H_2}\otimes X}\right]^+}$$

be regular. Then

$$\mathbf{S}_{\left[M_{M_{H_2}\otimes X}(\Sigma_0\otimes I)M_{M_{H_2}\otimes X}\right]^+} = [n - r(\mathbf{X})]\mathbf{S}_{\Sigma_0^+} + r(\mathbf{X})\mathbf{S}_{(P_{H_2}\Sigma_0 P_{H_2})^+}$$

and the $\boldsymbol{\vartheta}_0$*-MINQUE is*

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}} &= \left\{[n - r(\mathbf{X})]\mathbf{S}_{\Sigma_0^+} + r(\mathbf{X})\mathbf{S}_{(P_{H_2}\Sigma_0 P_{H_2})^+}\right\}^{-1}\boldsymbol{\gamma},\\
\boldsymbol{\gamma} &= (\gamma_1,\dots,\gamma_p)',\\
\gamma_i &= \mathrm{Tr}(\underline{\mathbf{Y}}'\mathbf{M}_X\underline{\mathbf{Y}}\boldsymbol{\Sigma}_0^+\mathbf{V}_i\boldsymbol{\Sigma}_0^+) + \mathrm{Tr}\left\{[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2']'\mathbf{P}_X\right.\\
&\quad \left.\times\, [\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2'](\mathbf{P}_{H_2}\boldsymbol{\Sigma}_0\mathbf{P}_{H_2})^+\right\}, \quad i = 1,\dots,p
\end{aligned}$$

and

$$\left\{\mathbf{S}_{P_{H_2}\Sigma_0 P_{H_2})^+}\right\}_{i,j} = \mathrm{Tr}\left[\mathbf{V}_i(\mathbf{P}_{H_2}\boldsymbol{\Sigma}_0\mathbf{P}_{H_2})^+\mathbf{V}_j(\mathbf{P}_{H_2}\boldsymbol{\Sigma}_0\mathbf{P}_{H_2})^+\right], \quad i,j = 1,\dots,p.$$

Proof (i) It is implied by Lemma 2.1 and by the equality

$$\left[\mathbf{M}_{I\otimes(XM_{H_1'})}(\boldsymbol{\Sigma}_0\otimes\mathbf{I})\mathbf{M}_{I\otimes(XM_{H_1'})}\right]^+ = \boldsymbol{\Sigma}_0^+\otimes\mathbf{M}_{XM_{H_1'}}.$$

In (ii) the equality

$$\left[\mathbf{M}_{M_{H_2}\otimes X}(\boldsymbol{\Sigma}_0\otimes\mathbf{I})\mathbf{M}_{M_{H_2}\otimes X}\right]^+ = \boldsymbol{\Sigma}_0^+\otimes\mathbf{M}_X + (\mathbf{P}_{H_2}\boldsymbol{\Sigma}_0\mathbf{P}_{H_2})^+\otimes\mathbf{P}_X$$

must be used. □

Theorem 2.3 *(i) Let in the model (2) with* $\mathbf{H}_2 = \mathbf{I}$*, the matrix*

$$\mathbf{S}_{M_{[I\otimes(XM_{H_1'})]}(I\otimes\Sigma_0)M_{[I\otimes(XM_{H_1'})]}+}$$

be regular. Then

$$\mathbf{S}_{M_{[I\otimes(XM_{H_1'})]}(I\otimes\Sigma_0)M_{[I\otimes(XM_{H_1'})]}+} = m\mathbf{S}_{(M_{XM_{H_1'}}\Sigma_0 M_{XM_{H_1'}})^+}$$

and the $\boldsymbol{\vartheta}_0$*-MINQUE is*

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}} &= \left(m\mathbf{S}_{(M_{XM_{H_1'}}\Sigma_0 M_{XM_{H_1'}})^+}\right)^{-1}\boldsymbol{\gamma},\\
\boldsymbol{\gamma} &= (\gamma_1,\dots,\gamma_p)',\\
\gamma_i &= \operatorname{Tr}\left\{[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0]'\left(\mathbf{M}_{XM_{H_1'}}\boldsymbol{\Sigma}_0\mathbf{M}_{XM_{H_1'}}\right)^+\mathbf{V}_i\right.\\
&\quad \left.\times\left(\mathbf{M}_{XM_{H_1'}}\boldsymbol{\Sigma}_0\mathbf{M}_{XM_{H_1'}}\right)^+[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0]\right\},\\
&\quad i = 1,\dots,p.
\end{aligned}$$

(ii) Let in the model (2) with $\mathbf{H}_1 = \mathbf{I}$*, the matrix*

$$\mathbf{S}_{[M_{(M_{H_2}\otimes X)}(I\otimes\Sigma_0)M_{(M_{H_2}\otimes X)}]^+}$$

be regular. Then

$$\mathbf{S}_{[M_{(M_{H_2}\otimes X)}(I\otimes\Sigma_0)M_{(M_{H_2}\otimes X)}]^+} = [m - r(\mathbf{H}_2)]\mathbf{S}_{(M_X\Sigma_0 M_X)^+} + r(\mathbf{H}_2)\mathbf{S}_{\Sigma^+}$$

and the $\boldsymbol{\vartheta}_0$*-MINQUE is*

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}} &= \left\{[m - r(\mathbf{H}_2)]\mathbf{S}_{(M_X\Sigma_0 M_X)^+} + r(\mathbf{H}_2)\mathbf{S}_{\Sigma^+}\right\}^{-1}\boldsymbol{\gamma},\\
\boldsymbol{\gamma} &= (\gamma_1,\dots,\gamma_p)',\\
\gamma_i &= \operatorname{Tr}[\underline{\mathbf{Y}}'(\mathbf{M}_X\boldsymbol{\Sigma}_0\mathbf{M}_X)^+\mathbf{V}_i(\mathbf{M}_X\boldsymbol{\Sigma}_0\mathbf{M}_X)^+\underline{\mathbf{Y}}\mathbf{M}_{H_2}]\\
&\quad + \operatorname{Tr}\left\{[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2']'\boldsymbol{\Sigma}_0^+\mathbf{V}_i\boldsymbol{\Sigma}_0^+[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2']\mathbf{P}_{H_2}\right\}
\end{aligned}$$

Proof (i) The obvious equality

$$\left[\mathbf{M}_{[I\otimes(XM_{H_1'})}(I\otimes\Sigma_0)\mathbf{M}_{[I\otimes(XM_{H_1'})}\right]^+ = \mathbf{I}\otimes\left(\mathbf{M}_{XM_{H_1'}}\boldsymbol{\Sigma}_0\mathbf{M}_{XM_{H_1'}}\right)^+$$

must be taken into account.

(ii) The equality

$$\left[\mathbf{M}_{M_{H_2}\otimes X}(\mathbf{I}\otimes\boldsymbol{\Sigma}_0)\mathbf{M}_{M_{H_2}\otimes X}\right]^+ = \mathbf{M}_{H_2}\otimes(\mathbf{M}_X\boldsymbol{\Sigma}_0\mathbf{M}_X)^+ + \mathbf{P}_{H_2}\otimes\boldsymbol{\Sigma}_0^+$$

must be taken into account. □

Theorem 2.4 *(i) Let in the model (3) with $\mathbf{H}_2 = \mathbf{I}$ the matrix*

$$\mathbf{S}_{[M_{(Z'\otimes X)(I\otimes M_{H_1'})}(\Sigma_0\otimes I)M_{(Z'\otimes X)(I\otimes M_{H_1'})}]^+}$$

be regular. Then

$$\mathbf{S}_{[M_{(Z'\otimes X)(I\otimes M_{H_1'})}(\Sigma_0\otimes I)M_{(Z'\otimes X)(I\otimes M_{H_1'})}]^+} = \\ = \mathrm{Tr}\left(\mathbf{M}_{XM_{H_1'}}\right)\mathbf{S}_{\Sigma_0^+} + \mathrm{Tr}\left(\mathbf{P}_{XM_{H_1'}}\right)\mathbf{S}_{(M_{Z'}\Sigma_0 M_{Z'})^+}$$

and the $\boldsymbol{\vartheta}_0$-MINQUE is

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}} &= \left[\mathrm{Tr}\left(\mathbf{M}_{XM_{H_1'}}\right)\mathbf{S}_{\Sigma_0^+} + \mathrm{Tr}\left(\mathbf{P}_{XM_{H_1'}}\right)\mathbf{S}_{(M_{Z'}\Sigma_0 M_{Z'})^+}\right]^{-1}\boldsymbol{\gamma},\\
\boldsymbol{\gamma} &= \gamma_1,\dots,\gamma_p)',\\
\gamma_i &= \mathrm{Tr}\Big\{[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\mathbf{Z}]'\mathbf{M}_{XM_{H_1'}}[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\mathbf{Z}]\\
&\quad \times \boldsymbol{\Sigma}_0^+\mathbf{V}_i\boldsymbol{\Sigma}_0^+\Big\} + \mathrm{Tr}\Big\{[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\mathbf{Z}]'\mathbf{P}_{XM_{H_1'}}\\
&\quad \times [\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\mathbf{Z}]\left(\mathbf{M}_{Z'}\boldsymbol{\Sigma}_0\mathbf{M}_{Z'}\right)^+\mathbf{V}_i\left(\mathbf{M}_{Z'}\boldsymbol{\Sigma}_0\mathbf{M}_{Z'}\right)^+\Big\},\\
&\quad i = 1,\dots,p.
\end{aligned}$$

(ii) Let in the model (3) with $\mathbf{H}_1 = \mathbf{I}$ the matrix

$$\mathbf{S}_{[M_{(Z'\otimes X)(M_{H_2}\otimes I)}(\Sigma_0\otimes I)M_{(Z'\otimes X)(M_{H_2}\otimes I)}]^+}$$

be regular. Then

$$\mathbf{S}_{[M_{(Z'\otimes X)(M_{H_2}\otimes I)}(\Sigma_0\otimes I)M_{(Z'\otimes X)(M_{H_2}\otimes I)}]^+} = \\ = [n - r(\mathbf{X})]\mathbf{S}_{\Sigma_0^+} + r(\mathbf{X})\mathbf{S}_{(M_{Z'M_{H_2}}\Sigma_0 M_{Z'M_{H_2}})^+}$$

and the $\boldsymbol{\vartheta}_0$-MINQUE is

$$\begin{aligned}
\widehat{\boldsymbol{\vartheta}} &= \left\{[n - r(\mathbf{X})]\mathbf{S}_{\Sigma_0^+} + r(\mathbf{X})\mathbf{S}_{(M_{Z'M_{H_2}}\Sigma_0 M_{Z'M_{H_2}})^+}\right\}^{-1}\boldsymbol{\gamma},\\
\boldsymbol{\gamma} &= (\gamma_1,\dots,\gamma_p)',\\
\gamma_i &= \mathrm{Tr}(\underline{\mathbf{Y}}'\mathbf{M}_X\underline{\mathbf{Y}}\boldsymbol{\Sigma}_0^+\mathbf{V}_i\boldsymbol{\Sigma}_0^+) + \mathrm{Tr}\Big\{[\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2'\mathbf{Z}]'\mathbf{P}_X\\
&\quad \times [\underline{\mathbf{Y}} + \mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2'\mathbf{Z}]\left(\mathbf{M}_{Z'M_{H_2}}\boldsymbol{\Sigma}_0\mathbf{M}_{Z'M_{H_2}}\right)^+\mathbf{V}_i\\
&\quad \times \left(\mathbf{M}_{Z'M_{H_2}}\boldsymbol{\Sigma}_0\mathbf{M}_{Z'M_{H_2}}\right)^+\Big\}, \quad i = 1,\dots,p.
\end{aligned}$$

Proof (i) It is necessary to take into account the equality

$$\left[\mathbf{M}_{(Z'\otimes X)(I\otimes M_{H_1'})}(\boldsymbol{\Sigma}_0\otimes\mathbf{I})\mathbf{M}_{(Z'\otimes X)(I\otimes M_{H_1'})}\right]^+ = \\ = \boldsymbol{\Sigma}_0^+\otimes\mathbf{M}_{XM_{H_1'}} + (\mathbf{M}_{Z'}\boldsymbol{\Sigma}_0\mathbf{M}_{Z'})^+\otimes\mathbf{P}_{XM_{H_1'}}.$$

(ii) The equality

$$\Big[\mathbf{M}_{(Z'\otimes X)(M_{H_2}\otimes I)}(\boldsymbol{\Sigma}_0\otimes\mathbf{I})\mathbf{M}_{(Z'\otimes X)(M_{H_2}\otimes I)}\Big]^+ = \\ = \boldsymbol{\Sigma}_0^+\otimes\mathbf{M}_X + (\mathbf{M}_{Z'M_{H_2}}\boldsymbol{\Sigma}_0\mathbf{M}_{Z'M_{H_2}})^+\otimes\mathbf{P}_X$$

must be utilized. □

Theorem 2.5 *(i) Let in the model (4) with* $\mathbf{H}_2=\mathbf{I}$ *the matrix*

$$\mathbf{S}_{[M_{(Z'\otimes X)(I\otimes M_{H_1'})}(I\otimes\Sigma_0)M_{(Z'\otimes X)(I\otimes M_{H_1'})}]^+}$$

be regular. Then

$$\begin{aligned}\widehat{\boldsymbol{\vartheta}} &= \mathbf{S}^{-1}_{[M_{(Z'\otimes X)(I\otimes M_{H_1'})}(I\otimes\Sigma_0)M_{(Z'\otimes X)(I\otimes M_{H_1'})}]^+}\boldsymbol{\gamma}, \qquad \boldsymbol{\gamma}=(\gamma_1,\dots,\gamma_p)',\\ \gamma_i &= \mathrm{Tr}\left\{\Big[\underline{\mathbf{Y}}+\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\Big]'\boldsymbol{\Sigma}_0^+\mathbf{V}_i\boldsymbol{\Sigma}_0^+\Big[\underline{\mathbf{Y}}+\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\Big]\mathbf{M}_{Z'}\right\}\\ &\quad+\mathrm{Tr}\left\{\Big[\underline{\mathbf{Y}}+\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\Big]'\Big(\mathbf{M}_{XM_{H_1'}}\boldsymbol{\Sigma}_0\mathbf{M}_{XM_{H_1'}}\Big)^+\mathbf{V}_i\right.\\ &\quad\left.\times\Big(\mathbf{M}_{XM_{H_1'}}\boldsymbol{\Sigma}_0\mathbf{M}_{XM_{H_1'}}\Big)^+\Big[\underline{\mathbf{Y}}+\mathbf{H}_1'(\mathbf{H}_1\mathbf{H}_1')^+\mathbf{H}_0\Big]\mathbf{P}_{Z'}\right\},\quad i=1,\dots,p\end{aligned}$$

and

$$\mathbf{S}_{[M_{(Z'\otimes X)(I\otimes M_{H_1'})}(I\otimes\Sigma_0)M_{(Z'\otimes X)(I\otimes M_{H_1'})}]^+} = \\ = \mathrm{Tr}(\mathbf{M}_{Z'})\mathbf{S}_{\Sigma_0^+}+\mathrm{Tr}(\mathbf{P}_{Z'})\mathbf{S}_{(M_{XM_{H_1'}}\Sigma_0M_{XM_{H_1'}})^+}.$$

(ii) If in the model (4) $\mathbf{H}_1=\mathbf{I}$ *and the matrix*

$$\mathbf{S}_{[M_{(Z'M_{H_2})\otimes X}(I\otimes\Sigma_0)M_{(Z'M_{H_2})\otimes X}]^+}$$

is regular, then

$$\widehat{\boldsymbol{\vartheta}} = \mathbf{S}^{-1}_{[M_{(Z'M_{H_2})\otimes X}(I\otimes\Sigma_0)M_{(Z'M_{H_2})\otimes X}]^+}\boldsymbol{\gamma}, \qquad \boldsymbol{\gamma}=(\gamma_1,\dots,\gamma_p)',$$

$$\gamma_i=\mathrm{Tr}\left\{\Big[\underline{\mathbf{Y}}+\mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2'\Big]'\boldsymbol{\Sigma}_0^+\mathbf{V}_i\boldsymbol{\Sigma}_0^+\Big[\underline{\mathbf{Y}}+\mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2'\Big]\mathbf{M}_{Z'M_{H_2}}\right\}$$
$$+\mathrm{Tr}\left\{\Big[\underline{\mathbf{Y}}+\mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2'\Big]'(\mathbf{M}_X\boldsymbol{\Sigma}_0\mathbf{M}_X)^+\Big[\underline{\mathbf{Y}}+\mathbf{X}\mathbf{H}_0(\mathbf{H}_2'\mathbf{H}_2)^+\mathbf{H}_2'\Big]\mathbf{P}_{Z'M_{H_2}}\right\}$$

and

$$\mathbf{S}_{[M_{(Z'M_{H_2})\otimes X}(I\otimes\Sigma_0)M_{(Z'M_{H_2})\otimes X}]^+}=\mathrm{Tr}(\mathbf{M}_{Z'M_{H_2}})\mathbf{S}_{\Sigma_0^+}+\mathrm{Tr}(\mathbf{P}_{Z'M_{H_2}})\mathbf{S}_{(M_X\Sigma_0M_X)^+}.$$

Proof In (i) the equality

$$\left[\mathbf{M}_{(Z'\otimes X)(I\otimes M_{H_1'})}(\mathbf{I}\otimes\mathbf{\Sigma}_0)\mathbf{M}_{(Z'\otimes X)(I\otimes M_{H_1'})}\right]^+ = \\ = \mathbf{M}_{Z'}\otimes\mathbf{\Sigma}_0 + \mathbf{P}_{Z'}\otimes(\mathbf{M}_{XM_{H_1'}}\mathbf{\Sigma}_0\mathbf{M}_{XM_{H_1'}})^+$$

must be used.

In (ii) the equality

$$\left[\mathbf{M}_{[(Z'M_{H_2})\otimes X]}(\mathbf{I}\otimes\mathbf{\Sigma}_0)\mathbf{M}_{[(Z'M_{H_2})\otimes X]}\right]^+ = \\ = \mathbf{M}_{Z'M_{H_2}}\otimes\mathbf{\Sigma}_0 + \mathbf{P}_{Z'M_{H_2}}\otimes(\mathbf{M}_X\mathbf{\Sigma}_0\mathbf{M}_X)^+$$

must be used. □

3 Confidence regions

3.1 The matrix $\mathbf{\Sigma}$ is given

In this section the observation matrix is assumed to be normally distributed. Since confidence regions for multivariate models can be directly rewritten from the formulae for univariate models, the following lemmas are given without proofs.

Lemma 3.1 *(i) Let in the model (1) with $\mathbf{H}_2 = \mathbf{I}$ the $s\times k$ matrix $\mathbf{G}_1$ and the $m\times t$ matrix $\mathbf{G}_2$ be given. Let $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ be unbiasedly estimable, i.e.*

$$\mathcal{M}(\mathbf{G}_2\otimes\mathbf{G}_1')\subset\mathcal{M}[\mathbf{I}\otimes(\mathbf{X}',\mathbf{H}_1')].$$

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

$$\mathcal{E} = \Big\{\mathbf{U}:\mathbf{U}\in\mathcal{M}_{s,t}, \mathrm{Tr}\Big((\mathbf{U}-\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}_2})'[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'})^+\mathbf{G}_1']^+ \\ \times(\mathbf{U}-\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}_2})(\mathbf{G}_2'\mathbf{\Sigma}\mathbf{G}_2)^+\Big)\le\chi_f^2(0,1-\alpha)\Big\},$$

$$f = r\{\mathrm{Var}[\mathrm{vec}(\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}_2})]\} = r(\mathbf{G}_2'\mathbf{\Sigma}\mathbf{G}_2)r[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'})^+\mathbf{G}_1'].$$

Here $\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}_2}$ is the BLUE of the matrix $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$.

(ii) Let in the model (1) with $\mathbf{H}_1 = \mathbf{I}$ the matrix $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ be unbiasedly estimable, i.e. $\mathcal{M}(\mathbf{G}_2\otimes\mathbf{G}_1')\subset\mathcal{M}(\mathbf{I}\otimes\mathbf{X}',\mathbf{H}_2\otimes\mathbf{I})$. Then the $(1-\alpha)$-confidence region is

$$\mathcal{E} = \Big\{\mathbf{U}:\mathbf{U}\in\mathcal{M}_{s,t}, \mathrm{Tr}\Big[(\mathbf{U}-\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}_2})'[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']^+(\mathbf{U}-\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}_2}) \\ \times\Big(\mathbf{G}_2'\{[\mathbf{M}_{H_2}(\mathbf{\Sigma}+\mathbf{M}_{H_2})^+\mathbf{M}_{H_2}]^+-\mathbf{M}_{H_2}\}\mathbf{G}_2\Big)^+\Big]\le\chi_f^2(0;1-\alpha)\Big\},$$

$$f = r\Big\{\mathrm{Var}\Big[\mathrm{vec}(\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}_2})\Big]\Big\} \\ = r\Big(\mathbf{G}_2'\{[\mathbf{M}_{H_2}(\mathbf{\Sigma}+\mathbf{M}_{H_2})^+\mathbf{M}_{H_2}]^+-\mathbf{M}_{H_2}\}\mathbf{G}_2\Big)r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1'].$$

Lemma 3.2 *(i) Let in the model (2) with* $\mathbf{H}_2 = \mathbf{I}$ *the* $s \times k$ *matrix* $\mathbf{G}_1$ *and the* $m \times t$ *matrix* $\mathbf{G}_2$ *be given and let* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable. Then the* $(1-\alpha)$*-confidence region is*

$$\begin{aligned}\mathcal{E} &= \Big\{\mathbf{U} : \mathbf{U} \in \mathcal{M}_{s,t} \operatorname{Tr}\Big[(\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)'\Big(\mathbf{G}_1\big\{[\mathbf{M}_{H_1'}\mathbf{X}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^+ \\ &\quad \times \mathbf{X}\mathbf{M}_{H_1'}]^+ - \mathbf{M}_{H_1'}\big\}\mathbf{G}_1'\Big)^+(\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)(\mathbf{G}_2'\mathbf{G}_2)^+\Big] \leq \chi^2_f(0; 1-\alpha)\Big\}, \\ f &= r\Big\{\operatorname{Var}\big[\operatorname{vec}(\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)\big]\Big\} \\ &= r(\mathbf{G}_2)r\Big(\mathbf{G}_1\big\{[\mathbf{M}_{H_1'}\mathbf{X}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^+\mathbf{X}\mathbf{M}_{H_1'}]^+ - \mathbf{M}_{H_1'}\big\}\mathbf{G}_1'\Big).\end{aligned}$$

(ii) Let in the model (2) with $\mathbf{H}_1 = \mathbf{I}$ *the* $s \times k$ *matrix* $\mathbf{G}_1$ *and the* $m \times t$ *matrix* $\mathbf{G}_2$ *be given and let* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable. Then the* $(1-\alpha)$*-confidence region is*

$$\begin{aligned}\mathcal{E} &= \Bigg\{\mathbf{U} : \mathbf{U} \in \mathcal{M}_{s,t}, \operatorname{Tr}\Big((\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)'\big\{\mathbf{G}_1[(\mathbf{X}'\mathbf{T}^+\mathbf{X})^+ - \mathbf{I}]\mathbf{G}_1'\big\}^+ \\ &\quad \times (\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)(\mathbf{G}_2'\mathbf{M}_{H_2}\mathbf{G}_2)^+\Big) \leq \chi^2_f(0; 1-\alpha)\Bigg\}, \\ f &= r\Big\{\operatorname{Var}\big[\operatorname{vec}(\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)\big]\Big\} = r\big\{\mathbf{G}_1[(\mathbf{X}'\mathbf{T}^+\mathbf{X})^+ - \mathbf{I}]\mathbf{G}_1'\big\}r(\mathbf{G}_2'\mathbf{M}_{H_2}\mathbf{G}_2).\end{aligned}$$

Lemma 3.3 *(i) Let in the model (3) with* $\mathbf{H}_2 = \mathbf{I}$ *the* $s \times k$ *matrix* $\mathbf{G}_1$ *and the* $m \times t$ *matrix* $\mathbf{G}_2$ *be given and let* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2 \otimes \mathbf{G}_1') \subset \mathcal{M}(\mathbf{Z} \otimes \mathbf{X}', \mathbf{I} \otimes \mathbf{H}_1')$*. Then the* $(1-\alpha)$*-confidence region is*

$$\begin{aligned}\mathcal{E} &= \Big\{\mathbf{U} : \mathbf{U} \in \mathcal{M}_{s,t}, \operatorname{Tr}\Big((\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)'\Big[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'})^+\mathbf{G}_1'\Big]^+ \\ &\quad \times (\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)\Big\{\mathbf{G}_2'\big[(\mathbf{Z}\mathbf{U}^+\mathbf{Z}')^+ - \mathbf{I}\big]\mathbf{G}_2\Big\}\Big) \leq \chi^2_f(0; 1-\alpha)\Big\}, \\ f &= r\Big\{\operatorname{Var}\big[\operatorname{vec}(\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)\big]\Big\} \\ &= r\Big[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'})^+\mathbf{G}_1'\Big]\, r\Big\{\mathbf{G}_2'\big[(\mathbf{Z}\mathbf{U}^+\mathbf{Z}')^+ - \mathbf{I}\big]\mathbf{G}_2\Big\}.\end{aligned}$$

(ii) Let in the model (3) with $\mathbf{H}_1 = \mathbf{I}$ *the* $s \times k$ *matrix* $\mathbf{G}_1$ *and the* $m \times t$ *materix* $\mathbf{G}_2$ *be given and let* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2 \otimes \mathbf{G}_1') \subset \mathcal{M}(\mathbf{Z} \otimes \mathbf{X}', \mathbf{H}_2 \otimes \mathbf{I})$*. Then the* $(1-\alpha)$*-confidence region is*

$$\begin{aligned}\mathcal{E} &= \Big\{\mathbf{U} : \mathbf{U} \in \mathcal{M}_{s,t}, \operatorname{Tr}\Big[(\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)'[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']^+(\mathbf{U} - \widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2) \\ &\quad \times \Big(\mathbf{G}_2'\Big\{\big[\mathbf{M}_{H_2}\mathbf{Z}(\boldsymbol{\Sigma} + \mathbf{Z}'\mathbf{M}_{H_2}\mathbf{Z})^+\mathbf{Z}'\mathbf{M}_{H_2}\big]^+ - \mathbf{M}_{H_2}\Big\}\mathbf{G}_2\Big)^+\Big] \leq \chi^2_f(0; 1-\alpha)\Big\}, \\ f &= r\Big\{\operatorname{Var}\big[\operatorname{vec}(\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}_2)\big]\Big\} \\ &= r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']\, r\Big(\mathbf{G}_2'\Big\{\big[\mathbf{M}_{H_2}\mathbf{Z}(\boldsymbol{\Sigma} + \mathbf{Z}'\mathbf{M}_{H_2}\mathbf{Z})^+\mathbf{Z}'\mathbf{M}_{H_2}\big]^+ - \mathbf{M}_{H_2}\Big\}\mathbf{G}_2\Big).\end{aligned}$$

Lemma 3.4 *(i) Let in the model (4) with* $\mathbf{H}_2 = \mathbf{I}$ *the* $s \times k$ *matrix* $\mathbf{G}_1$ *and* $m \times t$ *matrix* $\mathbf{G}_2$ *be given and let the matrix* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2 \otimes \mathbf{G}_1') \subset \mathcal{M}(\mathbf{Z} \otimes \mathbf{X}', \mathbf{I} \otimes \mathbf{H}_1')$. *Then the* $(1-\alpha)$*-confidence region is*

$$\mathcal{E} = \Big\{\mathbf{U} : \mathbf{U} \in \mathcal{M}_{s,t}, \mathrm{Tr}\Big[(\mathbf{U} - \widehat{\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}}_2)'\Big(\mathbf{G}_1\big\{[\mathbf{M}_{H_1'}\mathbf{X}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H_1'}\mathbf{X})^+ \times \mathbf{X}\mathbf{M}_{H_1'}]^+ - \mathbf{M}_{H_1'}\big\}\mathbf{G}_1'\Big)^+ (\mathbf{U} - \widehat{\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}}_2)[\mathbf{G}_2'(\mathbf{Z}'\mathbf{Z})^+\mathbf{G}_2]^+\Big] \le \chi_f^2(0; 1-\alpha)\Big\},$$

$$f = r\{\mathrm{Var}[\mathrm{vec}(\widehat{\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}}_2)]\} = r\Big(\mathbf{G}_1\big\{[\mathbf{M}_{H_1'}\mathbf{X}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H_1'}\mathbf{X})^+\mathbf{X}\mathbf{M}_{H_1'}]^+ - \mathbf{M}_{H_1'}\big\}\mathbf{G}_1'\Big) r[\mathbf{G}_2'(\mathbf{Z}'\mathbf{Z})^+\mathbf{G}_2].$$

(ii) Let in the model (4) with $\mathbf{H}_1 = \mathbf{I}$ *the* $s \times k$ *matrix* $\mathbf{G}_1$ *and the* $m \times t$ *matrix* $\mathbf{G}_2$ *be given and let* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2 \otimes \mathbf{G}_1') \subset \mathcal{M}(\mathbf{Z} \otimes \mathbf{X}', \mathbf{H}_2 \otimes \mathbf{I})$. *Then the* $(1-\alpha)$*-confidence region is*

$$\mathcal{E} = \Big\{\mathbf{U} : \mathbf{U} \in \mathcal{M}_{s,t}, \mathrm{Tr}\Big((\mathbf{U} - \widehat{\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}}_2)'\Big\{\mathbf{G}_1\big[(\mathbf{X}'\mathbf{T}^+\mathbf{X})^+ - \mathbf{I}\big]\mathbf{G}_1'\Big\}^+ \times (\mathbf{U} - \widehat{\widehat{\mathbf{G}_1\mathbf{B}\mathbf{G}}}_2)\big[\mathbf{G}_2'(\mathbf{M}_{H_2}\mathbf{Z}\mathbf{Z}'\mathbf{M}_{H_2})^+\mathbf{G}_2\big]^+\Big) \le \chi_f^2(0; 1-\alpha)\Big\},$$

$$f = r\Big\{\mathbf{G}_1\big[(\mathbf{X}'\mathbf{T}^+\mathbf{X})^+ - \mathbf{I}\big]\mathbf{G}_1'\Big\} r\big[\mathbf{G}_2'(\mathbf{M}_{H_2}\mathbf{Z}\mathbf{Z}'\mathbf{M}_{H_2})^+\mathbf{G}_2\big].$$

3.2 The matrix $\boldsymbol{\Sigma}$ is of the form $\sum_{i=1}^p \vartheta_i \mathbf{V}_i$

If the estimators of the variance components $\vartheta_1, \dots, \vartheta_p$, are sufficiently accurate, then confidence regions cover the functions of parameter matrix with probability sufficiently near to prescribed confidence level $1-\alpha$. How rigorous conditions on the accuracy is, the nonsensitivity region can show.

In the first step let an univariate universal linear statistical model with constraints be considered, i.e.

$$\mathbf{Y} \sim N_n\left(\mathbf{X}\boldsymbol{\beta}, \sum_{i=1}^p \vartheta_i \mathbf{V}_i\right), \quad \mathbf{H}_{q,k}\boldsymbol{\beta} + \mathbf{h}_{q,1} = \mathbf{0}.$$

The $(1-\alpha)$-confidence region for the function $\mathbf{G}_{r,k}\boldsymbol{\beta}$, $\mathbf{H}\boldsymbol{\beta} + \mathbf{h} = \mathbf{0}$, is

$$\mathcal{C}_G = \left\{\mathbf{u} : \mathbf{u} \in R^k, (\mathbf{u} - \widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}})'[\mathrm{Var}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}})]^-(\mathbf{u} - \widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}}) \le \chi_f^2(0; 1-\alpha)\right\},$$

where

$$\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}} = \mathbf{G}\Big(\big[(\mathbf{M}_{H'}\mathbf{X}')^-_{m(\Sigma)}\big]'\mathbf{Y} - \Big\{\mathbf{I} - \big[(\mathbf{M}_{H'}\mathbf{X}')^-_{m(\Sigma)}\big]'\mathbf{X}\Big\}\mathbf{H}'(\mathbf{H}\mathbf{H}')^+\mathbf{h},$$

$$f = r[\mathrm{Var}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}})],$$

$$\mathrm{Var}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}})] = \mathbf{V}_G = \mathbf{G}\Big(\big[(\mathbf{M}_{H'}\mathbf{X}')^-_{m(\Sigma)})\big]'\boldsymbol{\Sigma}(\mathbf{M}_{H'}\mathbf{X}')^-_{m(\Sigma)})\mathbf{G}',$$

$$\boldsymbol{\Sigma} = \sum_{i=1}^p \vartheta_i \mathbf{V}_i.$$

Lemma 3.5 *Let*

$$T(\boldsymbol{\vartheta}) = (\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}} - \mathbf{G}\boldsymbol{\beta})'[\mathrm{Var}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}})]^{-}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}} - \mathbf{G}\boldsymbol{\beta}).$$

Then

$$\begin{aligned}
\frac{\partial T(\boldsymbol{\vartheta})}{\partial \vartheta_i} &= -2\mathbf{v}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H'}\mathbf{X}')^{+}\mathbf{V}_i\mathbf{T}'_G\mathbf{V}_G^{+}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}} - \mathbf{G}\boldsymbol{\beta}) \\
&\quad -(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}} - \mathbf{G}\boldsymbol{\beta})'\mathbf{V}_G^{+}\mathbf{T}_G\mathbf{V}_i\mathbf{T}'_G\mathbf{V}_G^{+}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}} - \mathbf{G}\boldsymbol{\beta}), \\
\mathbf{V}_G &= \mathrm{Var}(\widehat{\widehat{\mathbf{G}\boldsymbol{\beta}}}) = \mathbf{G}\Big\{\big[\mathbf{M}_{H'}\mathbf{X}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H'}\mathbf{X}')^{+}\mathbf{X}\mathbf{M}_{H'}\big]^{+} - \mathbf{M}_{H'}\Big\}\mathbf{G}', \\
\mathbf{T}_G &= \mathbf{G}\big[\mathbf{M}_{H'}\mathbf{X}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H'}\mathbf{X}')^{+}\mathbf{X}\mathbf{M}_{H'}\big]^{+}\mathbf{X}'(\boldsymbol{\Sigma} + \mathbf{X}\mathbf{M}_{H'}\mathbf{X}')^{+}, \\
\mathbf{v} &= \mathbf{Y} - \widehat{\widehat{\mathbf{X}\boldsymbol{\beta}}}, \\
E\left(\frac{\partial T(\boldsymbol{\vartheta})}{\partial \vartheta_i}\right) &= -\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}'_G\mathbf{V}_G^{+}\mathbf{T}_G) = -a_i.
\end{aligned}$$

Further

$$\begin{aligned}
\mathrm{cov}\left(\frac{\partial T(\boldsymbol{\vartheta})}{\partial \vartheta_i}, \frac{\partial T(\boldsymbol{\vartheta})}{\partial \vartheta_j}\right) &= 4\,\mathrm{Tr}\left[\mathbf{V}_i\mathbf{T}'_G\mathbf{V}_G^{+}\mathbf{T}_G\mathbf{V}_j\big(\mathbf{M}_{XM_{H'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{H'}}\big)^{+}\right] \\
&\quad + 2\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}'_G\mathbf{V}_G^{+}\mathbf{T}_G\mathbf{V}_j\mathbf{T}'_G\mathbf{V}_G^{+}\mathbf{T}_G) = \{\mathbf{A}\}_{i,j}.
\end{aligned}$$

Theorem 3.6 *Let* $\mathbf{a} = (a_1, \ldots, a_p)'$ *be the vector given by the preceding lemma and* $\mathbf{A}$ *be the matrix with the* (i, j) *entry equal to* $\{\mathbf{A}\}_{i,j}$ *given also by the preceding lemma. Then the nonsensitivity region for the confidence region* $\mathcal{C}_G$ *is*

$$\begin{aligned}
\mathcal{N}_G &= \Bigg\{\delta\boldsymbol{\vartheta} : \big[\delta\boldsymbol{\vartheta} - \delta_{\max}(t^2\mathbf{A} - \mathbf{a}\mathbf{a}')^{+}\mathbf{a}\big]'\,(t^2\mathbf{A} - \mathbf{a}\mathbf{a}') \\
&\qquad \big[\delta\boldsymbol{\vartheta} - \delta_{\max}(t^2\mathbf{A} - \mathbf{a}\mathbf{a}')^{+}\mathbf{a}\big] \le \delta_{\max}^2 \frac{\mathbf{a}'\mathbf{A}^{+}\mathbf{a}}{t^2 - \mathbf{a}'\mathbf{A}^{+}\mathbf{a}}\Bigg\}, \\
\delta_{\max} &= \chi_f^2(0; 1-\alpha) - \chi_f^2(0; 1-\alpha-\varepsilon)
\end{aligned}$$

and $t > 0$ *is sufficiently large real number. It is valid that*

$$\delta\boldsymbol{\vartheta} \in \mathcal{N}_G \Rightarrow P\{\mathbf{G}\boldsymbol{\beta} \in \mathcal{C}_B\} \ge 1 - \alpha - \varepsilon.$$

Proof Let t be sufficiently large, such that

$$\begin{aligned}
\sum_{i=1}^{p} \frac{\partial T(\boldsymbol{\vartheta})}{\partial \vartheta_i}\delta\vartheta_i &< E\left(\sum_{i=1}^{p} \frac{\partial T(\boldsymbol{\vartheta})}{\partial \vartheta_i}\delta\vartheta_i\right) + t\sqrt{Var\left(\sum_{i=1}^{p} \frac{\partial T(\boldsymbol{\vartheta})}{\partial \vartheta_i}\delta\vartheta_i\right)} \\
&= -\mathbf{a}'\delta\boldsymbol{\vartheta} + t\sqrt{\delta\boldsymbol{\vartheta}'\mathbf{A}\delta\boldsymbol{\vartheta}}.
\end{aligned}$$

If

$$-\mathbf{a}'\delta\boldsymbol{\vartheta} + t\sqrt{\delta\boldsymbol{\vartheta}'\mathbf{A}\delta\boldsymbol{\vartheta}} \le \delta_{\max},$$

where $\delta_{\max} = \chi^2_f(0; 1-\alpha) - \chi^2_f(0; 1-\alpha-\varepsilon)$, then

$$P\{T(\boldsymbol{\vartheta} + \delta\boldsymbol{\vartheta}) \le \chi^2_f(0; 1-\alpha)\} \ge 1 - \alpha - \varepsilon.$$

Thus

$$t^2\delta\boldsymbol{\vartheta}'\mathbf{A}\delta\boldsymbol{\vartheta} \le (\delta_{\max} + \mathbf{a}'\delta\boldsymbol{\vartheta})^2 \Leftrightarrow t^2\delta\boldsymbol{\vartheta}'\mathbf{A}\delta\boldsymbol{\vartheta} - \delta\boldsymbol{\vartheta}'\mathbf{a}\mathbf{a}'\delta\boldsymbol{\vartheta} - 2\delta_{\max}\mathbf{a}'\delta\boldsymbol{\vartheta} \le \delta^2_{\max}.$$

If $\mathbf{a} \in \mathcal{M}(t^2\mathbf{A} - \mathbf{a}\mathbf{a}')$, then the last inequality can be rearranged as

$$\begin{aligned}\left[\delta\boldsymbol{\vartheta} - \delta_{\max}(t^2\mathbf{A} - \mathbf{a}\mathbf{a}')^+\mathbf{a}\right]' (t^2\mathbf{A} - \mathbf{a}\mathbf{a}') \left[\delta\boldsymbol{\vartheta} - \delta_{\max}(t^2\mathbf{A} - \mathbf{a}\mathbf{a}')^+\mathbf{a}\right]& \\ \le \delta^2_{\max}\frac{\mathbf{a}'\mathbf{A}^+\mathbf{a}}{t^2 - \mathbf{a}'\mathbf{A}^+\mathbf{a}}&.\end{aligned}$$

Here the equality

$$\mathbf{a}'(t^2\mathbf{A} - \mathbf{a}\mathbf{a}')^+\mathbf{a} = \frac{\mathbf{a}'\mathbf{A}^+\mathbf{a}}{t^2 - \mathbf{a}'\mathbf{A}^+\mathbf{a}}$$

is used. It is valid that $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2$, where

$$\begin{aligned}\{\mathbf{A}_1\}_{i,j} &= 4\,\mathrm{Tr}\left[\mathbf{V}_i\mathbf{T}'_G\mathbf{V}^+_G\mathbf{T}_G\mathbf{V}_j\left(\mathbf{M}_{XM_{H'}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{H'}}\right)^+\right],\\ \{\mathbf{A}_2\}_{i,j} &= 2\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}'_G\mathbf{V}^+_G\mathbf{T}_G\mathbf{V}_j\mathbf{T}'_G\mathbf{V}^+_G\mathbf{T}_G) = \{\mathbf{A}\}_{i,j}.\end{aligned}$$

Both matrices are p.s.d., i.e. $\mathcal{M}(\mathbf{A}_1 + \mathbf{A}_2) = \mathcal{M}(\mathbf{A}_1, \mathbf{A}_2)$. Since $\mathbf{a} \in \mathcal{M}(\mathbf{A}_2)$ because of the relationship

$$\begin{pmatrix}\mathrm{Tr}(\mathbf{U}\mathbf{V}_1)\\ \mathrm{Tr}(\mathbf{U}\mathbf{V}_2)\\ \vdots\\ \mathrm{Tr}(\mathbf{U}\mathbf{V}_p)\end{pmatrix} \in \mathcal{M}\begin{pmatrix}\mathrm{Tr}(\mathbf{U}\mathbf{V}_1\mathbf{U}\mathbf{V}_1), & \mathrm{Tr}(\mathbf{U}\mathbf{V}_1\mathbf{U}\mathbf{V}_2, & \dots, & \mathrm{Tr}(\mathbf{U}\mathbf{V}_1\mathbf{U}\mathbf{V}_p)\\ \mathrm{Tr}(\mathbf{U}\mathbf{V}_2\mathbf{U}\mathbf{V}_1), & \mathrm{Tr}(\mathbf{U}\mathbf{V}_2\mathbf{U}\mathbf{V}_2), & \dots, & \mathrm{Tr}(\mathbf{U}\mathbf{V}_2\mathbf{U}\mathbf{V}_p)\\ \dots & \dots & \dots & \dots\\ \mathrm{Tr}(\mathbf{U}\mathbf{V}_p\mathbf{U}\mathbf{V}_1), & \mathrm{Tr}(\mathbf{U}\mathbf{V}_p\mathbf{U}\mathbf{V}_2), & \dots, & \mathrm{Tr}(\mathbf{U}\mathbf{V}_p\mathbf{U}\mathbf{V}_p)\end{pmatrix},$$

where $\mathbf{U} = \mathbf{T}'_G\mathbf{V}^+_G\mathbf{T}_G$, what can be easily proved, and the number t can be suitably chosen, the assumption $\mathbf{a} \in \mathcal{M}(t^2\mathbf{A} - \mathbf{a}\mathbf{a}')$ can be ensured. □

A construction of the nonsensitivity region for different multivariate linear statistical models with constraints can be derived from the last theorem. It is sufficient to find the vector $\mathbf{a}$ and the matrix $\mathbf{A}$ for different situations.

Theorem 3.7 *Let in the model (1) with* $\mathbf{H}_2 = \mathbf{I}$ *the matrix* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2 \otimes \mathbf{G}'_1) \subset \mathcal{M}[\mathbf{I} \otimes (\mathbf{X}', \mathbf{H}'_1)]$. *Then*

$$\begin{aligned}\mathbf{a} &= (a_1, \dots, a_p)',\\ a_i &= r[\mathbf{G}_1(\mathbf{M}_{H'_1}\mathbf{X}'\mathbf{X})^+\mathbf{G}'_1]\,\mathrm{Tr}[\mathbf{V}_i\mathbf{G}_2(\mathbf{G}'_2\boldsymbol{\Sigma}\mathbf{G}_2)^+\mathbf{G}'_2], \quad i = 1, \dots, p,\\ \mathbf{A} &= r[\mathbf{G}_1(\mathbf{M}_{H'_1}\mathbf{X}'\mathbf{X})^+\mathbf{G}'_1]\mathbf{S}_{G_2(G'_2\Sigma G_2)^+G'_2},\end{aligned}$$

where

$$\left\{\mathbf{S}_{G_2(G'_2\Sigma G_2)^+G'_2}\right\}_{i,j} = \mathrm{Tr}[\mathbf{V}_i\mathbf{G}_2(\mathbf{G}'_2\boldsymbol{\Sigma}\mathbf{G}_2)^+\mathbf{G}'_2\mathbf{V}_j\mathbf{G}_2(\mathbf{G}'_2\boldsymbol{\Sigma}\mathbf{G}_2)^+\mathbf{G}'_2].$$

Proof With respect to Lemma 3.5 the following scheme

$$\mathbf{V}_i \to \mathbf{V}_i \otimes \mathbf{I}, \quad \mathbf{T}_G \to \mathbf{T}_{G_2' \otimes G_1}, \quad \mathbf{V}_G \to \mathbf{V}_{G_2' \otimes G_1},$$

will be used. Here

$$\begin{aligned} \mathbf{T}_{G_2' \otimes G_1} &= \mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1} = \mathbf{G}_2' \otimes \mathbf{G}_1[(\mathbf{M}_{H_1'}\mathbf{X}')^-_{m(I)}]', \\ \mathbf{T}_{G_2} &= \mathbf{G}_2, \quad \mathbf{T}_{G_1} = \mathbf{G}_1[(\mathbf{M}_{H_1'}\mathbf{X}')^-_{m(I)}]', \\ \mathbf{V}_{G_2' \otimes G_1} &= \mathbf{V}_{G_2} \otimes \mathbf{V}_{G_1}, \\ \mathbf{V}_{G_1} &= \mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'})^+\mathbf{G}_1', \\ \mathbf{V}_{G_2} &= \mathbf{G}_2'\mathbf{\Sigma}\mathbf{G}_2. \end{aligned}$$

Now we use the formulae from Lemma 3.5 and thus we obtain

$$\begin{aligned} a_i &= \mathrm{Tr}[(\mathbf{V}_1 \otimes \mathbf{I})(\mathbf{T}_{G_2} \otimes \mathbf{T}_{G_1}')(\mathbf{V}_{G_2}^+ \otimes \mathbf{V}_{G_1}^+)(\mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1})] \\ &= r[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X})^+\mathbf{M}_{H_1'})^+\mathbf{G}_1']\,\mathrm{Tr}[\mathbf{V}_i\mathbf{G}_2(\mathbf{G}_2'\mathbf{\Sigma}\mathbf{G}_2)^+\mathbf{G}_2], \end{aligned}$$

since

$$\begin{aligned} &\mathrm{Tr}(\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^+\mathbf{T}_{G_1}) = \\ &= \mathrm{Tr}\left\{(\mathbf{M}_{H_1'}\mathbf{X}')^-_{m(I)}\mathbf{G}_1'[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']^+\mathbf{G}_1[(\mathbf{M}_{H_1'}\mathbf{X}')^-_{m(I)}]'\right\} \\ &= \mathrm{Tr}\left\{\mathbf{G}_1[(\mathbf{M}_{H_1'}\mathbf{X}')^-_{m(I)}]'(\mathbf{M}_{H_1'}\mathbf{X}')^-_{m(I)}\mathbf{G}_1'[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']^+\right\} \\ &= \mathrm{Tr}\left\{[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X})^+\mathbf{G}_1'][\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']^+\right\} \\ &= r[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']. \end{aligned}$$

As far as the matrix $\mathbf{A}$ be concerned, it is valid that

$$\begin{aligned} \{\mathbf{A}\}_{i,j} &= 4\,\mathrm{Tr}\Big\{(\mathbf{V}_i \otimes \mathbf{I})(\mathbf{T}_{G_2} \otimes \mathbf{T}_{G_1}')(\mathbf{V}_{G_2}^+ \otimes \mathbf{V}_{G_1}^+)(\mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1})(\mathbf{V}_j \otimes \mathbf{I}) \\ &\quad \times [\mathbf{M}_{I\otimes(XM_{H_1'})}(\mathbf{\Sigma} \otimes \mathbf{I})\mathbf{M}_{I\otimes(XM_{H_1'})}]^+\Big\} \\ &\quad + 2\,\mathrm{Tr}\Big\{(\mathbf{V}_i \otimes \mathbf{I})(\mathbf{T}_{G_2} \otimes \mathbf{T}_{G_1}')(\mathbf{V}_{G_2}^+ \otimes \mathbf{V}_{G_1}^+)(\mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1})(\mathbf{V}_j \otimes \mathbf{I}) \\ &\quad \times (\mathbf{T}_{G_2} \otimes \mathbf{T}_{G_1}')(\mathbf{V}_{G_2}^+ \otimes \mathbf{V}_{G_1}^+)(\mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1})(\mathbf{V}_j \otimes \mathbf{I})\Big\} \\ &= r[\mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X})^+\mathbf{G}_1'] \\ &\quad \times \mathrm{Tr}\left\{\mathbf{V}_i\mathbf{G}_2(\mathbf{G}_2'\mathbf{\Sigma}\mathbf{G}_2)^+\mathbf{G}_2'\mathbf{V}_j\mathbf{G}_2(\mathbf{G}_2'\mathbf{\Sigma}\mathbf{G}_2)^+\mathbf{G}_2'\right\}, \end{aligned}$$

since

$$\begin{aligned} [\mathbf{M}_{I\otimes(XM_{H_1'})}(\mathbf{\Sigma} \otimes \mathbf{I})\mathbf{M}_{I\otimes(XM_{H_1'})}]^+ &= [(\mathbf{I} \otimes \mathbf{M}_{XM_{H_1'}})(\mathbf{\Sigma} \otimes \mathbf{I})(\mathbf{I} \otimes \mathbf{M}_{XM_{H_1'}})]^+ \\ &= \mathbf{\Sigma}^+ \otimes \mathbf{M}_{XM_{H_1'}}. \end{aligned}$$

Now it is easy to finish the proof. □

Theorem 3.8 *Let in the model (1) with* $\mathbf{H}_1 = \mathbf{I}$ *the matrix* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2 \otimes \mathbf{G}_1') \subset \mathcal{M}(\mathbf{I} \otimes \mathbf{X}', \mathbf{H}_2 \otimes \mathbf{I})$. *Then*

$$\begin{aligned}
\mathbf{a} &= (a_1, \dots, a_p)',\\
a_i &= r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'),\\
\mathbf{A} &= 4r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']\mathbf{C}_{T_{G_2}V_{G_2}^+T_{G_2}',(P_{H_2}\Sigma P_{H_2})^+} + 2r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']\mathbf{S}_{T_{G_2}V_{G_2}^+T_{G_2}'},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{T}_{G_2'\otimes G_1} &= \mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1}, \quad \mathbf{T}_{G_1} = \mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{X}', \quad \mathbf{T}_{G_2} = (\mathbf{M}_{H_2})_{m(\Sigma)}^-\mathbf{G}_2,\\
\mathbf{V}_{G_2'\otimes G_1} &= \mathbf{V}_{G_2} \otimes \mathbf{V}_{G_1},\\
\mathbf{V}_{G_1} &= \mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1', \quad \mathbf{V}_{G_2} = \mathbf{G}_2'\Big\{[\mathbf{M}_{H_2}(\mathbf{\Sigma} + \mathbf{M}_{H_2})^+\mathbf{M}_{H_2}]^+ - \mathbf{M}_{H_2}\Big\}
\end{aligned}$$

and

$$\begin{aligned}
\Big\{\mathbf{C}_{T_{G_2}V_{G_2}^+T_{G_2}',(P_{H_2}\Sigma P_{H_2})^+}\Big\}_{i,j} &= \mathrm{Tr}[\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'\mathbf{V}_j(\mathbf{P}_{H_2}\mathbf{\Sigma}\mathbf{P}_{H_2})^+],\\
\Big\{\mathbf{S}_{T_{G_2}V_{G_2}^+T_{G_2}'}\Big\}_{i,j} &= \mathrm{Tr}[\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'\mathbf{V}_j\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'].
\end{aligned}$$

Proof It is analogous as in preceding theorem. The formulae from Lemma 3.5 must be used. The equality

$$\big[\mathbf{M}_{M_{H_2}\otimes X}(\mathbf{\Sigma} \otimes \mathbf{I})\mathbf{M}_{M_{H_2}\otimes X}\big] = \mathbf{\Sigma}^+ \otimes \mathbf{M}_X + (\mathbf{P}_{H_2}\mathbf{\Sigma}\mathbf{P}_{H_2})^+ \otimes \mathbf{P}_X$$

must be taken into account. □

Theorem 3.9 *Let in the model (2) with* $\mathbf{H}_2 = \mathbf{I}$ *the matrix* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2' \otimes \mathbf{G}_1) \subset \mathcal{M}[\mathbf{I} \otimes (\mathbf{X}', \mathbf{H}_1')]$. *Then*

$$\begin{aligned}
\mathbf{a} &= (a_1, \dots, a_p)',\\
a_i &= r(\mathbf{G}_2)\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^+\mathbf{T}_{G_1}), \quad i = 1, \dots, p,\\
\mathbf{A} &= 4r(\mathbf{G}_2)\mathbf{C}_{T_{G_1}'V_{G_1}^+T_{G_1},(M_{XM_{H_1'}}\Sigma M_{XM_{H_1'}})^+} + 2r(\mathbf{G}_2)\mathbf{S}_{T_{G_1}'V_{G_1}^+T_{G_1}},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{T}_{G_2'\otimes G_1} &= \mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1},\\
\mathbf{T}_{G_1} &= \mathbf{G}_1[\mathbf{M}_{H_1'}\mathbf{X}')_{m(\Sigma)}^-]'\\
&= \mathbf{G}_1[\mathbf{M}_{H_1'}\mathbf{X}'(\mathbf{\Sigma} + \mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^+\mathbf{X}\mathbf{M}_{H_1'}]^+\mathbf{M}_{H_1'}\mathbf{X}'(\mathbf{\Sigma} + \mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^+,\\
\mathbf{T}_{G_2} &= \mathbf{G}_2,\\
\mathbf{V}_{G_1} &= \mathbf{G}_1\big\{[\mathbf{M}_{H_1'}\mathbf{X}'(\mathbf{\Sigma} + \mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^+\mathbf{X}\mathbf{M}_{H_1'}]^+ - \mathbf{M}_{H_1'}\big\}\mathbf{G}_1',\\
\mathbf{V}_{G_2} &= \mathbf{G}_2'\mathbf{G}_2
\end{aligned}$$

and

$$\left\{\mathbf{C}_{T'_{G_1}V^+_{G_1}T_{G_1},(M_{XM_{H'_1}}\Sigma M_{XM_{H'_1}})^+}\right\}_{i,j} =$$
$$= \mathrm{Tr}[\mathbf{V}_i\mathbf{T}'_{G_1}\mathbf{V}^+_{G_1}\mathbf{T}_{G_1}\mathbf{V}_j(\mathbf{M}_{XM_{H'_1}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{H'_1}})^+],$$
$$\left\{\mathbf{S}_{T'_{G_1}V^+_{G_1}T_{G_1}}\right\}_{i,j} = \mathrm{Tr}(\mathbf{V}_i\mathbf{T}'_{G_1}\mathbf{V}^+_{G_1}\mathbf{T}_{G_1}\mathbf{V}_j\mathbf{T}'_{G_1}\mathbf{V}^+_{G_1}\mathbf{T}_{G_1}).$$

Proof The obvious equality

$$\left[\mathbf{M}_{I\otimes(XM_{H'_1})}(\mathbf{I}\otimes\boldsymbol{\Sigma})\mathbf{M}_{I\otimes(XM_{H'_1})}\right] = \mathbf{I}\otimes\left(\mathbf{M}_{XM_{H'_1}}\boldsymbol{\Sigma}\mathbf{M}_{XM_{H'_1}}\right)^+$$

and Lemma 3.5 must be used. □

Theorem 3.10 *Let in the model (2) with* $\mathbf{H}_1 = \mathbf{I}$ *the matrix* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2\otimes\mathbf{G}'_1)\subset\mathcal{M}(\mathbf{I}\otimes\mathbf{X}',\mathbf{H}_2\otimes\mathbf{I})$*. Then*

$$\begin{aligned}
\mathbf{a} &= (a_1,\dots,a_p)',\\
a_i &= r(\mathbf{M}_{H_2}\mathbf{G}_2)\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}'_{G_1}\mathbf{V}^+_{G_1}\mathbf{T}_{G_1}),\quad i=1,\dots,p,\\
\mathbf{A} &= 4r(\mathbf{M}_{H_2}\mathbf{G}_2)\mathbf{C}_{T'_{G_1}V^+_{G_1}T_{G_1},(M_X\Sigma M_X)^+} + 2r(\mathbf{M}_{H_2}\mathbf{G}_2)\mathbf{S}_{T'_{G_1}V^+_{G_1}T_{G_1}},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{T}_{G'_2\otimes G_1} &= \mathbf{T}'_{G_2}\otimes\mathbf{T}_{G_1},\quad \mathbf{T}_{G_1} = \mathbf{G}_1\left[(\mathbf{X}')^-_{M(\Sigma)}\right]' = \mathbf{G}_1\mathbf{X}'\mathbf{T}^+\mathbf{X})^+\mathbf{X}'\mathbf{T}^+,\\
\mathbf{T} &= \boldsymbol{\Sigma}+\mathbf{X}\mathbf{X},\quad \mathbf{T}_{G_2} = \mathbf{M}_{H_2}\mathbf{G}_2,\\
\mathbf{V}_{G_1} &= \mathbf{G}_1[\mathbf{X}'\mathbf{T}^+\mathbf{X})^+ - \mathbf{I}]\mathbf{G}'_1,\quad \mathbf{V}_{G_2} = \mathbf{G}'_2\mathbf{M}_{H_2}\mathbf{G}_2
\end{aligned}$$

and

$$\begin{aligned}
\left\{\mathbf{C}_{T'_{G_1}V^+_{G_1}T_{G_1},(M_X\Sigma M_X)^+}\right\}_{i,j} &= \mathrm{Tr}[\mathbf{V}_i\mathbf{T}'_{G_1}\mathbf{V}^+_{G_1}\mathbf{T}_{G_1}\mathbf{V}_j(\mathbf{M}_X\boldsymbol{\Sigma}\mathbf{M}_X)^+],\\
\left\{\mathbf{S}_{T'_{G_1}V^+_{G_1}T_{G_1}}\right\}_{i,j} &= \mathrm{Tr}(\mathbf{V}_i\mathbf{T}'_{G_1}\mathbf{V}^+_{G_1}\mathbf{T}_{G_1}).
\end{aligned}$$

Proof The equalities

$$\left[\mathbf{M}_{M_{H_2}\otimes X}(\mathbf{I}\otimes\boldsymbol{\Sigma})\mathbf{M}_{M_{H_2}\otimes X}\right] = \mathbf{M}_{H_2}\otimes(\mathbf{M}_X\boldsymbol{\Sigma}\mathbf{M}_X)^+ + \mathbf{P}_{H_2}\otimes\boldsymbol{\Sigma}^+,$$
$$\mathbf{T}'_{G_2}\mathbf{P}_{H_2} = \mathbf{0}$$

and Lemma 3.5 must be taken into account. □

Theorem 3.11 *Let in the model (3) with* $\mathbf{H}_2 = \mathbf{I}$ *the matrix* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2\otimes\mathbf{G}'_1)\subset\mathcal{M}(\mathbf{Z}\otimes\mathbf{X}',\mathbf{I}\otimes\mathbf{H}'_1)$*. Then*

$$\begin{aligned}
\mathbf{a} &= (a_1,\dots,a_p)',\\
a_i &= r[\mathbf{V}^+_{G_1}\mathbf{G}_1(\mathbf{M}_{H'_1}\mathbf{X}'\mathbf{X}\mathbf{M}_{H'_1})^+\mathbf{G}'_1]\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}^+_{G_2}\mathbf{T}'_{G_2}),\quad i=1,\dots,p,\\
\mathbf{A} &= 4r[\mathbf{V}^+_{G_1}\mathbf{G}_1(\mathbf{M}_{H'_1}\mathbf{X}'\mathbf{X}\mathbf{M}_{H'_1})^+\mathbf{G}'_1]\mathbf{C}_{T_{G_2}V^+_{G_2}T'_{G_2},(M_{Z'}\Sigma M_{Z'})^+}\\
&\quad + 2r[\mathbf{V}^+_{G_1}\mathbf{G}_1(\mathbf{M}_{H'_1}\mathbf{X}'\mathbf{X}\mathbf{M}_{H'_1})^+\mathbf{G}'_1]\mathbf{S}_{T_{G_2}V^+_{G_2}T'_{G_2}},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{T}_{G_2'\otimes G_1} &= \mathbf{T}_{G_2}'\otimes\mathbf{T}_{G_1}, \quad \mathbf{T}_{G_1}=\mathbf{G}_1\left(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'}\right)^+\mathbf{M}_{H_1'}\mathbf{X}',\\
\mathbf{T}_{G_2} &= =\mathbf{Z}_{m(\Sigma)}^-\mathbf{G}_2=\mathbf{U}^+\mathbf{Z}'(\mathbf{Z}\mathbf{U}^+\mathbf{Z}')^+, \quad \mathbf{U}=\mathbf{\Sigma}+\mathbf{Z}'\mathbf{Z},\\
\mathbf{V}_{G_1} &= \mathbf{G}_1(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'})^+\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{\Sigma}\mathbf{X}\mathbf{M}_{H_1'}(\mathbf{M}_{H_1'}\mathbf{X}'\mathbf{X}\mathbf{M}_{H_1'})^+\mathbf{G}_1',\\
\mathbf{V}_{G_2} &= \mathbf{G}_2'[(\mathbf{Z}\mathbf{U}^+\mathbf{Z}')^+-\mathbf{I}]\mathbf{G}_2
\end{aligned}$$

and

$$\begin{aligned}
\left\{\mathbf{C}_{T_{G_2}V_{G_2}^+T_{G_2}',(M_{Z'}\Sigma M_{Z'})^+}\right\}_{i,j} &= \mathrm{Tr}[\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'\mathbf{V}_j(\mathbf{M}_{Z'}\mathbf{\Sigma}\mathbf{M}_{Z'})^+],\\
\left\{\mathbf{S}_{T_{G_2}V_{G_2}^+T_{G_2}'}\right\}_{i,j} &= \mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'\mathbf{V}_j\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}').
\end{aligned}$$

Proof Lemma 3.5 and the equalities

$$\left[\mathbf{M}_{Z'\otimes(XM_{H_1'})}(\mathbf{\Sigma}\otimes\mathbf{I})\mathbf{M}_{Z'\otimes(XM_{H_1'})}\right]=\mathbf{\Sigma}^+\otimes\mathbf{M}_{XM_{H_1'}}+(\mathbf{M}_{Z'}\mathbf{\Sigma}\mathbf{M}_{Z'})^+\otimes\mathbf{P}_{XM_{H_1'}},$$

$$\mathbf{T}_{G_1}\mathbf{M}_{XM_{H_1'}}=\mathbf{0}$$

must be used. □

Theorem 3.12 *Let in the model (3) with $\mathbf{H}_1=\mathbf{I}$ the matrix $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ be unbiasedly estimable, i.e. $\mathcal{M}(\mathbf{G}_2\otimes\mathbf{G}_1')\subset\mathcal{M}(\mathbf{Z}\otimes\mathbf{X}',\mathbf{H}_2\otimes\mathbf{I})$. Then*

$$\begin{aligned}
\mathbf{a} &= (a_1,\dots,a_p)',\\
a_i &= r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'), \quad i=1,\dots,p,\\
\mathbf{A} &= 4r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']\mathbf{C}_{T_{G_2}V_{G_2}^+T_{G_2}',(M_{Z'M_{H_2}}\Sigma M_{Z'M_{H_2}})^+}\\
&\quad +2r[\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1']\mathbf{S}_{T_{G_2}V_{G_2}^+T_{G_2}'},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{T}_{G_2'\otimes G_1} &= \mathbf{T}_{G_2}'\otimes\mathbf{T}_{G_1}, \quad \mathbf{T}_{G_1}=\mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{X}',\\
\mathbf{T}_{G_2} &= (\mathbf{M}_{H_2}\mathbf{Z})_{m(\Sigma)}^-\mathbf{G}_2\\
&= (\mathbf{\Sigma}+\mathbf{Z}'\mathbf{M}_{H_2}\mathbf{Z}')^+\mathbf{Z}'\mathbf{M}_{H_2}[\mathbf{M}_{H_2}\mathbf{Z}(\mathbf{\Sigma}+\mathbf{Z}'\mathbf{M}_{H_2}\mathbf{Z})^+\mathbf{Z}'\mathbf{M}_{H_2}]^+\mathbf{G}_2,\\
\mathbf{V}_{G_1} &= \mathbf{G}_1(\mathbf{X}'\mathbf{X})^+\mathbf{G}_1',\\
\mathbf{V}_{G_2} &= \mathbf{G}_2'\{[\mathbf{M}_{H_2}\mathbf{Z}'(\mathbf{\Sigma}+\mathbf{Z}'\mathbf{M}_{H_2}\mathbf{Z})^+\mathbf{Z}'\mathbf{M}_{H_2}]^+-\mathbf{M}_{H_2}\}\mathbf{G}_2
\end{aligned}$$

and

$$\begin{aligned}
&\left\{\mathbf{C}_{T_{G_2}V_{G_2}^+T_{G_2}',(M_{Z'M_{H_2}}\Sigma M_{Z'M_{H_2}})^+}\right\}_{i,j}=\\
&= \mathrm{Tr}[\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'\mathbf{V}_j(\mathbf{M}_{Z'M_{H_2}}\mathbf{\Sigma}\mathbf{M}_{Z'M_{H_2}})^+],\\
\left\{\mathbf{S}_{T_{G_2}V_{G_2}^+T_{G_2}'}\right\} &= \mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}'\mathbf{V}_j\mathbf{T}_{G_2}\mathbf{V}_{G_2}^+\mathbf{T}_{G_2}').
\end{aligned}$$

Proof The equalities

$$\left[\mathbf{M}_{(Z'M_{H_2})\otimes X}(\mathbf{\Sigma}\otimes\mathbf{I})\mathbf{M}_{(Z'M_{H_2})\otimes X}\right]=\mathbf{\Sigma}^{+}\otimes\mathbf{M}_X+\left(\mathbf{M}_{Z'M_{H_2}}\mathbf{\Sigma}\mathbf{M}_{Z'M_{H_2}}\right)^{+}\otimes\mathbf{P}_X,$$

$$\mathbf{T}_{G_1}\mathbf{M}_X=\mathbf{0}$$

and Lemma 3.5 must be taken into account. □

Theorem 3.13 *Let in the model (4) with $\mathbf{H}_2=\mathbf{I}$ the matrix $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ be unbiasedly estimable, i.e. $\mathcal{M}(\mathbf{G}_2\otimes\mathbf{G}_1')\subset\mathcal{M}(\mathbf{Z}\otimes\mathbf{X}',\mathbf{I}\otimes\mathbf{H}_1')$. Then*

$$\begin{aligned}\mathbf{a}&=(a_1,\dots,a_p)',\\ a_i&=r[\mathbf{G}_2'(\mathbf{Z}\mathbf{Z}')^{+}\mathbf{G}_2]\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^{+}\mathbf{T}_{G_1}),\quad i=1,\dots,p,\\ \mathbf{A}&=4r[\mathbf{G}_2'(\mathbf{Z}\mathbf{Z}')^{+}\mathbf{G}_2]\mathbf{C}_{T_{G_1}'V_{G_1}^{+}T_{G_1}',(M_{XM_{H_1'}}\Sigma M_{XM_{H_1'}})^{+}}\\ &\quad+2r[\mathbf{G}_2'(\mathbf{Z}\mathbf{Z}')^{+}\mathbf{G}_2]\mathbf{S}_{T_{G_1}'V_{G_1}^{+}T_{G_1}'},\end{aligned}$$

where

$$\begin{aligned}\mathbf{T}_{G_2'\otimes G_1}&=\mathbf{T}_{G_2}'\otimes\mathbf{T}_{G_1},\\ \mathbf{T}_{G_1}&=\mathbf{G}_1\left[(\mathbf{M}_{H_1'}\mathbf{X}')^{-}_{m(\Sigma)}\right]'\\ &=\mathbf{G}_1[\mathbf{M}_{H_1'}\mathbf{X}'(\mathbf{\Sigma}+\mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^{+}\mathbf{X}\mathbf{M}_{H_1'}^{+}\mathbf{M}_{H_1'}\mathbf{X}'(\mathbf{\Sigma}+\mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^{+},\\ \mathbf{T}_{G_2}&=\mathbf{Z}'(\mathbf{Z}\mathbf{Z})^{+}\mathbf{G}_2,\\ \mathbf{V}_{G_1}&=\mathbf{G}_1\{[\mathbf{M}_{H_1'}\mathbf{X}'(\mathbf{\Sigma}+\mathbf{X}\mathbf{M}_{H_1'}\mathbf{X}')^{+}\mathbf{X}\mathbf{M}_{H_1'}]^{+}-\mathbf{M}_{H_1'}\}\mathbf{G}_1',\\ \mathbf{V}_{G_2}&=\mathbf{G}_2'(\mathbf{Z}\mathbf{Z}')^{+}\mathbf{G}_2\end{aligned}$$

and

$$\begin{aligned}&\left\{\mathbf{C}_{T_{G_1}'V_{G_1}^{+}T_{G_1}',(M_{XM_{H_1'}}\Sigma M_{XM_{H_1'}})^{+}}\right\}_{i,j}=\\ &=\mathrm{Tr}[\mathbf{V}_i\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^{+}\mathbf{T}_{G_1}2\mathbf{V}_j(\mathbf{M}_{XM_{H_1'}}\mathbf{\Sigma}\mathbf{M}_{XM_{H_1'}})^{+}],\\ \left\{\mathbf{S}_{T_{G_1}'V_{G_1}^{+}T_{G_1}}\right\}&=\mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^{+}\mathbf{T}_{G_1}\mathbf{V}_j\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^{+}\mathbf{T}_{G_1}).\end{aligned}$$

Proof The equalities

$$\left[\mathbf{M}_{Z'\otimes(XM_{H_1'}}(\mathbf{I}\otimes\mathbf{\Sigma})\mathbf{M}_{Z'\otimes(XM_{H_1'}}\right]^{+}=\mathbf{M}_{Z'}\otimes\mathbf{\Sigma}^{+}+\mathbf{P}_{Z'}\otimes(\mathbf{M}_{XM_{H_1'}}\mathbf{\Sigma}\mathbf{M}_{XM_{H_1'}})^{+},$$

$$\mathbf{T}_{G_2}'\mathbf{M}_{Z'}=\mathbf{0}$$

and Lemma 3.5 must be utilized. □

Theorem 3.14 *Let in the model (4) with* $\mathbf{H}_1 = \mathbf{I}$ *the matrix* $\mathbf{G}_1\mathbf{B}\mathbf{G}_2$ *be unbiasedly estimable, i.e.* $\mathcal{M}(\mathbf{G}_2 \otimes \mathbf{G}_1') \subset \mathcal{M}(\mathbf{Z} \otimes \mathbf{X}', \mathbf{H}_2 \otimes \mathbf{I})$. *Then*

$$\begin{aligned}
\mathbf{a} &= (a_1, \dots, a_p)', \\
a_i &= r[\mathbf{G}_2'(\mathbf{M}_{H_2}\mathbf{Z}\mathbf{Z}'\mathbf{M}_{H_2})^+\mathbf{G}_2]\,\mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^+\mathbf{T}_{G_1}'), \quad i = 1, \dots, p, \\
\mathbf{A} &= 4r[\mathbf{G}_2'(\mathbf{M}_{H_2}\mathbf{Z}\mathbf{Z}'\mathbf{M}_{H_2})^+\mathbf{G}_2]\mathbf{C}_{T_{G_1}'V_{G_1}^+T_{G_1},(M_X\Sigma M_X)^+} \\
&\quad + 2r[\mathbf{G}_2'(\mathbf{M}_{H_2}\mathbf{Z}\mathbf{Z}'\mathbf{M}_{H_2})^+\mathbf{G}_2]\mathbf{S}_{T_{G_1}'V_{G_1}^+T_{G_1}},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{T}_{G_2'\otimes G_1} &= \mathbf{T}_{G_2}' \otimes \mathbf{T}_{G_1}, \\
\mathbf{T}_{G_1} &= \mathbf{G}_1[(\mathbf{X}')^-_{m(\Sigma)}]' = \mathbf{G}_1(\mathbf{X}'\mathbf{T}^+\mathbf{X})^+\mathbf{X}'\mathbf{T}^+, \quad \mathbf{T} = \mathbf{\Sigma} + \mathbf{X}\mathbf{X}', \\
\mathbf{T}_{G_2} &= \mathbf{Z}'\mathbf{M}_{H_2}(\mathbf{M}_{H_2}\mathbf{Z}\mathbf{Z}'\mathbf{M}_{H_2})^+\mathbf{G}_2, \\
\mathbf{V}_{G_1} &= \mathbf{G}_1[(\mathbf{X}'\mathbf{T}^+\mathbf{X})^+ - \mathbf{I}]\mathbf{G}_1', \\
\mathbf{V}_{G_2} &= \mathbf{G}_2'(\mathbf{M}_{H_2}Z\mathbf{Z}\mathbf{Z}'\mathbf{M}_{H_2})^+\mathbf{G}_2
\end{aligned}$$

and

$$\begin{aligned}
\left\{\mathbf{C}_{T_{G_1}'V_{G_1}^+T_{G_1},(M_X\Sigma M_X)^+}\right\}_{i,j} &= \mathrm{Tr}[\mathbf{V}_i\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^+\mathbf{T}_{G_1}\mathbf{V}_j(\mathbf{M}_X\mathbf{\Sigma}\mathbf{M}_X)^+], \\
\left\{\mathbf{S}_{T_{G_1}'V_{G_1}^+T_{G_1}}\right\}_{i,j} &= \mathrm{Tr}(\mathbf{V}_i\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^+\mathbf{T}_{G_1}\mathbf{V}_j\mathbf{T}_{G_1}'\mathbf{V}_{G_1}^+\mathbf{T}_{G_1}.
\end{aligned}$$

Proof The equalities

$$\left[\mathbf{M}_{(Z'M_{H_2})\otimes X}(\mathbf{I}\otimes\mathbf{\Sigma})\mathbf{M}_{(Z'M_{H_2})\otimes X}\right]^+ = \mathbf{M}_{Z'M_{H_2}}\otimes\mathbf{\Sigma}^+ + \mathbf{P}_{Z'M_{H_2}}\otimes(\mathbf{M}_X\mathbf{\Sigma}\mathbf{M}_X)^+,$$

$$\mathbf{T}_{G_2}'\mathbf{M}_{Z'M_{H_2}} = \mathbf{0}$$

and Lemma 3.5 must be used. □

References

[1] Anderson, T. W.: Introduction to Multivariate Statistical Analysis. *J. Wiley, New York*, 1958.

[2] Fišerová, E., Kubáček, L., Kunderová, P.: Linear Statistical Models: Regularity and Singularities. *Academia, Praha*, 2007.

[3] Kshirsagar, A. M.: Multivariate Analysis. *M. Dekker, New York*, 1972.

[4] Kubáček, L., Kubáčková, L., Volaufová, J.: Statistical Models with Linear Structures. *Veda (Publishing House of Slovak Academy of Sciences), Bratislava*, 1995.

[5] Rao, C. R.: Linear Statistical Inference and Its Applications. *J. Wiley, New York–London–Sydney*, 1965.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 101–113

Linearization Regions for Confidence Ellipsoids*

Lubomír KUBÁČEK [1], Eva TESAŘÍKOVÁ [2]

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

[2] *Department of Algebra and Geometry, Faculty of Science, Palacký University*
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: tesariko@inf.upol.cz

(Received February 12, 2008)

Abstract

If an observation vector in a nonlinear regression model is normally distributed, then an algorithm for a determination of the exact $(1-\alpha)$-confidence region for the parameter of the mean value of the observation vector is well known. However its numerical realization is tedious and therefore it is of some interest to find some condition which enables us to construct this region in a simpler way.

Key words: Confidence ellipsoid; nonlinear regression model; linearization region.

2000 Mathematics Subject Classification: 62F10, 62J05

1 Introduction

In a linear statistical model with normally distributed observation vector the construction of the confidence regions is a simple problem. If the statistical model is nonlinear, i.e. the mean value of the observation vector is a nonlinear vector function of the parameters, then the problem can be also solved, however

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

it is much more complicated. Therefore it is reasonable to find another, much more simpler procedure, which can be used at least under some conditions.

The aim of the paper is to find such conditions which enables us to use the procedure from the theory of linear statistical models.

More on the problem of linearization of regression models cf. [3], [4], [5], [6], [10], [11].

2 Notation and auxiliary statement

Let $\mathbf{Y}$ be an n-dimensional random vector (observation vector) which is normally distributed. Its mean value is equal to $\mathbf{f}(\boldsymbol{\beta})$, where $\boldsymbol{\beta} \in R^k$ (k-dimensional real linear space) is an unknown vector parameter and $\mathbf{f}(\cdot)\colon R^k \to R^n$ is a vector function. It is assumed that it can be expressed with sufficiently high accuracy as

$$\mathbf{f}(\mathbf{u}) = \mathbf{f}_0 + \mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0) + \frac{1}{2}\boldsymbol{\kappa}(\mathbf{u} - \boldsymbol{\beta}_0), \quad \mathbf{u} \in R^k,$$

where

$$\mathbf{f}_0 = \mathbf{f}(\boldsymbol{\beta}_0), \quad \mathbf{F} = \frac{\partial \mathbf{f}(\mathbf{u})}{\partial \mathbf{u}'}\Big|_{u=\boldsymbol{\beta}_0},$$

$$\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = [\kappa_1(\delta\boldsymbol{\beta}), \dots, \kappa_n(\delta\boldsymbol{\beta})]', \quad \delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$$

$$\kappa_i(\delta\boldsymbol{\beta}) = \delta\boldsymbol{\beta}' \frac{\partial^2 f_i(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}\Big|_{u=\boldsymbol{\beta}_0} \delta\boldsymbol{\beta}, \quad i = 1, \dots, n.$$

The covariance matrix of the vector $\mathbf{Y}$ is $\sigma^2\mathbf{V}$, where $\sigma^2 \in (0, \infty)$ is either known or unknown parameter and the $n \times n$ matrix $\mathbf{V}$ is given.

The notation

$$\mathbf{Y} \sim N_n[\mathbf{f}(\boldsymbol{\beta}), \sigma^2\mathbf{V}], \quad \boldsymbol{\beta} \in R^k, \tag{1}$$

will be used in the following text.

The quadratized version of the model (1) is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\left[\mathbf{F}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}), \sigma^2\mathbf{V}\right], \quad \boldsymbol{\beta} \in R^k, \tag{2}$$

and the linearized version is

$$\mathbf{Y} - \mathbf{f}_0 \sim N_n\left(\mathbf{F}\delta\boldsymbol{\beta}, \sigma^2\mathbf{V}\right), \quad \boldsymbol{\beta} \in R^k. \tag{3}$$

Assumption The regularity of the model (3) is assumed in the following text, i.e. the rank of the matrix $\mathbf{F}$ is $r(\mathbf{F}) = k < n$ and the matrix $\mathbf{V}$ is positive definite.

Lemma 2.1 *The $(1-\alpha)$-confidence region for the vector $\boldsymbol{\beta}$ in the model (3) is*

$$\mathcal{E} = \left\{\mathbf{u}\colon (\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}}) \leq \sigma^2\chi^2_k(0; 1-\alpha)\right\}, \tag{4}$$

if the parameter σ^2 *is known.*

If it is estimated, then

$$\mathcal{E} = \left\{\mathbf{u}\colon (\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}}) \leq \widehat{\sigma^2} F_{k,n-k}(0; 1-\alpha)\right\}. \quad (5)$$

Here

$$\widehat{\delta\boldsymbol{\beta}} = (\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0),$$
$$\widehat{\sigma^2} = (\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})'\,\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})/(n-k),$$

$\chi^2_k(0; 1-\alpha)$ *is* $(1-\alpha)$*-quantile of the central chi-squared distribution with* k *degrees of freedom and* $F_{k,n-k}(0; 1-\alpha)$ *is* $(1-\alpha)$*-quantile of the central Fisher–Snedecor distribution with* k *and* $n-k$ *degrees of freedom.*

Proof Proof is well known (cf. e.g. [2]) and therefore it is omitted. □

Lemma 2.1 is not valid in the model (2). However if $\delta\boldsymbol{\beta} = \boldsymbol{\beta}^* - \boldsymbol{\beta}_0$ is sufficiently small, where $\boldsymbol{\beta}^*$ is the actual value of the vector parameter $\boldsymbol{\beta}$, it can be expected that the region $\mathcal{E}$ from (4) and (5), respectively, covers the actual value $\boldsymbol{\beta}^*$ with a probability larger than $1-\alpha-\varepsilon$, where $\varepsilon > 0$ is a sufficiently small real number.

3 Linearization region

Consider the quadratized model (2) with the given covariance matrix $\boldsymbol{\Sigma} = \sigma^2\mathbf{V}$ (i.e. σ^2 is known).

Definition 3.1 The Bates and Watts [1] parametric curvature $K^{(par)}$ and the intrinsic curvature $K^{(int)}$ of the model (1) at the point $\boldsymbol{\beta}_0$ are given as

$$K^{(par)} = \sigma \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta}}\colon \delta\boldsymbol{\beta} \in R^k\right\} = \sigma K_0^{(par)}$$

and

$$K^{(int)} = \sigma \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta}}\colon \delta\boldsymbol{\beta} \in R^k\right\} = \sigma K_0^{(int)}.$$

Here $\mathbf{P}_F^{V^{-1}} = \mathbf{F}(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}$ and $\mathbf{M}_F^{V^{-1}} = \mathbf{I} - \mathbf{P}_F^{V^{-1}}$.

Let an r-dimensional vector function $\mathbf{d}\colon R^k \to R^r$

$$\mathbf{d}(\boldsymbol{\beta}) = \mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta}, \quad \boldsymbol{\beta} \in R^k,$$

where $r(\mathbf{D}_{r,k}) = r \leq k$, be under consideration.

Theorem 3.1 *Let α and ε be sufficiently small positive real numbers and let $\delta_{\max}$ be solution of the equation*

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

If

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{\mathcal{E}} = \left\{\delta\boldsymbol{\beta} \colon \delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \le \frac{2\sqrt{\delta_{\max}}}{K^{(par)}(\boldsymbol{\beta}_0)}\right\}, \quad \textit{where} \;\; \mathbf{C} = \frac{\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}}{\sigma^2},$$

then

$$\mathcal{E}^* = \left\{\mathbf{u} \colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}) \le \left(\sqrt{\chi_r^2(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2\right\}$$

covers $\mathbf{D}\delta\boldsymbol{\beta}$ with probability at least $(1 - \alpha - \varepsilon)$.

Proof Let

$$\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \le \frac{2\sqrt{\delta_{\max}}}{K^{(par)}(\boldsymbol{\beta}_0)}.$$

Then

$$\forall\{\mathbf{u} \in R^r\} |\mathbf{u}'\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}]| \le \sqrt{\delta_{\max}}\sqrt{\mathbf{u}'\mathbf{D}\mathbf{C}^{-1}\mathbf{D}'\mathbf{u}},$$

what is equivalent, with respect to the Scheffé theorem [9], to

$$[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}]'\mathbf{D}'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}] \le \delta_{\max}.$$

Let

$$\left\{\delta\boldsymbol{\beta} \colon [E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}]'\mathbf{D}(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}] \le \delta_{\max}\right\}. \tag{6}$$

Let $(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1} = \sum_{i=1}^r \lambda_i \mathbf{f}_i\mathbf{f}_i'$ be the spectral decomposition.

The $(1-\alpha)$-confidence ellipsoid in the linearized model is

$$\left\{\mathbf{u} \colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}) \le \chi_r^2(0; 1-\alpha)\right\}$$

and its semiaxes are $a_i = \sqrt{\chi_r^2(0; 1-\alpha)}/\sqrt{\lambda_i}$, $i = 1, \dots, r$. The semiaxes of the ellipsoid (6) are $\pi_i = \sqrt{\delta_{\max}}/\sqrt{\lambda_i}, i = 1, \dots, r$.

The semiaxes of the ellipsoid $\mathcal{E}^*$ are $a_i + \pi_i$, $i = 1, \dots, r$ and it covers all $(1-\alpha)$-ellipsoids in the linearized model with centers

$$\mathbf{D}\widehat{\delta\boldsymbol{\beta}} + \mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}], E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta} \in \mathcal{E}^*.$$

The random variable

$$(\mathbf{D}\delta\boldsymbol{\beta} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{D}\delta\boldsymbol{\beta} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})$$

is chi-squared with r degrees of freedom and with the parameter of noncentrality equal to

$$\delta = \left\{ \mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}] \right\}' (\mathbf{DC}^{-1}\mathbf{D}')^{-1}\mathbf{D}[E(\widehat{\delta\boldsymbol{\beta}}) - \delta\boldsymbol{\beta}] < \delta_{\max}.$$

If $\delta_{\max}$ satisfies the equality

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

then the ellipsoid $\mathcal{E}^*$ covers the vector $\mathbf{D}\delta\boldsymbol{\beta}$ with probability larger or equal to $1 - \alpha - \varepsilon$. □

Corollary 3.1 *Let* $\mathbf{d}(\boldsymbol{\beta}) = \boldsymbol{\beta}$. *If*

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{\mathcal{E}} = \left\{ \delta\boldsymbol{\beta} \colon \delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \le \frac{2\sqrt{\delta_{\max}}}{K^{(par)}(\boldsymbol{\beta}_0)} \right\},$$

then

$$\mathcal{E}^* = \left\{ \mathbf{u} \colon (\mathbf{u} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{C}(\mathbf{u} - \widehat{\delta\boldsymbol{\beta}}) \le \left(\sqrt{\chi_r^2(0; 1-\alpha)} + \sqrt{\delta_{\max}} \right)^2 \right\}$$

covers $\delta\boldsymbol{\beta}$ *with probability at least* $(1 - \alpha - \varepsilon)$.

Let the function $\mathbf{d}(\boldsymbol{\beta})$ be of the quadratic form, i.e.

$$\mathbf{d}(\boldsymbol{\beta}) = \mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\delta}(\delta\boldsymbol{\beta}),$$

where $\boldsymbol{\delta}(\delta\boldsymbol{\beta}) = [\delta_1(\delta\boldsymbol{\beta}), \ldots, \delta_r(\delta\boldsymbol{\beta})]'$, $\delta_i(\delta\boldsymbol{\beta}) = \delta\boldsymbol{\beta}'\mathbf{A}_i\delta\boldsymbol{\beta}$, $\mathbf{A}_i = \mathbf{A}_i'$, $i = 1, \ldots, r$.

Definition 3.2 The measure of nonlinearity for the confidence ellipsoid is

$$C_{d(\cdot),\mathrm{conf}} =$$
$$= \sup \left\{ \frac{\sqrt{(\boldsymbol{\delta} - \mathbf{DC}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa})'(\mathbf{DC}^{-1}\mathbf{D}')^{-1}(\boldsymbol{\delta} - \mathbf{DC}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa})}}{\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta}} \colon \delta\boldsymbol{\beta} \in R^k \right\}.$$

Theorem 3.2 *Let* $\delta_{\max}$ *satisfies the equality*

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

where α *and* ε *are positive sufficiently small real numbers. Let*

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{d(\cdot),\mathrm{conf}} = \left\{ \delta\boldsymbol{\beta} \colon \delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \le \frac{2\sqrt{\delta_{\max}}}{C_{d(\cdot),\mathrm{conf}}} \right\}.$$

Then the ellipsoid

$$\mathcal{E}_{d(\cdot)} = \Big\{\mathbf{u} \in R^r\colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})' \leq \left(\sqrt{\chi^2_r(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2\Big\} + \mathbf{D}\widehat{\delta\boldsymbol{\beta}}$$

covers the vector

$$\mathbf{D}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\delta}(\delta\boldsymbol{\beta})$$

with probability larger or equal at least to $(1-\alpha-\varepsilon)$.

Proof The random variable

$$[\mathbf{d}(\boldsymbol{\beta}) - \mathbf{d}(\boldsymbol{\beta}_0) - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}]'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}[\mathbf{d}(\boldsymbol{\beta}) - \mathbf{d}(\boldsymbol{\beta}_0) - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}]$$

is chi-squared distributed with the parameter of noncentrrality

$$\delta = \frac{1}{4}(\boldsymbol{\delta} - \mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\boldsymbol{\delta} - \mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}).$$

From Definition 3.2 we have

$$4\delta \leq (C_{d(\cdot),\mathrm{conf}})^2(\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta})^2.$$

If $\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq 2\sqrt{\delta_{\max}}/C_{d(\cdot),\mathrm{conf}}$, then $\delta \leq \delta_{\max}$ and then the vector

$$E[\delta(\boldsymbol{\beta}_0) + \mathbf{D}\widehat{\delta\boldsymbol{\beta}}] - [\mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\delta}(\delta\boldsymbol{\beta})]$$

is an element of the ellipsoid

$$\Big\{\mathbf{u}\colon \mathbf{u} \in R^r, \mathbf{u}'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{u} \leq \delta_{\max}\Big\}$$

with probability at least $1-\alpha-\varepsilon$. Now it is obvious how to finish the proof. □

Corollary 3.2 *If the function* $\mathbf{d}(\cdot)$ *is linear, i.e.* $\mathbf{d}(\boldsymbol{\beta}) = \mathbf{d}(\boldsymbol{\beta}_0) + \mathbf{D}\delta\boldsymbol{\beta}$, *then*

$$\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \leq \frac{2\sqrt{\delta_{\max}}}{C_{D\beta}} \Rightarrow P\{\mathbf{d}(\boldsymbol{\beta}) \in \mathcal{E}\} \geq 1-\alpha-\varepsilon,$$

where

$$C_{D\beta} = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'\boldsymbol{\Sigma}^{-1}\mathbf{F}\mathbf{C}^{-1}\mathbf{D}'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}\mathbf{D}\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}}}{\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta}}\colon \delta\boldsymbol{\beta} \in R^k\right\},$$

$$\mathcal{E} = \Big\{\mathbf{u} + \mathbf{d}(\boldsymbol{\beta}_0)\colon (\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}})'(\mathbf{D}\mathbf{C}^{-1}\mathbf{D}')^{-1}(\mathbf{u} - \mathbf{D}\widehat{\delta\boldsymbol{\beta}}) \leq \left(\sqrt{\chi^2_r(0; 1-\alpha)} + \sqrt{\delta_{\max}}\right)^2\Big\}$$

and

$$P\{\chi^2_r(\delta_{\max}) \le \chi^2_r(0; 1-\alpha)\} = 1 - \alpha - \varepsilon$$

(cf. Theorem 3.2).

Corollary 3.3 *If the function $d(\cdot)$ is scalar, i.e. $d(\boldsymbol{\beta}) = d(\boldsymbol{\beta}_0) + \mathbf{d}'\delta\boldsymbol{\beta} + \frac{1}{2}\delta(\delta\boldsymbol{\beta})$, then*

$$\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta} \le \frac{2\sqrt{\delta_{\max}}}{C_{d(\beta)}} \Rightarrow P\{d(\boldsymbol{\beta}) \in \mathcal{E}\} \ge 1 - \alpha - \varepsilon,$$

where

$$\begin{aligned} C_{d(\beta)} &= \sup\left\{ \frac{\sqrt{\boxed{a}'(\mathbf{d}'\mathbf{C}^{-1}\mathbf{d})^{-1}\boxed{a}}}{\delta\boldsymbol{\beta}'\mathbf{C}\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k \right\}, \\ \mathcal{E} &= \left\{ u + d(\boldsymbol{\beta}_0) \colon (u - \mathbf{d}'\widehat{\delta\boldsymbol{\beta}})^2/(\mathbf{d}'\mathbf{C}^{-1}\mathbf{d}) \le \left(\sqrt{\chi^2_1(0;1-\alpha)} + \sqrt{\delta_{\max}}\right)^2 \right\}, \end{aligned}$$

and

$$\boxed{a} = [\delta(\delta\boldsymbol{\beta}) - \mathbf{d}'\mathbf{C}^{-1}\mathbf{F}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})].$$

Until now the parameter σ^2 is assumed to be known. Let

$$T(\delta\boldsymbol{\beta}) = U\frac{n-k}{k}, \qquad U = \frac{(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}})}{(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})}.$$

Then $T(\delta\boldsymbol{\beta}^*)$ has the Fisher–Snedecor distribution $F_{k,n-k}(\cdot)$ in case of the linearized version (3) of the regression model.

Lemma 3.1 *In the the quadratized model (2) we have*

(i) $(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}}) \sim \sigma^2\chi^2_k(\delta_1)$,

where $\delta_1 = \boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})/(4\sigma^2)$ *and*

(ii) $(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}})'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}}) \sim \sigma^2\chi^2_{n-k}(\delta_2)$,

where $\delta_2 = \boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta})/(4\sigma^2)$.

Proof (i) The parameter of noncentrality δ_1 is

$$\delta_1 = E(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}E(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}})/\sigma^2.$$

Since in the quadratized model (2)

$$\begin{aligned} E(\delta\boldsymbol{\beta}^* - \widehat{\delta\boldsymbol{\beta}}) &= \delta\boldsymbol{\beta}^* - (\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}\left[\mathbf{F}\delta\boldsymbol{\beta}^* + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*)\right] \\ &= -\frac{1}{2}(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*), \end{aligned}$$

the statement (i) is valid.

(ii) Analogously

$$E(\mathbf{Y} - \mathbf{f}_0 - \mathbf{F}\widehat{\delta\boldsymbol{\beta}}) = \mathbf{F}\delta\boldsymbol{\beta}^* + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*) - \mathbf{F}(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{V}^{-1} \times \left[\mathbf{F}\delta\boldsymbol{\beta}^* + \frac{1}{2}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*)\right] = \frac{1}{2}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}^*)$$

and therefore also the statement (ii) is valid. □

The probability density of the random variable $\chi^2_f(\delta)$ is [7]

$$g_{f,\delta}(y) = \begin{cases} \exp[-(y+\delta)/2] \sum_{r=0}^{\infty} \frac{1}{r!}\left(\frac{\delta}{2}\right)^r \frac{y^{r+(f/2)-1}}{2^{r+f/2}\Gamma(r+f/2)}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Thus the density of the random variable U is

$$g(u;\delta_1,\delta_2) = \int_0^\infty g_{k,\delta_1}(uv)g_{n-k,\delta_2}(v)v dv.$$

Let the set $C_{\delta_1^*,\delta_2^*}$ be defined as follows.

$$C_{\delta_1^*,\delta_2^*} = \left\{(\delta_1^*,\delta_2^*)\colon \int_0^{[(n-k)/k]F_{k,n-k}(0;1-\alpha)} g(u;\delta_1^*,\delta_2^*)du = 1-\alpha-\varepsilon\right\}.$$

Theorem 3.3 *The linearization region for the confidence ellipsoid in the case of the estimated* σ^2 *is*

$$\mathcal{L}_{\delta_1,\delta_2} = \left\{\delta\boldsymbol{\beta}\colon \delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} < \sigma\frac{2\sqrt{\delta_1^*}}{K_0^{(par)}} \quad \& \quad \delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} < \sigma\frac{2\sqrt{\delta_2^*}}{K_0^{(int)}}\right\}$$

i.e.

$$\delta\boldsymbol{\beta} \in \mathcal{L}_{\delta_1,\delta_2} \Rightarrow P\{\delta\boldsymbol{\beta} \in \mathcal{E}\} \geq 1-\alpha-\varepsilon,$$

where $\mathcal{E}$ *is given by (5).*

Proof With respect to Definition 3.3 it is valid that

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) \leq \frac{1}{4\sigma^4}(\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta})^2\left(\sigma K_0^{(par)}\right)^2.$$

Thus the inequality

$$\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} \leq \sigma\frac{2\sqrt{\delta_1^*}}{K_0^{(par)}}$$

implies

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{P}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \delta_1 \leq \delta_1^*.$$

Analogously

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) \leq \frac{1}{4\sigma^4}(\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta})^2\left(\sigma K_0^{(int)}\right)^2.$$

Thus the inequality

$$\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\delta\boldsymbol{\beta} \leq \sigma\frac{2\sqrt{\delta_2^*}}{K_0^{(int)}}$$

implies

$$\frac{1}{4\sigma^2}\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{V}^{-1}\mathbf{M}_F^{V^{-1}}\boldsymbol{\kappa}(\delta\boldsymbol{\beta}) = \delta_2 \leq \delta_2^*.$$

□

In order not to prefer one of the parameter noncentrality for the other one, the condition

$$\delta_1^*/\delta_2^* = (K_0^{(par)})^2/(K_0^{(int)})^2$$

can be used. Thus

$$\frac{2\sqrt{\delta_1^*}}{K_0^{(par)}} = \frac{2\sqrt{\delta_2^*}}{K_0^{(int)}}.$$

In some cases the Bates and Watts intrinsic curvature is zero and thus the random variable $T = [(n-k)/k]U$ has the the noncentral Fisher–Snedecor distribution

$$F_{k,n-k}(\delta_1) = [\chi_k^2(\delta_1)/k]/[\chi_{n-k}^2(0)/(n-k)],$$

since $\delta_2 = 0$. Let $\delta_{1,\max}$ be solution of the equation

$$P\{F_{k,n-k}(\delta_{1,\max}) \geq F_{k,n-k}(0;1-\alpha)\} = \alpha + \varepsilon.$$

If

$$C^{(ell,D\delta\beta)} = \sigma C_0^{(ell,D\delta\beta)},$$

where

$$C_0^{(ell,D\delta\beta)} = $$
$$= \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'\mathbf{V}^{-1}\mathbf{F}\mathbf{C}_0^{-1}\mathbf{D}'(\mathbf{D}\mathbf{C}_0^{-1}\mathbf{D}')^{-1}\mathbf{D}\mathbf{C}_0^{-1}\mathbf{F}'\mathbf{V}^{-1}\boldsymbol{\kappa}}}{\delta\boldsymbol{\beta}'\mathbf{C}_0\delta\boldsymbol{\beta}} : \delta\boldsymbol{\beta} \in R^k\right\},$$

where $\mathbf{C}_0 = \mathbf{F}'\mathbf{V}^{-1}\mathbf{F}$, then the following implication is valid.

$$\begin{aligned}\delta\boldsymbol{\beta}'\mathbf{C}_0\delta\boldsymbol{\beta} \;\leq\; &\sigma\frac{2\sqrt{\delta_{1,\max}}}{C_0^{(ell,D\delta\beta)}}\\ &\Rightarrow P\Big\{(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{D}'(\mathbf{D}\mathbf{C}_0\mathbf{D}')^{-1}\mathbf{D}(\delta\boldsymbol{\beta} - \widehat{\delta\boldsymbol{\beta}})\\ &\leq k\widehat{\sigma^2}F_{k,n-k}(0;1-\alpha)\Big\} \geq 1-\alpha-\varepsilon.\end{aligned}$$

It is to be remarked that in the case $\mathbf{D} = \mathbf{I}$, i.e. the confidence ellipsoid for the parameter $\delta\boldsymbol{\beta}$ must be determined, then the equality $C^{(ell,\delta\boldsymbol{\beta})} = K^{(par)}$ can be used.

In the case that σ^2 must be estimated, the decision whether linearization of the model with respect to the confidence ellipsoid can be used, is made with some uncertainty. Therefore a comparison of the given procedure with the exact determination, which is in this case known (cf. [8]), is interesting.

Lemma 3.2 *Let in the model* $\mathbf{Y} \sim N_n[\mathbf{f}(\boldsymbol{\beta}), \boldsymbol{\Sigma}]$, $\boldsymbol{\beta} \in R^k$, *the matrix* $\boldsymbol{\Sigma}$ *be known.*

(i) Then the set

$$\mathcal{C}_\beta = \left\{\boldsymbol{\beta} \colon [\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}]' \left(\mathbf{P}^{\Sigma^{-1}}_{F(\beta)}\right)' \boldsymbol{\Sigma}^{-1} \mathbf{P}^{\Sigma^{-1}}_{F(\beta)} [\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}] \le \chi^2_k(0; 1-\alpha)\right\}$$

is the exact $(1-\alpha)$*-confidence region for the parameter* $\boldsymbol{\beta}$.

(ii) If the matrix $\boldsymbol{\Sigma}$ *is of the form* $\boldsymbol{\Sigma} = \sigma^2\mathbf{V}$, *where* $\mathbf{V}$ *is a given* $n \times n$ *p.d. matrix and* σ^2 *is unknown parameter, then the exact* $(1-\alpha)$*-confidence set for the parameter* $\boldsymbol{\beta}$ *is*

$$\mathcal{D}_\beta = \left\{\boldsymbol{\beta} \colon [\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}]' \left(\mathbf{P}^{V^{-1}}_{F(\beta)}\right)' \mathbf{V}^{-1} \mathbf{P}^{V^{-1}}_{F(\beta)} [\mathbf{f}(\boldsymbol{\beta}) - \mathbf{Y}] \right.$$
$$\left. \le \frac{k}{n-k} [\mathbf{Y} - \mathbf{f}(\boldsymbol{\beta})]' \left(\mathbf{M}^{V^{-1}}_{F(\beta)}\right)' \mathbf{V}^{-1} \mathbf{M}^{V^{-1}}_{F(\beta)} [\mathbf{Y} - \mathbf{f}(\boldsymbol{\beta})] F_{k,n-k}(0; 1-\alpha)\right\}.$$

Numerical determination of the exact confidence regions is tedious and time consuming unlike procedure given by a linearization.

4 Numerical example

Consider the Michaelis–Menten model, i.e.

$$f_i(\beta_1, \beta_2) = \frac{x_i\beta_1}{x_i + \beta_2}, \quad x_i = 1, 2, 3, 4, 6$$

and $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$, $\sigma = 0.1$.

If $\beta_1 = 4$ and $\beta_2 = 1$, then (cf. [12])

$$\{\mathbf{F}\}_{i,\cdot} = \left(\frac{x_i}{1+x_i}, -\frac{4x_i}{(1+x_i)^2}\right), \quad i = 1, 2, 3, 4, 5,$$

$$\mathbf{F}_i = \begin{pmatrix} 0, & -\frac{x_i}{(1+x_i)^2} \\ -\frac{x_i}{(1+x_i)^2}, & \frac{8x_i}{(1+x_i)^3} \end{pmatrix}, \quad i = 1, 2, 3, 4, 5,$$

$$K_0^{(int)} = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{M}_F\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{F}\delta\boldsymbol{\beta}} \colon \delta\boldsymbol{\beta} \in R^k\right\} = 0.3326,$$

$$K_0^{(par)} = \sup\left\{\frac{\sqrt{\boldsymbol{\kappa}'(\delta\boldsymbol{\beta})\mathbf{P}_F\boldsymbol{\kappa}(\delta\boldsymbol{\beta})}}{\delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{F}\delta\boldsymbol{\beta}} \colon \delta\boldsymbol{\beta} \in R^k\right\} = 1.3212.$$

Let $\sigma(= 0.1)$ be known and let $\varepsilon = 0.05$, i.e. $\delta_{\max} = 0.6398$. Then the linearization region for the confidence ellipsoid is

$$\mathcal{L}_{\mathcal{E}} = \left\{ \delta\boldsymbol{\beta} \colon \delta\boldsymbol{\beta}'\mathbf{F}'\mathbf{F}\delta\boldsymbol{\beta} \leq \sigma \frac{2\sqrt{\delta_{\max}}}{K_0^{(par)}(\boldsymbol{\beta}_0)} \right\}$$

and the 0.95-confidence ellipsoid for $\delta\boldsymbol{\beta}$ is

$$\mathcal{E} = \left\{ \mathbf{u} \colon (\mathbf{u} - \widehat{\delta\boldsymbol{\beta}})'\mathbf{F}'\mathbf{F}(\mathbf{u} - \boldsymbol{\beta}_0 - \widehat{\delta\boldsymbol{\beta}}) \leq 0.0599 \right\}$$

cf. Fig. 1.

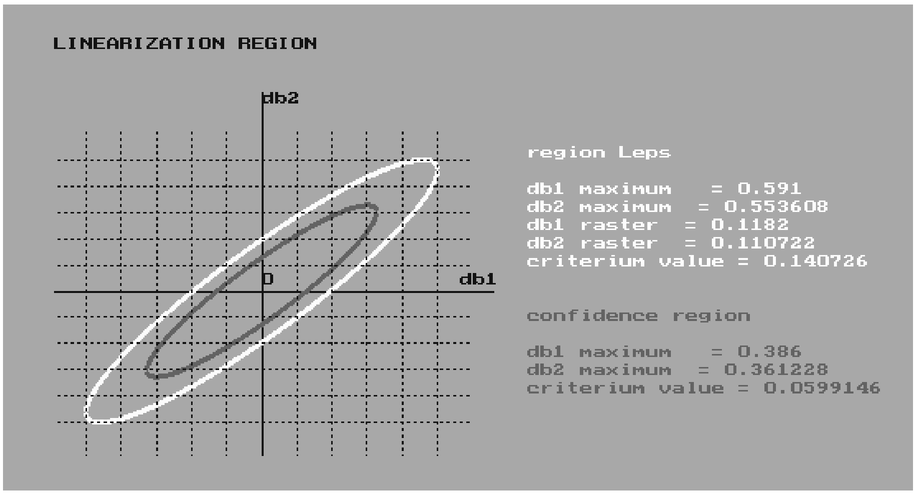


Fig. 1: 0.95-confidence ellipse for $\delta\boldsymbol{\beta}$ and the region $\mathcal{L}_{\mathcal{E}}$

Let set of measured data $\mathbf{y}$ are simulated for $\sigma = 0.1$, i.e.

$$\mathbf{y} = (1.90, 2.57, 3.08, 3.13, 3.58)'.$$

If $\delta_1/\delta_2 = (K^{(par)}/K^{(int)})^2 = 15.779\,478 = t$, then the set $C_{\delta_1^*,\delta_2^*}$ consists of a single point which is a solution of the equations

$$\int_0^{\frac{3}{2}9.552} \left[\int_0^{\infty} g_{2,\delta_1^*}(uv) g_{3,\delta_2^*}(v) v dv \right] du = 0.95 - 0.05, \quad \delta_1^* = t\delta_2^*,$$

where

$$g_{f,\delta}(y) = \begin{cases} \exp[-(y+\delta)/2] \displaystyle\sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{\delta}{2}\right)^r \frac{y^{r+(f/2)-1}}{2^{r+f/2}\Gamma(r+f/2)}, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

In this case the linearization region from Theorem 3.3 is given in Fig. 2.

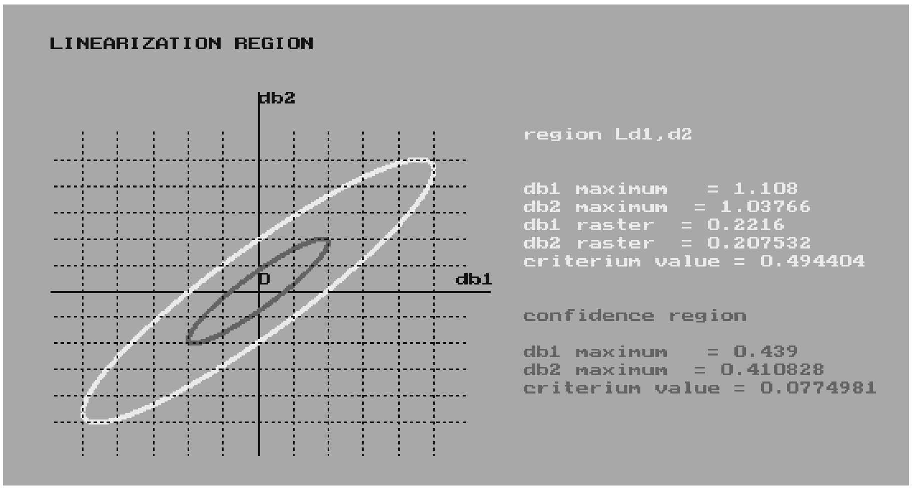


Fig. 2: The region $\mathcal{L}_{\delta_1,\delta_2}$ and 0.95-confidence ellipse (5)

The set $\mathcal{D}_\beta$ from Lemma 3.2 is given for $1-\alpha = 0.90$ at Fig. 3.

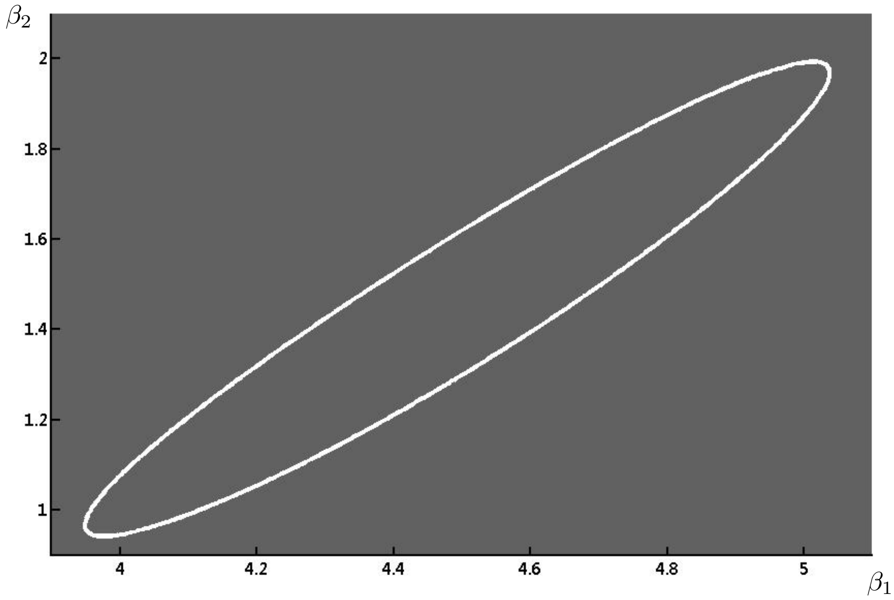


Fig. 3: The set $\mathcal{D}_\beta$ from Lemma 3.2

Acknowledgement Authors would like to express their thanks to Mgr. J. Marek, Ph.D. for his thoroughly check of the manuscript and a help with numerical calculation.

References

[1] Bates, D. M., Watts, D. G.: *Relative curvature measures of nonlinearity.* J. Roy. Stat. Soc. **B 42** (1980), 1–25.

[2] Kubáček, L., Kubáčková, L., Volaufová, J.: Statistical Models with Linear Structures. *Veda (Publishing House of Slovak Academy of Science), Bratislava*, 1995.

[3] Kubáček, L.: *On a linearization of regression models.* Applications of Mathematics **40** (1995), 61–78.

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

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

[6] Kubáček, L., Tesaříková, E.: *Confidence regions in nonlinear models with constraints.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **42**, 2003, 407–426.

[7] Kubáčková, L.: Foundations of Experimental Data Analysis. *CRC-Press, Boca Raton-Ann Arbor–London–Tokyo*, 1992.

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

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

[10] Tesaříková, E., Kubáček, L.: *How to deal with regression models with a weak nonlinearity.* Discuss. Math., Probab. Stat. **21**, 2001, 21–48.

[11] Tesaříková, E., Kubáček, L.: *Estimators of dispersion in models with constraints (demoprogram).* Dept. Algebra and Geometry, Fac. Sci., Palacký Univ., Olomouc, 2003.

[12] Tesaříková, E., Kubáček, L.: *Linearization regions for confidence ellipsoids (demoprogram).* Department of Algebra and Geometry, Faculty of Dept. Algebra and Geometry, Fac. Sci., Palacký Univ., Olomouc, 2007.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 115–119

A Result on Segmenting Jungck–Mann Iterates

Memudu Olaposi OLATINWO

*Department of Mathematics, Obafemi Awolowo University,
Ile-Ife, Nigeria
e-mail: polatinwo@oauife.edu.ng*

(Received May 5, 2007)

Abstract

In this paper, following the concepts in [5, 7], we shall establish a convergence result in a uniformly convex Banach space using the Jungck–Mann iteration process introduced by Singh et al [13] and a certain general contractive condition. The authors of [13] established various stability results for a pair of nonself-mappings for both Jungck and Jungck–Mann iteration processes. Our result is a generalization and extension of that of [7] and its corollaries. It is also an improvement on the result of [7].

Key words: Jungck–Mann iteration process; uniformly convex Banach space.

2000 Mathematics Subject Classification: 47H06, 47H10

1 Introduction

Suppose that $A = (a_{nk})$ is an infinite, lower triangular, regular row-stochastic matrix, E a closed convex subset of a Banach space and T a continuous mapping of E into itself and $x_1 \in E$. Then, the general Mann iteration process $M(x_1, A, T)$ which was introduced in Mann [9] is defined by

$$v_n = \sum_{k=1}^{n} a_{nk} x_k, \quad x_{n+1} = T v_n, \quad n = 1, 2, \dots, \tag{1}$$

If A is the identity matrix, then each sequence of $M(x_1, A, T)$ becomes the sequence of Picard iterates of T at x_1. It was established in [9] that if either of the sequences $\{x_n\}$ a nd $\{v_n\}$ converges, then the other also converges to the same point, and their common limit is a fixed point of T.

In [5, 7], it is said that the matrix A is *segmenting* for the Mann process if $a_{n+1,k} = (1 - a_{n+1,n+1})a_{nk}$ for $k \leq n$. In this case, v_{n+1} lies on the segment joining v_n and Tv_n:

$$v_{n+1} = (1 - d_n)v_n + d_nTv_n, \quad n = 1, 2, \ldots, \tag{2}$$

where $d_n = a_{n+1,n+1}$. A segmenting matrix is determined by its sequence of diagonal elements. Some authors including [3, 11, 12] have investigated the case $d_n = \lambda$, $0 < \lambda < 1$, while Mann [9] approximated the fixed points of continuous functions on a closed interval of the real line using the segmenting matrix determined by $d_n = \frac{1}{n}\ \forall\ n$. Dotson [6] considered the case when d_n is bounded away from 0 and 1. Groetsch [7] generalized the results of [3, 6, 9, 11, 12] in a uniformly convex Banach space by employing (2) and assuming that A is a segmenting matrix for which $\sum_{n=1}^{\infty} d_n(1 - d_n) = \infty$.

We shall give another definition of a segmenting matrix in the next section with a view to generalizing and extending Groetsch [7] and others mentioned earlier in this paper.

2 Preliminaries

Singh et al [13] introduced the following iteration process: Let $(E, \|.\|)$ be a normed linear space, $S, T \colon Y \to E$ and $T(Y) \subseteq S(Y)$. Then, for $x_0 \in Y$, consider the iteration process

$$Sx_{n+1} = (1 - \alpha_n)Sx_n + \alpha_nTx_n, \quad n = 0, 1, 2, \ldots, \tag{3}$$

where $\{\alpha_n\}_{n=0}^{\infty}$ satisfies

(i) $\alpha_0 = 1$,

(ii) $0 \leq \alpha_n \leq 1$ for $n > 0$,

(iii) $\sum \alpha_n = \infty$, and

(iv) $\sum_{j=0}^{n} \alpha_j \Pi_{i=j+1}^{n}(1 - \alpha_i + a\alpha_i)$ converges.

The iteration process (3) is called the *Jungck–Mann iteration.*

For $Y = E$, $S = I$ (identity operator) in (3) with $\{\alpha_n\}_{n=0}^{\infty}$ satisfying (i)–(iv), then we have the Mann iteration process introduced by Mann [9]. Also, if in (3), $Y = E$, $S = I$ (identity operator) and $\alpha_n = 1$, then we obtain the Jungck iteration introduced by Jungck [8].

Following (3), we shall generalize and extend Groetsch [7] and others mentioned earlier in this paper by assuming that A is a segmenting matrix for which

$$Sv_{n+1} = (1 - d_n)Sv_n + d_nTv_n, \quad n = 1, 2, \ldots, \tag{$\star$}$$

such that $\sum_{n=1}^{\infty} d_n(1-d_n) = \infty$ and $S, T\colon C \to C$ are selfmappings on a nonempty convex subset C of a uniformly convex Banach space E. The operators S and T are assumed to have a common fixed point and satisfy in addition the contractive condition

$$\|Tx - Ty\| \le \|Sx - Sy\|, \quad \forall x, y \in C. \tag{$\star\star$}$$

If $S = I$ (identity operator) in ($\star$), then we obtain (2) and if $S = I$ in ($\star\star$) then we have $\|Tx - Ty\| \le \|x - y\|$, $\forall x, y \in C$ (that is, T becomes a nonexpansive mapping).

We shall establish our main result in the next section. However, the following lemma is required in the sequel.

Lemma 2.1 (Groetsch [7]) *Let X be a uniformly convex Banach space and let $x, y \in X$. If $\|x\| \le 1$, $\|y\| \le 1$ and $\|x - y\| \ge \epsilon > 0$, then*

$$\|\lambda x + (1-\lambda)y\| \le 1 - 2\lambda(1-\lambda)\delta(\epsilon)$$

for $0 \le \lambda < 1$ and $\delta(\epsilon) > 0$.

The proof of this Lemma is contained in [4, 7].

3 The Main Result

Theorem 3.1 *Let C be a convex subset of a uniformly convex Banach space E and $S, T\colon C \to C$ selfmappings satisfying condition ($\star\star$) and $T(C) \subseteq S(C)$. Suppose that S and T have at least a common fixed point. Let $\{Sv_n\}_{n=1}^{\infty}$ be the sequence defined by ($\star$). Then, the sequence $\{(S-T)v_n\}_{n=1}^{\infty}$ converges strongly to 0 for each $x_1 \in C$ such that $\sum_{n=1}^{\infty} d_n(1-d_n) = \infty$.*

Proof If p is a common fixed point of S and T (i.e. $Sp = Tp = p$), then

$$\begin{aligned}
\|Sv_{n+1} - p\| &= \|(1-d_n)Sv_n + d_nTv_n - (1 - d_n + d_n)p\| \\
&= \|(1-d_n)(Sv_n - p) + d_n(Tv_n - p)\| \\
&\le (1-d_n)\|Sv_n - p\| + d_n\|Tv_n - p\| \\
&= (1-d_n)\|Sv_n - p\| + d_n\|Tv_n - Tp\| \\
&\le (1-d_n)\|Sv_n - p\| + d_n\|Sv_n - Sp\| \\
&= (1-d_n)\|Sv_n - p\| + d_n\|Sv_n - p\| \\
&= \|Sv_n - p\| \le \|Sv_{n-1} - p\| \le \cdots \le \|Sv_1 - p\|,
\end{aligned} \tag{4}$$

from which we have that the sequence $\{Sv_n - p\}_{n=1}^{\infty}$ is decreasing.

Now,

$$\begin{aligned}
&\|(S-T)v_n\| = \|Sv_n - Tv_n\| \le \|Sv_n - p\| + \|p - Tv_n\| \\
&\quad = \|Sv_n - p\| + \|Tp - Tv_n\| \le \|Sv_n - p\| + \|Sp - Sv_n\| = 2\|Sv_n - p\|.
\end{aligned}$$

Suppose on the contrary that $\{(S-T)v_n\}_{n=1}^{\infty}$ does not converge to 0. Since $\|Sv_n - Tv_n\| \leq 2\|Sv_n - p\|$, we may assume that there is an $a > 0$, $a \in (0,1)$ such that $\|Sv_n - p\| \geq a$ for any n. If $\{(S-T)v_n\}_{n=1}^{\infty}$ does not converge to 0, then there is an $\epsilon > 0$ such that $\|Sv_n - Tv_n\| \geq \epsilon$ for any n.

Let

$$b = 2\delta\left(\frac{\epsilon}{\|Sv_1 - p\|}\right), \quad x_n = \frac{Sv_n - p}{\|Sv_n - p\|} \quad \text{and} \quad y_n = \frac{Tv_n - p}{\|Sv_n - p\|}.$$

Then, we have

$$\|x_n\| = \left\|\left(\frac{Sv_n - p}{\|Sv_n - p\|}\right)\right\| \leq \frac{\|Sv_n - p\|}{\|Sv_n - p\|} = 1$$

and

$$\|y_n\| = \left\|\left(\frac{Tv_n - p)}{\|Sv_n - p\|}\right)\right\| \leq \frac{\|Tv_n - Tp)\|}{\|Sv_n - p\|} \leq \frac{\|Sv_n - Sp\|}{\|Sv_n - p\|} = \frac{\|Sv_n - p\|}{\|Sv_n - p\|} = 1.$$

Hence, we have by $(\star)$ that

$$\begin{aligned} \|Sv_{n+1} - p\| &= \|(1-d_n)Sv_n + d_nTv_n - (1 - d_n + d_n)p\| \\ &= \|(1-d_n)(Sv_n - p) + d_n(Tv_n - p)\| \\ &= \left\|(\|Sv_n - p\|)\left[(1-d_n)\frac{(Sv_n - p)}{\|Sv_n - p\|} + d_n\frac{(Tv_n - p)}{\|Sv_n - p\|}\right]\right\| \\ &= \|(\|Sv_n - p\|)[(1-d_n)x_n + d_ny_n]\| \\ &\leq \|Sv_n - p\| \; \|(1-d_n)x_n + d_ny_n\|. \end{aligned} \tag{5}$$

Using (4) and Lemma 2.1 in (5) yield

$$\begin{aligned} &\|Sv_{n+1} - p\| \leq \\ &\leq [1 - d_n(1-d_n)b]\|Sv_n - p\| \\ &= \|Sv_n - p\| - bd_n(1-d_n)\|Sv_n - p\| \\ &\leq \|Sv_{n-1} - p\| - bd_{n-1}(1-d_{n-1})\|Sv_{n-1} - p\| - bd_n(1-d_n)\|Sv_n - p\| \\ &\leq \|Sv_{n-1} - p\| - bd_{n-1}(1-d_{n-1})\|Sv_n - p\| - bd_n(1-d_n)\|Sv_n - p\| \\ &= \|Sv_{n-1} - p\| - b[d_{n-1}(1-d_{n-1}) + d_n(1-d_n)]\|Sv_n - p\|. \end{aligned}$$

Repeating this process inductively leads to

$$a \leq \|Sv_{n+1} - p\| \leq \|Sv_1 - p\|$$

$$-\, b\Big[d_1(1-d_1)\|Sv_n - p\| + d_2(1-d_2)\|Sv_n - p\| + \cdots + d_n(1-d_n)\|Sv_n - p\|\Big]$$

$$= \|Sv_1 - p\| - b\sum_{j=1}^{n} d_j(1-d_j)\|Sv_n - p\| \leq \|Sv_1 - p\| - ab\sum_{j=1}^{n} d_j(1-d_j).$$

Therefore, we obtain

$$a\left[1+b\sum_{j=1}^{n} d_j(1-d_j)\right] \leq \|Sv_1 - p\|,$$

from which it follows that

$$a \leq \frac{\|Sv_1 - p\|}{1+b\sum_{j=1}^{n} d_j(1-d_j)} \to 0 \quad \text{as } n\to\infty,$$

leading to a contradiction. Therefore, we have $a = 0$. Hence,

$$\lim_{n\to\infty} \|Sv_n - Tv_n\| = 0.$$

Remark 3.1 Theorem 3.1 is also a generalization of the results of [3, 6, 7, 9, 11, 12].

References

[1] Browder, F. E.: *Nonexpansive nonlinear operators in a Banach space.* Proc. Acad. Sci. **54** (1965), 1041–1044.

[2] Browder, F. E., Petryshyn, W. V.: *The solution by iteration of nonlinear functional equations in Banach spaces.* Bull. Amer. Math. Soc. **72** (1966), 571–575.

[3] Browder, F. E., Petryshyn, W. V.: *Construction of fixed points of nonlinear mappings in a Hilbert space.* J. Math. Anal. Appl. **20** (1967), 197–228.

[4] Chidume, C. E.: *Geometric properties of Banach spaces and nonlinear iterations.* The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy, October 2000.

[5] Outlaw, C., Groetsch, C. W.: *Averaging iteration in a Banach space.* Bull. Amer. Math. Soc. **75** (1969), 430–432.

[6] Dotson, W. G., Jr.: *On the Mann iterative process.* Trans. Amer. Math. Soc. **149** (1970), 65–73.

[7] Groetsch, C. W.: *A Note on Segmenting Mann Iterates.* J. Math. Anal. Appl. **40** (1972), 369–372.

[8] Jungck, G.: *Commuting mappings and fixed points.* Amer. Math. Monthly **83**, 4 (1976), 261–263.

[9] Mann, W. R.: *Mean value methods in iteration.* Proc. Amer. Math. Soc. **4** (1953), 506–510.

[10] Opial, Z.: *Weak convergence of the successive approximations for nonexpansive mappings in Banach spaces.* Bull. Amer. Math. Soc. **73** (1967), 591–597.

[11] Petryshyn, W. V.: *Construction of fixed points of demicompact mappings in Hilbert space.* J. Math. Anal. Appl. **14** (1966), 276–284.

[12] Schaefer, H.: *Über die Methode Sukzessiver Approximationen.* Jahresber. Deutsch. Math. Verein. **59** (1957), 131–140.

[13] Singh, S. L., Bhatnagar, C., Mishra, S. N.: *Stability of Jungck-type iterative procedures.* Internat. J. Math. & Math. Sci. **19** (2005), 3035–3043.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 121–127

Further Results on Global Stability of Solutions of Certain Third-order Nonlinear Differential Equations

Mathew Omonigho OMEIKE

Department of Mathematics, University of Agriculture, Abeokuta, Nigeria
e-mail: moomeike@yahoo.com

(Received July 17, 2008)

Abstract

Sufficient conditions are established for the global stability of solutions of certain third-order nonlinear differential equations. Our result improves on Tunc's [10].

Key words: Nonlinear differential equation; trivial solution; global stability; Lyapunov's method.

2000 Mathematics Subject Classification: 34D23

1 Introduction

We consider the third-order nonlinear ordinary differential equation

$$\dddot{x} + \psi(x, \dot{x}, \ddot{x})\ddot{x} + f(x, \dot{x}) = 0 \tag{1.1}$$

or its equivalent system

$$\dot{x} = y, \quad \dot{y} = z, \quad \dot{z} = -\psi(x, y, z)z - f(x, y), \tag{1.2}$$

where

$$\psi, \psi_x, \psi_z \in C(\mathbb{R} \times \mathbb{R} \times \mathbb{R}, \mathbb{R}) \quad \text{and } f, f_x, f_y \in C(\mathbb{R} \times \mathbb{R}, \mathbb{R}). \tag{1.3}$$

It is assumed that solutions of (1.1) exist and are unique.

Stability is a very important problem in the theory and application of differential equations, and an effective method for studying the stability of nonlinear differential equations is the second method of Lyapunov (see [1–14]).

In a recently paper, Tunc [10] obtained the global stability of (1.1) and the following result was proved.

Theorem A (Tunc [10]). *Further to the basic assumptions on the functions ψ and f suppose the following:*

(i) $xf(x,0) > 0$ *for* $x \neq 0$*;*

(ii) $\int_0^y f(0,v)\,dv \geq 0$*;*

(iii) $\lim_{|x|\to\infty} \sup \int_0^x f(u,0)\,du = \infty$*;*

(iv) there is a positive constant B *such that* $\psi(x,y,z) \geq B$ *for all* x, y, z*;*

(v)

$$B\left[f(x,y) - f(x,0) - \int_0^y \psi_x(x,v,0)v\,dv\right] y \geq y\int_0^y f_x(x,v)\,dv$$

for all x, y*;*

(vi)

$$B\left[f(x,y) - f(x,0) - \int_0^y \psi_x(x,v,0)v\,dv\right] y + \psi(x,y,z) \geq y\int_0^y f_x(x,v)\,dv + B$$

for all $x, y \neq 0, z$*;*

(vii)

$$4B\int_0^x f(u,0)\,du \left\{\int_0^y [f(x,v) - f(x,0)]\,dv + B\int_0^y [\psi(x,v,0) - B]\,v\,dv\right\} \geq y^2 f^2(x,0)$$

for all $x, y \neq 0$*;*

(viii) $y\psi_z(x,y,z) \geq 0$ *for all* x, y, z

Then the trivial solution of equation (1.1) is globally asymptotically stable.

Interestingly, (1.1) is a rather general third-order nonlinear differential equation. In particular, many third-order differential equations which have been discussed in [12] are special cases of (1.1), and some known results can be obtained using this theorem. However, it is not easy to apply Theorem A to these special cases to obtain new or better results since Theorem A has some hypotheses which are not necessary for the stability of many nonlinear equations.

Our aim in this paper is to further study the global stability of (1.1). In the next section, we establish a criterion for the stability of (1.1), which extends and improve Theorem A. Finally, in Section 3, we apply our result to some examples.

In the following discussion, we always assume (1.3) holds without further mention.

2 Main result

Our main result in this section is the following theorem.

Theorem *Let δ_0, a, b, c be positive constants such that $ab > c$. Assume that*

(1) $\frac{f(x,0)}{x} \geq \delta_0, \quad x \neq 0, \quad f(0,0) = 0,$

(2) $f'(x,0) \leq c,$

(3) $f_y(x, \theta y) \geq b$ *for* $0 \leq \theta \leq 1,$

(4) $\psi(x, y, z) > a,$

(5) $y\psi_z(x, y, \theta z) \geq 0,$ *for* $0 \leq \theta \leq 1,$

(6) $a\left[f(x,y) - f(x,0) - \int_0^y \psi_x(x,v,0)v\,dv\right] y \geq y \int_0^y f_x(x,v)\,dv.$

Then, the trivial solution of (1.1) is globally asymptotically stable.

Remark 1 The theorem just stated above improves the theorem established in [1] and includes the result established in [9]. The results of Ezeilo [2], Ogurtsov [5] and Goldwyn and Narendra [3] are also direct consequences of our result.

Proof Clearly, (1.1) is equivalent to the system (1.2) and $(0,0,0)$ is a solution. Now, consider the Lyapunov function

$$V(x,y,z) = \int_0^x f(u,0)\,du + \int_0^y \psi(x,v,0)v\,dv + a^{-1}\int_0^y f(x,v)\,dv + \frac{1}{2}a^{-1}z^2 + yz \tag{2.1}$$

This is rewritten as

$$V(x,y,z) = \frac{1}{2a}(ay+z)^2 + \frac{1}{2ab}(f(x,0)+by)^2 + \int_0^y [\psi(x,v,0)-a]v\,dv + \frac{1}{a}\int_0^y [f_v(x,\theta v) - b]v\,dv + \int_0^x [1 - \frac{1}{ab}f'(u,0)]f(u,0)\,du$$

where $f_v(x,\theta v) = v^{-1}\{f(x,v) - f(x,0)\}$, $v \neq 0$.

On using hypotheses (1)–(4) of the theorem,

$$V(x,y,z) \geq \frac{1}{2a}(ay+z)^2 + \frac{1}{2ab}(f(x,0)+by)^2 + \frac{1}{2}\delta_1 x^2,$$

where $\delta_1 = \frac{1}{ab}(ab-c)\delta_0 > 0$. It follows that there exists a constant $K > 0$ small enough that

$$V(x,y,z) \geq K(x^2+y^2+z^2).$$

Hence $V(x,y,z)$ is a positive definite function.

Next, we show that the derivative of $V(x,y,z)$ with respect to t along the solution path of (1.2) is negative semi definite.

$$\dot{V}_{(1.2)} = V_x\dot{x} + V_y\dot{y} + V_z\dot{z}, \tag{2.2}$$

where V_x, V_y, V_z are partial derivatives of V with respect to x, y and z respectively, and $\dot{x}, \dot{y}$ and $\dot{z}$ are as in (1.2).
Thus,

$$\begin{aligned} V_x &= f(x,0) + \int_0^y \psi_x(x,v,0)v\,dv + \frac{1}{a}\int_0^y f_x(x,v)\,dv, \\ V_y &= \psi(x,y,0)y + \frac{1}{a}f(x,y) + z, \qquad V_z = \frac{1}{a}z + y. \end{aligned}$$

Then, substituting V_x, V_y, V_z in (2.2) and using (1.2) yield

$$\dot{V}_{(1.2)}(x,y,z) = -\left\{f(x,y) - f(x,0) - \int_0^y \psi_x(x,v,0)v\,dv\right\}y$$
$$+\frac{1}{a}y\int_0^y f_x(x,v)\,dv - \psi_z(x,y,\theta z)yz^2 - \left[\frac{1}{a}\psi(x,y,z) - 1\right]z^2,$$

where $\psi_z(x,y,\theta z) = z^{-1}\{\psi(x,y,z) - \psi(x,y,0)\}$, $z \neq 0$. From hypotheses (4), (5) and (6) of theorem, we see that

$$\dot{V}_{(1.2)}(x,y,z) \leq 0, \tag{2.3}$$

and the rest of the proof may now follow as in [2, 9].

Let Ω denote a trajectory $x(t), y(t), z(t)$ of (1.2) satisfying the initial conditions $x(0) = x_0$, $y(0) = y_0$, $z(0) = z_0$, where (x_0, y_0, z_0) is an arbitrary point of the (x,y,z)-space. Then, by (2.3),

$$V(t) \equiv V(x(t),y(t),z(t)) \leq V(x_0,y_0,z_0) \qquad (t \geq 0). \tag{2.4}$$

Further, $V(t)$, being non-increasing and non-negative, tends to a non-negative limit, $V(\infty)$ say, as $t \to \infty$. To prove the theorem, it is sufficient to show that

$$V(\infty) \not> 0; \tag{2.5}$$

for, in that event, we should have $V(\infty) = 0$, and this would imply $x(\infty) = 0$, $y(\infty) = 0$, $z(\infty) = 0$, which is the required result.

Suppose on the contrary that (2.5) is not true: that is, assume that $V(\infty) > 0$. Since the set points (x, y, z) for which

$$V(x, y, z) \leq V(x_0, y_0, z_0)$$

is bounded, it is clear from (2.4) that the trajectory Ω has limit points; and the set of all its limit points consists of whole trajectories of (1.2) lying on the surface $V(x, y, z) = V(\infty)$. Thus, in particular, if Q is a limit point of Ω, there is a half-trajectory, Ω_Q say, of (1.2) issuing from Q and lying on the surface $V(x, y, z) = V(\infty)$. Evidently, we must have

$$\dot{V}_{(1.2)} \equiv 0 \tag{2.6}$$

on Ω_Q; for otherwise there would exist points (x, y, z) of Ω_Q at which

$$V(x, y, z) < V(\infty).$$

From (2.4) and (2.6) it follows readily that $z = 0$ and hence also that $y \equiv \gamma$, $x = \gamma t + \xi$ (γ, ξ constants), $\dot{z} = 0$ for any (x, y, z) on Ω_Q.

Also, since from (1.2),

$$\dot{z} = -\psi(x, y, z)z - f(x, y),$$

it follows that $f(x, y) = 0$, that is

$$f(\gamma t + \xi, \gamma) = 0. \tag{2.7}$$

Since $f(0, 0) = 0$, (2.7) clearly holds if and only if $\xi = \gamma = 0$ (see, for example, [12, p. 370]). Hence $x = 0$, $y = 0$. We have therefore that $x = y = z = 0$; this implies that the origin is a point of the surface $V(x, y, z) = V(\infty)$, which contradicts our assumption that $V(\infty) > 0$. This proves (2.5) and hence the theorem. □

Remark 2 Clearly our theorem is an improvement and extension of Theorem A. In particular, from our theorem we see that (ii), (vi) and (viii) assumed in Theorem A are not necessary, and (i) can be replaced by (1) for the global stability of the trivial solution of (1.1).

3 Examples

In this section, we consider certain examples which are particular cases of (1.1).

Example 1 Consider the equation

$$\dddot{x} + [(\sin x)\dot{x} + (\dot{x})^2 + e^{\dot{x}\ddot{x}} + 2]\ddot{x} + (\dot{x})^3 + \dot{x} + \frac{x}{1 + x^2} = 0. \tag{3.1}$$

(3.1) is in the form of (1.1) with

$$\psi(x, y, z) = (\sin x)y + y^2 + e^{yz} + 2, \qquad f(x, y) = y^3 + y + \frac{x}{1 + x^2}.$$

With $a = 2$, $b = 1$, $c = 1$, we observe that

$$\left[f(x,y) - f(x,0) - \int_0^y \psi_x(x,v,0)v dv\right] y = \left[y^3 + y - \frac{1}{3}(\cos x)y^3\right] y$$
$$> y^2 \frac{1-x^2}{(1+x^2)^2} = y^2 f'(x,0), \quad \text{for } y \neq 0.$$

Then it is easy to check all the hypotheses in Theorem are satisfied and so the trivial solution of (3.1) is globally asymptotically stable.

Example 2 Consider the equation

$$\dddot{x} + [\ln(1+x^2) + e^{\dot{x}\ddot{x}} + 2] + \frac{x}{1+x^2}(1+(\dot{x})^2) + \dot{x} + \frac{1}{3}(\dot{x})^3 = 0. \tag{3.2}$$

(3.2) is in the form (1.1) with

$$\psi(x,y,z) = \ln(1+x^2) + e^{yz} + 2, \qquad f(x,y) = \frac{x}{1+x^2}(1+y^2) + y + \frac{1}{3}y^3.$$

With $a = 2$, $b = 1$, $c = 1$, we observe that

$$\left[f(x,y) - f(x,0) - \int_0^y \psi_x(x,v,0)v\, dv\right] y = \left[y + \frac{1}{3}y^3\right] y$$
$$> y^2 \frac{(1-x^2)}{(1+x^2)^2} = y^2 f'(x,0), \quad \text{for } y \neq 0.$$

Then it is easy to check all the hypotheses in Theorem are satisfied and so the trivial solution of (3.2) is globally asymptotically stable.

References

[1] Barbashin, E. A.: Lyapunov Functions. *Nauka, Moscow*, 1970.

[2] Ezeilo, J. O. C.: *On the stability of solutions of certain differential equations of the third order.* Quart. J. Math. Oxford Ser. **11** (1960), 64–69.

[3] Goldwyn, M., Narendra, S.: Stability of Certain Nonlinear Differential Equation Using the Second Method of Lyapunov. *Craft Lab. Harvard Univ., Cambridge, MA*, 1963, pp. 1–14.

[4] Kraovkii, N. N.: Stability of Motion. *Stanford University Press, Stanford, CA*, 1963.

[5] Ogurtsov, A. I.: *On the stability of the solutions of some nonlinear differential equations of the third and forth order.* Ivz. Vyssh. Uchebn. Zaved. Mat. **10** (1959), 200–209.

[6] Qin, Y., Wan, M., Wang, L.: Theory and Applications of Stability of Motions. *Academic Press, Beijing*, 1981.

[7] Qian, C.: *On global stability of third-order nonlinear differential equations.* Nonlinear Analysis **42** (2000), 651–661.

[8] Qian, C.: *Asymptotic behavior of a third-order nonlinear differential equation.* J. Math. Anal. Appl. **284** (2003), 191–205.

[9] Omeike, M. O.: *Further results on global stability of third-order nonlinear differential equations.* Nonlinear Analysis **67** (2007) 3394–3400.

[10] Tunc, C.: *Global stability of solutions of certain third-order nonlinear differential equations.* Panamer. Math. J. **14**, 4 (2004), 31–35.

[11] Tunc, C.: *On the asymptotic behavior of solutions of certain third-order nonlinear differential equations.* J. Appl. Math. Stoch. Anal. **1** (2005), 29–35.

[12] Reissig, R, Sansone, G., Conti, R.: Nonlinear Differential Equations of Higher Order. *Noordhoff Inter. Pub., Leyden*, 1974.

[13] Shimanov, S. N.: *On the stability of the solution of a nonlinear equation of the third order.* Prikl. Mat. Mekh. **17** (1953), 369–372.

[14] Wang, L., Wang, M.: *Analysis of construction of Lyapunov functions of third-order nonlinear systems.* Acta Math. Appl. Sin. **6** (1983), 309–323.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **47** (2008) 129–138

On Weakly and Pseudo Concircular Symmetric Structures on a Riemannian Manifold

Füsun ÖZEN ZENGIN[1], Sezgin ALTAY DEMIRBAG[2]

Istanbul Technical University, Faculty of Sciences and Letters
Department of Mathematics, Maslak-Istanbul, Turkey
[1] *e-mail: fozen@itu.edu.tr*
[2] *e-mail: saltay@itu.edu.tr*

(Received February 27, 2008)

Abstract

In this paper, we examine the properties of hypersurfaces of weakly and pseudo concircular symmetric manifolds and we give an example for these manifolds.

Key words: Weakly symmetric manifold, pseudo symmetric manifold, weakly and pseudo symmetric concircular manifold, totally umbilical, totally geodesic, mean curvature, scalar curvature.

2000 Mathematics Subject Classification: 53B20, 53B15

1 Introduction

Firstly, Tamassy and Binh introduced weakly symmetric manifolds, [1].

A non-flat Riemannian manifold (M_n, g), $(n > 2)$ whose the curvature tensor satisfies the following relation is called weakly symmetric

$$\nabla_l R_{hijk} = A_l R_{hijk} + B_h R_{lijk} + D_i R_{hljk} + E_j R_{hilk} + F_k R_{hijl} \tag{1.1}$$

where A, B, D, E, F are non-zero 1-forms and ∇ denotes the covariant differentiation with respect to the metric tensor of the manifold. These 1-forms are called the associated 1-forms of the manifold and an n-dimensional manifold of this kind is denoted by $(WS)_n$. It may be mentioned in this connection that although the definition of a $(WS)_n$ is similar to that of a generalized pseudo-symmetric space studied by Chaki and Mondal, [2], the defining condition of a $(WS)_n$ is weaker than that of a generalized pseudo-symmetric manifold. De and Bandyopadhyay, [3], proved that 1-forms of $(WS)_n$ can not be all different. Then the equation (1.1) reduces to the form

$$\nabla_l R_{hijk} = A_l R_{hijk} + B_h R_{lijk} + B_i R_{hljk} + D_j R_{hilk} + D_k R_{hijl} \tag{1.2}$$

Let us consider a subspace V_m immersed in a Riemannian manifold V_n whose parametric representation is $u^\lambda = u^\lambda(u^1, u^2, \dots, u^m)$ where (u^λ) and (u^i) $(i, j, k, \dots = 1, 2, \dots, m)$ denote the coordinate systems of V_n and V_m, respectively. A conformal transformation $\bar{g}_{ij} = \rho^2 g_{ij}$ of the fundamental tensor of V_n, being a concircular one with the function ρ satisfying the equations

$$\rho_{ij} = \nabla_j \rho_i - \rho_i \rho_j + \frac{1}{2} g^{\alpha\beta} \rho_\alpha \rho_\beta g_{ij} = \phi g_{ij}, \qquad \rho_j = \frac{\partial}{\partial u^j} \ln \rho \tag{1.3}$$

this transformation is called concircular transformation where ϕ is a function of u^i.

The present paper deals with non-concircular flat Riemannian manifold (M_n, g) whose concircular curvature tensor Z_{hijk} satisfies the condition $(n > 2)$

$$\nabla_l Z_{hijk} = A_l Z_{hijk} + B_h Z_{lijk} + D_i Z_{hljk} + E_j Z_{hilk} + F_k Z_{hijl}$$

where

$$Z_{hijk} = R_{hijk} - \frac{R}{n(n-1)} (g_{hk} g_{ij} - g_{hj} g_{ik})$$

R_{hijk} is the curvature tensor and R is the scalar curvature. Such a manifold will be called a weakly concircular symmetric manifold and denoted by $(WZS)_n$, [4]. It was shown that, in [5], Z^h_{ijk} is invariant under a concircular transformation.

Desa and Amur studied the concircular recurrent Riemannian manifold, [6]. The authors proved that the defining condition of a $(WZS)_n$ can always be expressed in the following form, [4]

$$\nabla_l Z_{hijk} = A_l Z_{hijk} + B_h Z_{lijk} + B_i Z_{hljk} + D_j Z_{hilk} + D_k Z_{hijl} \tag{1.4}$$

where A, B, D 1-forms (non-zero simultaneously).

From the first Bianchi identity, we get

$$R_{hijk} + R_{hjki} + R_{hkij} = 0 \tag{1.5}$$

The second Bianchi identity for a Riemannian manifold is

$$\nabla_s R_{hijk} + \nabla_j R_{hiks} + \nabla_k R_{hisj} = 0 \tag{1.6}$$

Let $(\bar{M}, \bar{g})$ be an $(n+1)$-dimensional Riemannian manifold covered by a system of coordinate neighborhoods $\{U, y^\alpha\}$. Let (M, g) be a hypersurface of $(\bar{M}, \bar{g})$ defined via a system of parametric equation $y^\alpha = y^\alpha(x^i)$, where Greek indices take the values $1, 2, \ldots, n+1$ and Latin indices take the values $1, 2, \ldots, n$ a locally coordinate system. Then, we have

$$g_{ij} = \bar{g}_{\alpha\beta} y_i^\alpha y_j^\beta \tag{1.7}$$

Let n^α be a local unit normal to (M, g). Thus, we obtain $\bar{g}_{\alpha\beta} n^\alpha y_i^\beta = 0$, $g_{\alpha\beta} n^\alpha n^\beta = 1$ and it is easily seen that there are the following conditions between the contrary metric tensors of the hypersurface (M, g) and $(\bar{M}, \bar{g})$

$$g^{\alpha\beta} = g^{ij} y_i^\alpha y_j^\beta + n^\alpha n^\beta, \;\; y_i^\alpha = \frac{\partial y^\alpha}{\partial x^i}, \;\; (i, j = 1, 2, \ldots, n; \; \alpha = \beta = 1, 2, \ldots, n+1) \tag{1.8}$$

A point of a hypersurface, at which the principal directions of the curvature are indeterminate, is called an umbilical point. In order that the lines of curvature may be indeterminate at every point of the hypersurface, it is necessary and sufficient that $\Omega_{ij} = \omega g_{ij}$, where ω is an invariant. According to [7],

$$M = \Omega_{ij} g^{ij} = n\omega \tag{1.9}$$

where the scalar M is called the mean curvature of such a hypersurface, so that the conditions for indeterminate lines of curvature are expressible as

$$\Omega_{ij} = \frac{M}{n} g_{ij} \tag{1.10}$$

If all the geodesics of a hypersurface (M, g) are also geodesics of $(\bar{M}, \bar{g})$, the former is called a totally geodesic hypersurface of the latter. Such hypersurfaces are generalizations of planes in ordinary space. A necessary and sufficient condition that (M, g) be a totally geodesic hypersurface is that the normal curvature should vanish for all directions in (M, g), and at every point. This requires

$$\Omega_{ij} = 0 \tag{1.11}$$

Consequently,

$$M = 0 \tag{1.12}$$

and (1.10) is satisfied.

The structure equations of Gauss and Mainardi–Codazzi, [8]

$$R_{ijkl} = \bar{R}_{\alpha\beta\gamma\theta} B_{ijkl}^{\alpha\beta\gamma\theta} + \Omega_{ijkl}$$

and

$$\nabla_k \Omega_{ij} - \nabla_j \Omega_{ik} + \bar{R}_{\beta\gamma\delta\theta} n^\beta B_{ijk}^{\gamma\delta\theta} = 0$$

where $\Omega_{ijkl} = \Omega_{lj}\Omega_{ik} - \Omega_{il}\Omega_{jk}$.

From (1.9), the above equations reduce to the following forms

$$R_{ijkl} = \bar{R}_{\alpha\beta\gamma\theta} B_{ijkl}^{\alpha\beta\gamma\theta} + \frac{M^2}{n^2}(g_{lj}g_{ik} - g_{li}g_{jk}) \tag{1.13}$$

and

$$\bar{R}_{\alpha\gamma\delta\theta} n^{\alpha} B^{\gamma\delta\theta}_{ijk} = \frac{1}{n}(g_{ik}\nabla_j M - g_{ij}\nabla_k M) \tag{1.14}$$

respectively, where R_{ijkl} and $\bar{R}_{\alpha\beta\gamma\theta}$ are the curvature tensors (M, g) and $(\bar{M}, \bar{g})$, and $B^{\alpha\beta\gamma\theta}_{ijkl} = B^{\alpha}_i B^{\beta}_j B^{\gamma}_k B^{\theta}_l$, $B^{\alpha}_i = y^{\alpha}_i$.

From the Gauss equation, we get

$$\bar{R} = R + 2\bar{R}_{\alpha\beta} n^{\alpha} n^{\beta} - \Omega_{ijkl} g^{il} g^{jk} \tag{1.15}$$

The concircular curvature tensors of (M, g) and $(\bar{M}, \bar{g})$ can be written in the form

$$Z_{hijk} = R_{hijk} + \frac{R}{n(n-1)} G_{hijk} \tag{1.16}$$

and

$$\bar{Z}_{\alpha\beta\gamma\theta} = \bar{R}_{\alpha\beta\gamma\theta} + \frac{\bar{R}}{n(n+1)} G_{\alpha\beta\gamma\theta} \tag{1.17}$$

where $G_{hijk} = g_{hj} g_{ik} - g_{hk} g_{ij}$ and $G_{\alpha\beta\gamma\theta} = \bar{g}_{\alpha\gamma}\bar{g}_{\beta\theta} - \bar{g}_{\alpha\theta}\bar{g}_{\beta\gamma}$. On account of (1.7), (1.13), (1.16) and (1.17), we get

$$Z_{hijk} = \bar{Z}_{\alpha\beta\gamma\theta} B^{\alpha\beta\gamma\theta}_{hijk} + \frac{M^2}{n^2} G_{hijk} + \frac{1}{n}(\frac{R}{n-1} - \frac{\bar{R}}{n+1}) G_{hijk} \tag{1.18}$$

2 Totally umbilical hypersurface of a weakly concircular symmetric manifold

Now, we consider an $(n+1)$-dimensional weakly concircular symmetric Riemannian manifold and we denote this manifold by $(WZS)_{n+1}$. For a $(WZS)_{n+1}$, we have

$$\nabla_e \bar{Z}_{abcd} = A_e \bar{Z}_{abcd} + B_a \bar{Z}_{ebcd} + B_b \bar{Z}_{aecd} + D_c \bar{Z}_{abed} + D_d \bar{Z}_{abce} \tag{2.1}$$

Using (1.17), we obtain

$$\bar{Z}_{abcd} n^a B^{bcd}_{ijk} = \bar{R}_{abcd} n^a B^{bcd}_{ijk} \tag{2.2}$$

We assume that the scalar curvature of $(WZS)_n$ is not constant and $(WZS)_n$ is a totally umbilical hypersurface. In this case, we find that

$$\begin{aligned}
\nabla_s Z_{hijk} &= A_s \bar{Z}_{abcd} B^{abcd}_{hijk} + B_h \bar{Z}_{ebcd} B^{ebcd}_{sijk} + B_i \bar{Z}_{aecd} B^{aecd}_{hsjk} \\
&+ D_j \bar{Z}_{abed} B^{abed}_{hisk} + D_k \bar{Z}_{abce} B^{abce}_{hijs} + \frac{1}{n^2} G_{hijk} \nabla_s M^2 \\
&+ \frac{1}{n} G_{hijk} \nabla_s \Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big) + \frac{M}{n}\Big(g_{hs}\bar{R}_{abcd} B^{bcd}_{ijk} n^a + g_{is}\bar{R}_{badc} B^{adc}_{hkj} n^b \\
&+ g_{js}\bar{R}_{cdab} B^{dab}_{khi} n^c + g_{ks}\bar{R}_{dcba} B^{cba}_{jih} n^d\Big)
\end{aligned} \tag{2.3}$$

By the aid of the Gauss equation, (2.3) can be written as

$$\begin{aligned}\nabla_s Z_{hijk} &= A_s\Big(Z_{hijk} - \frac{M^2}{n^2}G_{hijk} - \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)G_{hijk}\Big)\\ &+ B_h\Big(Z_{sijk} - \frac{M^2}{n^2}G_{sijk} - \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)G_{sijk}\Big)\\ &+ B_i\Big(Z_{hsjk} - \frac{M^2}{n^2}G_{hsjk} - \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)G_{hsjk}\Big)\\ &+ D_j\Big(Z_{hisk} - \frac{M^2}{n^2}G_{hisk} - \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)G_{hisk}\Big)\\ &+ D_k\Big(Z_{hijs} - \frac{M^2}{n^2}G_{hijs} - \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)G_{hijs}\Big)\\ &+ \frac{1}{n^2}G_{hijk}\nabla_s M^2 + \frac{1}{n}G_{hijk}\nabla_s\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\\ &+ \frac{M}{n^2}[(g_{hs}g_{ik} - g_{is}g_{hk})\nabla_j M + (g_{is}g_{hj} - g_{ij}g_{hs})\nabla_k M\\ &+ (g_{js}g_{ik} - g_{ij}g_{sk})\nabla_h M + (g_{ks}g_{hj} - g_{js}g_{hk})\nabla_i M] \end{aligned} \tag{2.4}$$

Now, we suppose that (M, g) is $(WZS)_n$.
By the aid of (1.4) and (2.4), we have

$$\begin{aligned}&\Big[\frac{M^2}{n^2} + \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\Big](A_s G_{hijk} + B_h G_{sijk} + B_i G_{hsjk} + D_j G_{hisk} + D_k G_{hijs})\\ &\quad - G_{hijk}\nabla_s\left(\frac{M^2}{n^2} + \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\right)\\ &- \frac{M}{n^2}(G_{hisk}\nabla_j M + G_{ihsj}\nabla_k M + G_{sijk}\nabla_h M + G_{kjsh}\nabla_i M) = 0\end{aligned} \tag{2.5}$$

Multiplying (2.5) by $g^{hk}g^{ij}$, we can obtain

$$\begin{aligned}&\left(\frac{M^2}{n^2} + \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\right)(2B_s + 2D_s + nA_s)\\ &- \frac{(n+2)}{n^2}\nabla_s M^2 - \nabla_s\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big) = 0\end{aligned} \tag{2.6}$$

Similarly, multiplying (2.5) by $g^{ik}g^{hs}$, it is easily obtained that

$$\begin{aligned}&\left(\frac{M^2}{n^2} + \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\right)(B_s + A_s + (n-1)D_s)\\ &- \frac{(n+2)}{2n^2}\nabla_s M^2 - \frac{1}{n}\nabla_s\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big) = 0\end{aligned} \tag{2.7}$$

Let us suppose that

$$R = (1 - \frac{2}{n+1})\bar{R} \tag{2.8}$$

where the scalar curvature R is not constant.

From (2.6) and (2.7), we get

$$A_s = 2D_s \quad \text{or} \quad M = 0 \tag{2.9}$$

We assume that $A_s = 2D_s$. Transvecting (1.4) with g^{lk} and g^{ij}, we get

$$g^{ks}\nabla_s G_{hk} = (A_k - B_k + D_k)G_{hs}g^{sk} \tag{2.10}$$

where $G_{hk} = R_{hk} - \frac{R}{n}g_{hk}$ $(n > 2)$ is the Einstein tensor.

Similarly, transvecting (1.4) with g^{hk} and g^{ij}, we have

$$(B_k + D_k)G_{hs}g^{sk} = 0 \tag{2.11}$$

Hence, using the equations $(2.9)_1$ and (2.10), it can be obtained that

$$(A_k + 2B_k)G_{hs}g^{sk} = 0 \tag{2.12}$$

Now, multiplying the equation (1.4) by g^{hl} and g^{ij} and using the result $\nabla_s R^s_h = \frac{1}{2}\nabla_h R$, we obtain $R \equiv$ const. In the beginning, we suppose that $R \neq const.$ Thus, $A_s \neq 2D_s$. From (2.9), we have $M = 0$, i.e., the hypersurface is totally geodesic. Thus, we can state the following theorem:

Theorem 2.1 *In the totally umbilical hypersurface* $(WZS)_n$ *of* $(WZS)_{n+1}$, *if the expression* $R = (1 - \frac{2}{n+1})\bar{R}$, $(R \neq const.)$ *is satisfied then the hypersurface is totally geodesic.*

Theorem 2.2 *If the totally umbilical hypersurface* $(WZS)_n$ *of a* $(WZS)_{n+1}$ *satisfies the condition* $\frac{R}{n-1} - \frac{\bar{R}}{n+1} = c$ $(c < 0,\ const.)$ *then either the mean curvature or the scalar curvature of this hypersurface is constant.*

Proof We assume that the totally umbilical hypersurface $(WZS)_n$ of $(WZS)_{n+1}$ satisfies the condition

$$-\frac{\bar{R}}{n+1} + \frac{R}{n-1} = c \tag{2.13}$$

From (2.5) and (2.13), we obtain

$$\begin{aligned}\Big(\frac{M^2}{n^2} + \frac{c}{n}\Big)&(A_sG_{hijk} + B_hG_{sijk} + B_iG_{hsjk} + D_jG_{hisk} + D_kG_{hijs})\\ &- \frac{1}{n^2}G_{hijk}\nabla_s M^2 - \frac{M}{n^2}(G_{hisk}\nabla_j M\\ &+ G_{ihsj}\nabla_k M + G_{sijk}\nabla_h M + G_{kjsh}\nabla_i M) = 0\end{aligned} \tag{2.14}$$

Multiplying (2.14) by $g^{hk}g^{ij}$, we find that

$$\Big(\frac{M^2}{n^2} + \frac{c}{n}\Big)(2B_s + 2D_s + nA_s) - \frac{(n+2)}{n^2}\nabla_s M^2 = 0 \tag{2.15}$$

Similarly, multiplying (2.14) by $g^{ik}g^{hs}$, we can easily obtain that

$$\Big(\frac{M^2}{n^2}+\frac{c}{n}\Big)(B_s+A_s+(n-1)D_s)-\frac{(n+2)}{2n^2}\nabla_s M^2=0 \qquad (2.16)$$

Using (2.15) and (2.16), we get

$$M^2=-cn \qquad \text{or} \qquad A_s=2D_s \qquad (2.17)$$

On the other hand, from (1.4), we have

$$\nabla_l Z_{hijk}=A_l Z_{hijk}+B_h Z_{lijk}+B_i Z_{hljk}+D_j Z_{hilk}+D_k Z_{hijl} \qquad (2.18)$$

Permutating j,k and l by cyclic in (2.18), adding the three equations and using the expression (1.5) and the first Bianchi Identity, we obtain

$$\begin{aligned}&(A_l-2D_l)Z_{hijk}+(A_j-2D_j)Z_{hikl}+(A_k-2D_k)Z_{hilj}\\&-\frac{1}{n(n-1)}(G_{hijk}\nabla_l R+G_{hikl}\nabla_j R+G_{hilj}\nabla_k R)\end{aligned} \qquad (2.19)$$

Transvecting (2.19) with $g^{ij}g^{hk}$, we can obtain

$$2(A_k-2D_k)g^{hk}G_{hl}=\frac{(n-2)}{n}\nabla_l R \qquad (2.20)$$

If $A_k=2D_k$, from (2.20), then we say that the scalar curvature of this hypersurface is constant. If $A_k\neq 2D_k$, from (2.17), the mean curvature of this hypersurface must be constant. If $c=0$ then it is clear that this hypersurface is totally geodesic. Thus, the proof is completed. □

Theorem 2.3 *If a totally geodesic hypersurface of a* $(WZS)_{n+1}$ *satisfies the condition* $R=(1-\frac{2}{n+1})\bar{R}$ *then this hypersurface is* $(WZS)_n$.

Proof From (1.4) and (2.4), the proof is easily seen that.

3 Totally umbilical hypersurface of a pseudo concircular symmetric manifold

We consider a non-concircular flat Riemannian manifold (M,g) whose concircular curvature tensor Z_{hijk} satisfies the condition

$$\nabla_l Z_{hijk}=2\lambda_l Z_{hijk}+\lambda_h Z_{lijk}+\lambda_i Z_{hljk}+\lambda_j Z_{hilk}+\lambda_k Z_{hijl} \qquad (3.1)$$

where λ_l is a non-zero covariant vector. Such a manifold will be called a pseudo-concircular symmetric manifold and denoted by $(PZS)_n$. Permutating j,k,l by cyclic in (3.1), we obtain the following equations

$$\nabla_j Z_{hikl}=2\lambda_j Z_{hikl}+\lambda_h Z_{jikl}+\lambda_i Z_{hjkl}+\lambda_k Z_{hijl}+\lambda_l Z_{hikj} \qquad (3.2)$$

and

$$\nabla_k Z_{hilj} = 2\lambda_k Z_{hilj} + \lambda_h Z_{kilj} + \lambda_i Z_{hklj} + \lambda_l Z_{hikj} + \lambda_j Z_{hilk} \tag{3.3}$$

Adding the equations (3.1), (3.2) and (3.3) and by using the first and the second Bianchi identities, it is obtained that

$$G_{hijk}\nabla_l R + G_{hikl}\nabla_j R + G_{hilj}\nabla_k R = 0 \tag{3.4}$$

Transvecting (3.4) with $g^{hk}g^{ij}$, we get $(1-n)(2-n)\nabla_l R = 0$.

Since $n > 2$, we find that the scalar curvature of the hypersurface is constant. Now, we can state the following theorem:

Theorem 3.1 *The scalar curvature of a pseudo concircular symmetric manifold is constant.*

Theorem 3.2 *Let us suppose that a hypersurface $(PZS)_n$ of a pseudo concircular symmetric manifold $(PZS)_{n+1}$ be totally umbilical. Then the scalar curvature of $(PZS)_{n+1}$ is constant.*

Proof Taking the relation $\frac{A_s}{2} = B_s = D_s = \lambda_s$ in (2.3), (2.4) and (2.5) and using the equation (3.1), we get

$$\begin{aligned}\Big(\frac{M^2}{n^2} + \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\Big)&(2\lambda_s G_{hijk} + \lambda_i G_{hsjk} + \lambda_j G_{hisk} + \lambda_k G_{hijs} + \lambda_h G_{sijk}) \\ &- \frac{1}{n^2} G_{hijk}\nabla_s M^2 - \frac{1}{n} G_{hijk}\nabla_s\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big) \\ &- \frac{M}{n^2}(G_{hisk}\nabla_j M + G_{ihsj}\nabla_k M + G_{sijk}\nabla_h M + G_{kjsh}\nabla_i M) = 0 \end{aligned}\tag{3.5}$$

Multiplying (3.5) by $g^{hk}g^{ij}$ and $g^{ik}g^{hs}$, respectively, we obtain

$$\begin{aligned}\Big(\frac{M^2}{n^2} + \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\Big)&2\lambda_s(2+n) - \frac{(n+2)}{n^2}\nabla_s M^2 \\ &- \nabla_s\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big) = 0\end{aligned}\tag{3.6}$$

and

$$\begin{aligned}\Big(\frac{M^2}{n^2} + \frac{1}{n}\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big)\Big)&\lambda_s(2+n) - \frac{(n+2)}{2n^2}\nabla_s M^2 \\ &- \frac{1}{n}\nabla_s\Big(\frac{R}{n-1} - \frac{\bar{R}}{n+1}\Big) = 0\end{aligned}\tag{3.7}$$

From (3.6) and (3.7), we obtain

$$-\frac{\bar{R}}{n+1} + \frac{R}{n-1} = c \tag{3.8}$$

where c is a positive constant. By using Theorem 3.1, we can say that

$$\bar{R} \equiv \text{const.} \tag{3.9}$$

□

Theorem 3.3 *If a totally geodesic hypersurface of $(PZS)_{n+1}$ satisfies the condition $R = (1 - \frac{2}{n+1})\bar{R}$ then the hypersurface is $(PZS)_n$.*

Proof Let us suppose that a hypersurface of $(PZS)_{n+1}$ be totally geodesic. From the expressions (1.12) and (2.4) and the condition $\frac{A_s}{2} = B_s = D_s = \lambda_s$, the proof is clear. □

4 An example of a $(WZS)_n$

In this section, we want to construct a $(WZS)_n$ spaces. On the coordinate space R^n (with coordinates $x^1, x^2, \ldots, x^n$), we define a Riemannian space V^n and calculate the components of the curvature tensor and its covariant derivative.

Let each Latin index run over $1, 2, \ldots, n$ and each Greek index over $2, 3, \ldots, n-1$. We define a Riemannian metric on R^n $(n > 3)$ by the formula

$$ds^2 = \phi(dx^1)^2 + k_{\alpha\beta}dx^\alpha dx^\beta + 2dx^1dx^n \tag{4.1}$$

where $[k_{\alpha\beta}]$ is a symmetric and non-singular matrix consisting of constants and ϕ is a function of $(x^1, x^2, \ldots, x^{n-1})$ and independent of x^n. In the metric considered, the only non-vanishing components of the curvature tensor, [9]

$$R_{1\alpha\beta 1} = \frac{1}{2}\phi_{.\alpha\beta} \tag{4.2}$$

where "." denotes the partial differentiation with respect to the coordinates and $k^{\alpha\beta}$ are the elements of the matrix inverse to $[k^{\alpha\beta}]$.

We consider V_n and

$$\phi = f(x^1)(V_{\alpha\beta}x^\alpha x^\beta \cos g(x^1) + w_{\alpha\beta}x^\alpha x^\beta \sin g(x^1) + k_{\alpha\beta}x^\alpha x^\beta h(x^1))$$

where f, g, h are functions of x^1 only and the matrices $[w_{\alpha\beta}]$, $[V_{\alpha\beta}]$ and $[k_{\alpha\beta}]$ are the form

$$w_{\alpha\beta} = -1 \ \text{ for } \alpha = \beta \qquad \text{and} \qquad w_{\alpha\beta} = 0 \ \text{ for } \alpha \neq \beta \tag{4.3}$$

$$V_{\alpha\beta} = 1 \ \text{ for } \alpha = \beta \qquad \text{and} \qquad V_{\alpha\beta} = 0 \ \text{ for } \alpha \neq \beta \tag{4.4}$$

and

$$k_{\alpha\beta} = \left\{ \begin{matrix} 1 & \text{for } \alpha = \beta \\ 0 & \text{otherwise} \end{matrix} \right\} \tag{4.5}$$

From (4.2), the only non-vanishing components of the concircular curvature tensor Z_{hijk} are

$$Z_{1\alpha\beta 1} = \left\{ \begin{matrix} f(\cos g - \sin g + h) & \text{for } \alpha = \beta \\ 0 & \text{for } \alpha \neq \beta \end{matrix} \right\} \tag{4.6}$$

Here, we consider

$$A_i = B_i = D_i = 0 \ \text{ for } i \neq 1 \text{ and } A_1 + B_1 + D_1 = c_1,\ c_1 \neq 0 \text{ and const.} \quad (4.7)$$

Thus, from (1.4), V_n will be $(WZS)_n$ if and only if the following relations

$$\nabla_1 Z_{1\alpha\alpha 1} = A_1 Z_{1\alpha\alpha 1} + B_1 Z_{1\alpha\alpha 1} + B_\alpha Z_{11\alpha 1} + D_\alpha Z_{1\alpha 11} + D_1 Z_{1\alpha\alpha 1} \quad (4.8)$$

$$\nabla_\alpha Z_{11\alpha 1} = A_\alpha Z_{11\alpha 1} + B_1 Z_{\alpha 1\alpha 1} + B_1 Z_{1\alpha\alpha 1} + D_\alpha Z_{11\alpha 1} + D_1 Z_{11\alpha\alpha} \quad (4.9)$$

$$\nabla_\alpha Z_{1\alpha 11} = A_\alpha Z_{1\alpha 11} + B_1 Z_{\alpha\alpha 11} + B_\alpha Z_{1\alpha 11} + D_1 Z_{1\alpha\alpha 1} + D_1 Z_{1\alpha 1\alpha} \quad (4.10)$$

Thus, using (4.8), (4.9) and (4.10), we find

$$\begin{aligned} &f'(x^1)(\cos g - \sin g + h) + f(x^1)(-g'\sin g - g'\cos g + h') \\ &\quad = (A_1 + B_1 + D_1) f(x^1)(\cos g - \sin g + h). \end{aligned} \quad (4.11)$$

By the aid of (4.11), we get

$$f(\cos g - \sin g + h) = c_2 e^{(A_1+B_1+D_1)x^1}, \qquad c_2 > 0. \quad (4.12)$$

So, the n-dimensional weakly concircular recurrent Riemannian manifold has the metric of the form

$$ds^2 = \phi (dx^1)^2 + k_{\alpha\beta} dx^\alpha dx^\beta + 2dx^1 dx^n,$$
$$\phi = c_2 e^{c_1 x^1} \sum_{k=2}^{n-1} (x^k)^2.$$

References

[1] Tamassy, L., Binh, T. Q.: *On weakly symmetric and weakly pseudo projective symmetric Riemannian manifolds.* Coll. Math. Soc. J. Bolyai **50** (1989), 663–670.

[2] Chaki, M. C., Mondal, S. P.: *On generalised pseudo symmetric manifolds.* Publ. Math. Debrecen **51**, 1-2 (1997), 35–42.

[3] De, U. C., Bandyopadhyay: *On weakly symmetric Riemannian spaces.* Publ. Math. Debrecen **54** (1999), 377–381.

[4] Ozen, F., Altay, S.: *On weakly concircular symmetric spaces.* Math. Pannonica **16**, 1 (2005), 29–38.

[5] Yano, K.: Concircular Geometry I. (1940) 195–200.

[6] Desa, P., Amur, K.: *On w-recurrent spaces.* Tensor **29** (1975), 98–102.

[7] Chen, B. Y.: Geometry of Submanifolds. *Marcel-Deker, New York*, 1973.

[8] Eisenhart, L. P.: Riemannian Geometry. *Princeton University Press*, 1949.

[9] Roter, W.: *On conformally symmetric Ricci-recurrent space.* Coll. Math. **31** (1974), 87–96.

ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
MATHEMATICA 47 (2008)

Published by the Palacký University Olomouc, Křížkovského 8, 771 47 Olomouc

First edition

ISBN 978-80-244-2121-6
ISSN 0231-9721