ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
2007

ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
MATHEMATICA 46 (2007)

MATHEMATICA 46

PALACKÝ UNIVERSITY OLOMOUC
OLOMOUC 2006

The autor are responsible for the translation of the papers.

ISBN 978-80-244-1819-3
ISSN 0231-9721

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007)

CONTENTS

Dušan BEDNAŘÍK, Karel PASTOR: Second-order sufficient condition for $\widetilde{\ell}$-stable functions 7
Ivan CHAJDA, Radomír HALAŠ: Varieties satisfying the triangular scheme need not be congruence distributive 19
Ivan CHAJDA, Miroslav KOLAŘÍK: Ideals, congruences and annihilators on nearlattices 25
Marina F. GREBENYUK, Josef MIKEŠ: Equipping distributions for linear distribution 35
Karel HRON: Estimation of the first order parameters in the twoepoch linear model 43
Pavla KUNDEROVÁ: Eliminating transformations for nuisance parameters in linear regression models with type I constraints 51
M. O. OMEIKE: Ultimate boundedness results for a certain third order nonlinear matrix differential equations 65
Irena RACHŮNKOVÁ, Jan TOMEČEK: Singular nonlinear problem for ordinary differential equation of the second order 75
Daniel WAGNER, Stefan WOPPERER: Bol-loops of order $3 \cdot 2^n$ 85

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 7–18

Second-order Sufficient Condition for $\tilde{\ell}$-stable Functions*

Dušan BEDNAŘÍK [1], Karel PASTOR [2]

[1] *Department of Mathematics, University of Hradec Králové,
Rokitanského 62, 500 03 Hradec Králové, Czech Republic
e-mail: dbednarik@seznam.cz*

[2] *Department of Mathematical Analysis and Applications of Mathematics
Faculty of Science, Palacký University
Tř. Svobody 26, 771 46 Olomouc, Czech Republic
e-mail: pastor@inf.upol.cz*

(Received December 7, 2006)

Abstract

The aim of our article is to present a proof of the existence of local minimizer in the classical optimality problem without constraints under weaker assumptions in comparisons with common statements of the result. In addition we will provide rather elementary and self-contained proof of that result.

Key words: Second-order derivative; $C^{1,1}$ function; stable function; isolated minimizer of order 2.

2000 Mathematics Subject Classification: 49K10, 26B05

1 Introduction

Past all doubt it is very important to be able to find the minimum or maximum of a function. Recall for example that by J. von Neumann every physical system tends to have its minimum of internal energy.

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

From the basic course of mathematical analysis we know the second-order condition for strict local minimum (see [Zo]) given in Theorem 1 below. We will use the following notation and terminology.

We denote by $f'(x;h)$, i.e.

$$f'(x;h) = \lim_{t \downarrow 0} \frac{f(x+th) - f(x)}{t},$$

the first-order directional derivative of $f : \mathbb{R}^N \to \mathbb{R}$ at $x \in \mathbb{R}^N$ in direction $h \in \mathbb{R}^N$

If there exists $f'(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R})$ (it means $f'(x)$ is an element of the set of all continuous linear mappings from $\mathbb{R}^N$ to $\mathbb{R}$) such that $f'(x)h = f'(x;h)$ for every $h \in S_{\mathbb{R}}^N = \{y \in \mathbb{R}^N; \|y\| = 1\}$, and the limit in the definition of $f'(x)h$ is uniform for $h \in S_{\mathbb{R}}^N$, then we say that $f : \mathbb{R}^N \to \mathbb{R}$ is *Fréchet differentiable at* $x \in \mathbb{R}^N$.

Further,

$$f''(x;u,v) = \lim_{t \downarrow 0} \frac{f'(x+tu;v) - f'(x;v)}{t}$$

denotes the second-order directional derivative of f at x in direction $(u,v) \in \mathbb{R}^N$.

We say that $f : \mathbb{R}^N \to \mathbb{R}$ is a C^2 function near $x \in \mathbb{R}^N$ if it has continuous second-order partial derivatives on some neighbourhood of x.

Analogously, we will say that a function $f : \mathbb{R}^N \to \mathbb{R}$ satisfies a p-property near $x \in \mathbb{R}^N$ if that p-property holds on some neighbourhood of x.

Recall that $x \in \mathbb{R}^N$ is an isolated minimizer of order 2 for a function $f : \mathbb{R}^N \to \mathbb{R}$ if there are a neighbourhood U of x and $A > 0$ satisfying $f(y) \geq f(x) + A\|y - x\|^2$ for every $y \in U$. We notice that each isolated minimizer of order 2 is a strict local minimizer.

Theorem 1 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a C^2 function near $x \in \mathbb{R}^N$. If $f'(x) = 0$, and*

$$f''(x;h,h) > 0,$$

for every $h \in S_{\mathbb{R}}^N$, then x is an isolated minimizer of order 2 for f.

Since in some problems of applied mathematics—as for example in variational inequalities, semi-infinite programming, penalty functions, proximal point methods, iterated local minimization by decomposition or augmented Lagrangian—differentiable functions which are not twice differentiable appear (see e.g. [HSN, KT, TR, Q1, Q2]), it was studied the following class of functions.

We say that $f : \mathbb{R}^N \to \mathbb{R}$ is a $C^{1,1}$ *function* near $x \in \mathbb{R}^N$ if it is differentiable on some neighbourhood of x and its derivative $f'(\cdot)$ is Lipschitz there.

It is clear that the class of $C^{1,1}$ functions includes the class of C^2 functions. On the other hand, considering a function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) = \int_0^x |t| dt, \quad \forall x \in \mathbb{R},$$

we have that $f'(x) = |x|$, for every $x \in \mathbb{R}$. It means that $f'(x)$ is Lipschitz function on $\mathbb{R}$, but f is not twice differentiable at 0.

R. Cominetti and R. Correa generalized Theorem 1 by the following way in 1990.

Theorem 2 [CC] *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a $C^{1,1}$ function near $x \in \mathbb{R}^N$. If $f'(x) = 0$, and*

$$f_\infty(x;h) := \liminf_{y\to x, t\downarrow 0} \frac{f'(y+th;h) - f'(y;h)}{t} > 0,$$

for every $h \in S^N_{\mathbb{R}}$, then x is an isolated minimizer of order 2 for f.

The second-order condition from Theorem 2 was improved by elimination of strict convergence in 2004. We used a certain derivative of the Dini type.

Theorem 3 [BP1] *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a $C^{1,1}$ function near $x \in \mathbb{R}^N$. If $f'(x) = 0$, and*

$$f''^{\ell}_D(x;h) := \liminf_{t\downarrow 0} \frac{f'(x+th;h) - f'(x;h)}{t} > 0,$$

for every $h \in S^N_{\mathbb{R}}$, then x is an isolated minimizer of order 2 for f.

I. Ginchev, A. Guerraggio and M. Rocca presented the generalization of Theorem 1 in terms of the Peano derivative in 2006.

Theorem 4 [GGR] *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a $C^{1,1}$ function near $x \in \mathbb{R}^N$. If $f'(x) = 0$, and*

$$f''^{\ell}_P(x;h) := \liminf_{t\downarrow 0} \frac{f(x+th) - f(x) - tf'(x;h)}{t^2/2} > 0,$$

for every $h \in S^N_{\mathbb{R}}$, then x is an isolated minimizer of order 2 for f.

It can be easily derived from [TR, Theorem 4] that $f''^{\ell}_P(x;h) \geq f''^{\ell}_D(x;h)$. Moreover, since the calculus with $f''^{\ell}_P(x;h)$ seems to be more comfortable than that with $f''^{\ell}_D(x;h)$, we can say that Theorem 3 lost its sense after Theorem 4. Example 1 confirms this fact.

Example 1 Let us consider a function

$$f(x) = \begin{cases} \int_0^{|x|} t(\frac{19}{20} + \sin\ln t)\,dt, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

In [BP2], we showed that f is a $C^{1,1}$ function, $f''^{\ell}_D(0;1) = \frac{19}{20} - 1 < 0$, and $f''^{\ell}_P(0;1) = f''^{\ell}_P(0;-1) = \frac{19}{20} + \frac{2}{5}(-\sqrt{5}) > 0$. Due to Theorem 4 the function f attains its strict local minimum, but Theorem 3 is not applicable.

Another result was stated by A. Ben-Tal and J. Zowe in 1985. A function $f : \mathbb{R}^N \to \mathbb{R}$ for which there exist a neighbourhood U of $x \in \mathbb{R}^N$ and $K > 0$ such that for all $y \in U$ there exists the Fréchet derivative $f'(y)$ and

$$\|f'(y) - f'(x)\| \le K\|y - x\|, \quad \forall y \in U,$$

is called *stable at x*. We note that if $f : \mathbb{R}^N \to \mathbb{R}$ is a $C^{1,1}$ function near $x \in \mathbb{R}^N$, then f is stable at x.

Theorem 5 [BZ] *Let $f : \mathbb{R}^N \to \mathbb{R}$ be Fréchet differentiable near $x \in \mathbb{R}^N$ and let f be stable at x. If $f'(x) = 0$ and*

$$f''_P(x;h) := \lim_{t\downarrow 0} \frac{f(x+th) - f(x) - tf'(x;h)}{t^2/2} > 0,$$

for every $h \in S^N_{\mathbb{R}}$, then x is a strict local minimizer 2 for f.

Finally, we note that we generalized both Theorems 4, 5 in terms of so called ℓ–stable functions as follows.

We say that a function $f : \mathbb{R}^N \to \mathbb{R}$ is *ℓ-stable at $x \in \mathbb{R}^N$* if there exist a neighbourhood U of x and $K > 0$ such that

$$|f^\ell(y;h) - f^\ell(x;h)| \le K\|y - x\|, \quad \forall y \in U, \forall h \in S^N_{\mathbb{R}},$$

where

$$f^\ell(x;h) = \liminf_{t\downarrow 0} \frac{f(x+th) - f(x)}{t}.$$

It is worth to note that the ℓ–stability at x implies the strict differentiability of the function at the point x.

Theorem 6 [BP2] *Let $f : \mathbb{R}^N \to \mathbb{R}$ be continuous near $x \in \mathbb{R}^N$ and let f be ℓ-stable at x. If $f'(x) = 0$, and*

$$f'^{\ell}_P(x;h) > 0,$$

then x is an isolated minimizer of order 2 for f.

The $C^{1,1}$ property can be generalized also in the following way.

Definition 1 We say that $f : \mathbb{R}^N \to \mathbb{R}$ is $\tilde{\ell}$-stable at $x \in \mathbb{R}^N$ if there exist a neighbourhood U of x and $K > 0$ such that

$$|f^\ell(z; z-y) - f^\ell(y; z-y)| \le K\|z - y\|^2, \quad \forall y, z \in U.$$

Remark 1 Notice that a function $f : \mathbb{R}^N \to \mathbb{R}$ is $\tilde{\ell}$-stable at $x \in \mathbb{R}^N$ if there exist a neighbourhood U of x and $K > 0$ such that

$$|f^\ell(y+th;h) - f^\ell(y;h)| \le Kt,$$

for every $h \in S^N_{\mathbb{R}}$, $y \in U$ and $t > 0$ satisfying $y + th \in U$.

Remark 2 Notice that verifying the ℓ-stability we compare $f^\ell(\cdot,\cdot)$ for points from U only with x but in all directions. Conversely, verifying the $\tilde{\ell}$-stability we compare $f^\ell(\cdot,\cdot)$ for every point from U with every point from U but only in the corresponding direction.

In [BP3], we passed the problem whether we can replace the condition to be ℓ-stable by the condition to be $\tilde{\ell}$-stable in Theorem 6. In Section 3, we will answer this question in the affirmative. Before it, in Section 2, we will examine some properties of $\tilde{\ell}$-stable functions.

2 $\tilde{\ell}$-stability

At first, we will derive that the $\tilde{\ell}$-stability together with continuity implies the Lipschitzness.

Lemma 1 [BP2, Lemma 4] *Let $f:\mathbb{R}^N\to\mathbb{R}$ be a continuous function, and let $a,b\in\mathbb{R}^N$. Then there exist $\xi_1,\xi_2\in(a,b)$ such that*

$$f^\ell(\xi_1;b-a)\le f(b)-f(a)\le f^\ell(\xi_2;b-a). \tag{1}$$

Lemma 2 *Let $f:\mathbb{R}^N\to\mathbb{R}$ be a continuous function near $x\in\mathbb{R}^N$ and let f be $\tilde{\ell}$–stable at x. Then there exists a neighbourhood V of x such that*

$$\sup_{h\in S^N_{\mathbb{R}},\ y\in V}|f^\ell(y;h)|<\infty.$$

Proof Suppose on the contrary that there are sequences $\{y_n\}_{n=1}^\infty\subset\mathbb{R}^N$, $\{h_n\}_{n=1}^\infty\subset S^N_{\mathbb{R}}$ such that $y_n\to x$ as $n\to\infty$ and

$$\lim_{n\to\infty}|f^\ell(y_n;h_n)|=\infty.$$

Without any loss of generality we can assume that either

$$\lim_{n\to\infty}f^\ell(y_n,h_n)=-\infty$$

or

$$\lim_{n\to\infty}f^\ell(y_n,h_n)=+\infty.$$

We suppose that the first case occurs (the second case can be treated by an analogous way).

Next we can assume that for certain $\gamma>0$ the condition in Definition 1 of the ℓ-stability is fulfilled on $B(x,\gamma)$, and moreover f is continuous and bounded on $B(x,\gamma)$. Let $\delta>0$ denote a constant such that for each sufficiently large $n\in\mathbb{N}$ we have : $y_n+\delta h_n\in B(x,\gamma)$.

Now, if we combine the $\tilde{\ell}$-stability and Lemma 1, for each sufficiently large $n\in\mathbb{N}$ we get $\xi_n\in(y_n,y_n+\delta h_n)$ such that the following holds :

$$\begin{aligned}f(y_n+\delta h_n)&\le f(y_n)+\delta f^\ell(\xi_n;h_n)\\&=f(y_n)+\delta[f^\ell(\xi_n;h_n)-f^\ell(y_n+\delta h_n;h_n)+f^\ell(y_n+\delta h_n;h_n)\\&\quad-f^\ell(y_n;h_n)+f^\ell(y_n;h_n)]\\&\le f(y_n)+2K\delta^2+\delta f^\ell(y_n;h_n).\end{aligned}$$

Since f is bounded on $B(x,\gamma)$ and $f^\ell(y_n;h_n) \to -\infty$, the previous inequality does not hold for any sufficiently large $n \in \mathbb{N}$,which is a contradiction. □

Proposition 1 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a continuous function near $x \in \mathbb{R}^N$ and let f be $\tilde{\ell}$-stable at x. Then f is Lipschitz near x.*

Proof Due to Lemma 2 there exists a ball $B(x,\delta)$ on which f is continuous and

$$L := \sup_{y\in B(x,\delta),h\in S_{\mathbb{R}^N}} |f^\ell(y;h)| < \infty.$$

Next by Lemma 1, for any pair of points $a, b \in B(x,\delta)$ there exists $\xi \in (a,b) \subset B(x,\delta)$ such that

$$|f(b) - f(a)| \le |f^\ell(\xi;(b-a)/\|b-a\|)|\|b-a\| \le L\|b-a\|.$$ □

Now, we will show some properties of $\tilde{\ell}$-stable functions concerning differentiability.

Lemma 3 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be L-Lipschitz near $x \in \mathbb{R}^N$. Then*

$$|f^\ell(x;h_2) - f^\ell(x;h_1)| \le L\|h_1 - h_2\|, \quad \forall h_1, h_2 \in \mathbb{R}^N.$$

Proof We consider arbitrary $h_1, h_2 \in \mathbb{R}^N$. For sufficiently small $t > 0$ it holds

$$\begin{aligned} -L\|h_2 - h_1\| &\le \frac{f(x+th_2) - f(x+th_1)}{t} \\ &= \frac{f(x+th_2) - f(x)}{t} - \frac{f(x+th_1) - f(x)}{t} \\ &= \frac{f(x+th_2) - f(x+th_1)}{t} \le L\|h_2 - h_1\|. \end{aligned}$$

Hence,

$$|f^\ell(x;h_2) - f^\ell(x;h_1)| \le L\|h_2 - h_1\|.$$ □

Proposition 2 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a continuous function near $x \in \mathbb{R}^N$ and let f be $\tilde{\ell}$-stable at x. Then for every y sufficiently close to x we have*

(1) f is directionally differentiable at y and $f'(y;-h) = -f'(y;h)$ for every $h \in S_{\mathbb{R}}^N$.

(2) the mapping $h \mapsto f'(y;h)$ from $\mathbb{R}^N$ to $\mathbb{R}$ is Lipschitz.

Proof Assume that f is continous on some neighborhood U of x and that the $\tilde{\ell}$-stability property holds on U, too. This means, that for some $K > 0$, we have:

$$|f^\ell(z;z-y) - f^\ell(y;z-y)| \le K\|z-y\|^2, \tag{2}$$

for every $z, y \in U$. Now fix $y_0 \in U$, $h_0 \in S^N_{\mathbb{R}}$ and show the existence of the directional derivative $f'(y_0; h_0)$. To do this, we will employ the following auxiliary function $g : \mathbb{R} \to \mathbb{R}$ defined as follows:

$$g(t) = f(y_0 + th_0), \quad t \in \mathbb{R}.$$

Now we can express its lower Dini directional derivative in the following form $g^\ell(t; 1) = f^\ell(y_0 + th_0; h_0)$. We have in particular that $g^\ell(0; 1) = f^\ell(y_0; h_0)$. Now we will show that the function g is $\tilde{\ell}$-stable at zero. So let us consider two arbitrary points $t_1, t_2 \in \mathbb{R}$ such that $y_0 + t_i h_0 \in U$, $i = 1, 2$. Then by (2) we have:

$$\begin{aligned} &|g^\ell(t_1; t_1 - t_2) - g^\ell(t_2; t_1 - t_2)| \\ &= |f^\ell(y_0 + t_1 h_0; (t_1 - t_2)h_0) - f^\ell(y_0 + t_2 h_0; (t_1 - t_2)h_0)| \leq K|t_1 - t_2|^2. \end{aligned} \tag{3}$$

Thus g is $\tilde{\ell}$-stable at $t = 0$. Now by the lipschitzness of f near x the function g must be Lipschitz near $t = 0$. Hence, from the Rademacher theorem it follows the existence of a sequence $\{t_n\}_{n=1}^\infty$ such that $t_n \downarrow 0$, and for every $n \in \mathbb{N}$ there exists $g'(t_n) \in \mathbb{R}$. By (3) the sequence $\{g'(t_n)\}_{n=1}^\infty$ is Cauchy and consequently there exists a limit

$$L = \lim_{n \to \infty} g'(t_n) \in \mathcal{L}(\mathbb{R}, \mathbb{R}). \tag{4}$$

In what follows, we will show that in fact $L = f'(y_0; h_0)$. This will be true if we prove that for each sequence $\{s_k\}_{k=1}^\infty$ such that $s_k \downarrow 0$ it holds

$$\left| L - \frac{f(y_0 + s_k h_0) - f(y_0)}{s_k} \right| \to 0 \quad \text{as } k \to \infty.$$

Now for every $k \in \mathbb{N}$ there is $n_k \in \mathbb{N}$ such that $t_{n_k} \in (0, s_k)$ and furthermore, by Lemma 1 and (3) there are $\xi_k, \xi'_k \in (0, s_k)$ such that

$$\begin{aligned} K|t_{n_k} - \xi_k| &\leq g'(t_{n_k}) - g^\ell(\xi_k; 1) \leq g'(t_{n_k}) - \frac{g(s_k) - g(0)}{s_k} \\ &\leq g'(t_{n_k}) - g^\ell(\xi'_k; 1) \leq K|t_{n_k} - \xi'_k|. \end{aligned}$$

This immediately implies that

$$\left| g'(t_{n_k}) - \frac{g(s_k) - g(0)}{s_k} \right| \to 0 \quad \text{as } k \to \infty. \tag{5}$$

Now since for every $k \in \mathbb{N}$ we have:

$$\left| L - \frac{f(y_0 + s_k h_0) - f(y_0)}{s_k} \right| \leq |L - g'(t_{n_k})| + \left| g'(t_{n_k}) - \frac{g(s_k) - g(0)}{s_k} \right|,$$

by (4), (5) we get the existence of limit

$$L = \lim_{k \to \infty} \frac{f(y_0 + s_k h_0) - f(x)}{s_k},$$

whenever $\{s_k\}_{k=1}^{\infty}$ is a sequence such that $s_k \downarrow 0$. Hence the following limit exists:

$$L = \lim_{s \downarrow 0} \frac{f(y_0 + sh_0) - f(y_0)}{s}.$$

The assertion (2) now follows immediately from Proposition 1 and Lemma 3. □

3 Main result

Theorem 7 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be continuous near $x \in \mathbb{R}^N$ and let f be $\tilde{\ell}$-stable at x. If $f'(x;h) = 0$ for all $h \in S_{\mathbb{R}^N}$, and*

$$f_P'^{\ell}(x;h) > 0, \quad \forall h \in S_{\mathbb{R}^N},$$

then x is an isolated minimizer of order 2 for f.

Proof Without loss of generality we can assume that $x = 0$, and $f(0) = 0$. In the proof, we will apply the mathematical induction on the dimension N. First put $N = 1$ and suppose on the contrary that the assertion does not hold. Then there exists a sequence $\{x_n\}_{n \in \mathbb{N}}$ such that $x_n \to 0$ as $n \to \infty$, and

$$f(x_n) \leq \frac{1}{n}|x_n|^2, \quad \forall n \in \mathbb{N}. \tag{6}$$

Suppose, for example, that $x_n > 0$ for every $n \in \mathbb{N}$. By the hypothesis of theorem it follows that there are $\delta > 0$, $\alpha > 0$ such that

$$\frac{f(t \cdot 1)}{t^2} \geq \alpha > 0, \quad \forall t \in (0, \delta). \tag{7}$$

By (6) and (7) we have for $n \in \mathbb{N}$ sufficiently large, that

$$\alpha \leq \frac{f(x_n)}{x_n^2} \leq \frac{1}{n},$$

hence a contradiction. Thus the assertion is true for $N = 1$.

Now let the assertion holds for $N \geq 1$ and we will prove it for $N + 1$. To do this, let us assume again that $\hat{x} = 0$ is not a minimizer of order 2 for f. This implies the existence of some sequence $\{x_n\}_{n \in \mathbb{N}}$ in $\mathbb{R}^{N+1}$ such that $x_n \to 0$ as $n \to \infty$, and

$$f(x_n) \leq \frac{1}{n}\|x_n\|^2, \quad \forall n \in \mathbb{N}. \tag{8}$$

Without loss of generality it can be assumed that for some neighbourhood $U(0)$ of zero, $x_n \in U(0)$ for every $n \in \mathbb{N}$, and on $U(0)$ the $\tilde{\ell}$-stability is valid. By the compactness of the unit sphere we can assume that for some $h_0 \in S_{N+1}$, $h_n := x_n/\|x_n\| \to h_0$ as $n \to \infty$. First suppose that for infinitely many $n \in \mathbb{N}$, x_n are contained in some linear subspace $L \subset \mathbb{R}^{N+1}$ of dimension $k \leq N$. Then, according to our induction assumption, $\hat{x} = 0$ is an isolated minimizer od order 2 for f which contradicts the property (8). Now let us suppose that this is not the case. Further let $0 < \rho < 1 - \frac{\sqrt{2}}{2}$ and let $v_1, \ldots, v_{N+1} \in S_{\mathbb{R}^{N+1}} \cap B(h_0; \rho)$ are linearly independent vectors generating a convex cone C with nonempty interior. Without loss of generality we can assume that $x_n \in \operatorname{int} C$, for every $n \in \mathbb{N}$. Let $n \in \mathbb{N}$ be fixed and $t_n = \|x_n\|$. Let F_i be the i-th boundary face of C such that

$$F_i \cap \{x_n + s(h_n - h_0) : s \geq 0\} = \{c_n\}. \tag{9}$$

Then $c_n/\|c_n\| \in S_{\mathbb{R}^{N+1}} \cap B(h_0; \rho)$. In view of our induction assumption, there exist some neighbourhood $V(0)$ and $A > 0$ such that:

$$\frac{f(c)}{\|c\|^2} \geq A > 0, \tag{10}$$

for every $c \in V(0) \cap \partial C$, where ∂C denotes the boundary of C. Further for some $\delta_0 > 0$, we have:

$$\frac{f(th_0)}{t^2} \geq A > 0, \quad \forall t \in (0, \delta_0). \tag{11}$$

From (8) and (10) it follows that if n is sufficiently large, then

$$\frac{f(c_n)}{t_n^2} \geq \frac{f(c_n)}{\|c_n\|^2} \geq A > \frac{1}{n} \geq \frac{f(x_n)}{t_n^2},$$

where $t_n := \|x_n\|$ and $\|c_n\| \geq t_n$. The last inequality can be shown as follows: $c_n = x_n + s_n(h_n - h_0)$, $s_n \geq 0 \Rightarrow c_n = (t_n + s_n)h_n - s_n h_0 \Rightarrow \|c_n\| = \|(t_n + s_n)h_n - s_n h_0\| \geq \|(t_n + s_n)h_n\| - \|s_n h_0\| = (t_n + s_n) - s_n = t_n$. Hence $\|c_n\| \geq t_n$. The last argument shows that $f(c_n) > f(x_n)$ if n is large enough. Next, Lemma 1 gives $\eta_n \in (c_n, x_n)$ such that

$$f^{\ell}(\eta_n; x_n - c_n) \leq f(x_n) - f(c_n) < 0. \tag{12}$$

Now again using Lemma 1, (8), (11), and (12), we have that for some $\xi_n \in (t_n h_n, t_n h_0)$ and n large enough, it holds:

$$\begin{aligned} 0 < \frac{A}{2} \leq \frac{f(t_n h_0) - f(t_n h_n)}{t_n^2} &\leq \frac{f^{\ell}(\xi_n; t_n(h_0 - h_n))}{t_n^2} \\ &< \frac{f^{\ell}(\xi_n; h_0 - h_n) - f^{\ell}(\eta_n; h_0 - h_n)}{t_n}. \end{aligned} \tag{13}$$

Claim 1 *If $n \in \mathbb{N}$ is large enough, then $\|\xi_n - \eta_n\| < t_n$.*

Proof We will now express the difference $\xi_n - \eta_n$:

$$\begin{aligned} \xi_n - \eta_n &= t_n h_n + \theta_1(t_n h_0 - t_n h_n) - [c_n + \theta_2(x_n - c_n)] \\ &= (\theta_1 t_n + s_n - \theta_2 s_n)(h_0 - h_n), \end{aligned}$$

where $\theta_1, \theta_2 \in (0,1),\ s_n \geq 0$ so that $c_n = x_n + s_n(h_n - h_0)$. For a two-dimensional picture see the Figure 1.

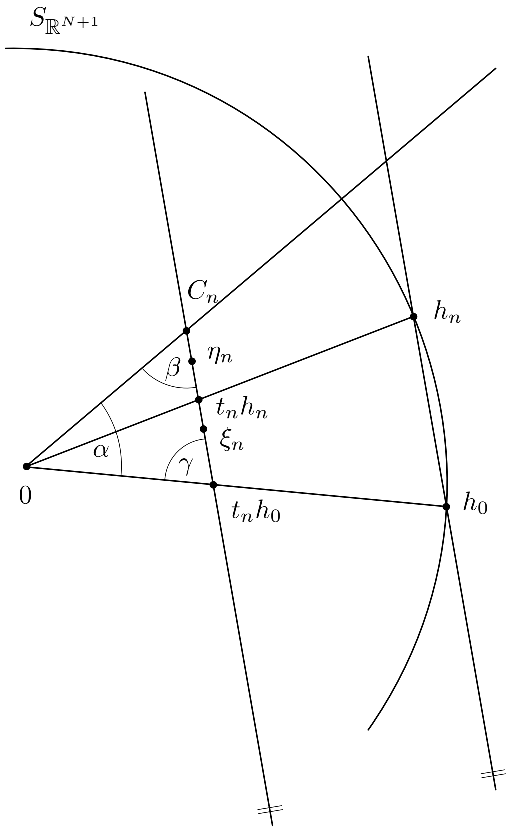

Fig. 1: Two-dimensional picture

Let us consider a triangle with vertices $0, t_n h_0, c_n$ and denote its inner angles by α, β, γ, respectively. It is clear that $\gamma \leq \pi/2$ and $\alpha + \beta + \gamma = \pi$. This implies:

$$\beta = \pi - \alpha - \gamma \geq \frac{\pi}{2} - \alpha. \tag{14}$$

Now let $v_i,\ i = 1 \ldots, N+1$, be one of the vectors generating the cone C and let φ denotes the angle between vectors h_0 and v_i, respectively. Then we have:

$$\begin{aligned} \cos\varphi &= \langle h_0, v_i \rangle = \langle h_0, v_i - h_0 + h_0 \rangle = \langle h_0, v_i - h_0 \rangle + \langle h_0, h_0 \rangle \\ &= 1 + \langle h_0, v_i - h_0 \rangle \geq 1 - \| v_i - h_0 \| \geq 1 - \rho > 1 - \left(1 - \frac{\sqrt{2}}{2}\right) \\ &= \frac{\sqrt{2}}{2} \Rightarrow \varphi < \frac{\pi}{4}. \end{aligned}$$

Suppose that as in the picture, α denotes the angle between vectors h_0 and $c_n/\|c_n\|$. Then also $\alpha < \pi/4$. Indeed, since we can express $c_n/\|c_n\|$ as a convex combination of the pair of vectors v_i, v_{i+1}, $i \in \{1, \dots, N\}$: $c_n/\|c_n\| = \lambda v_i + (1-\lambda)v_{i+1}$, where $\lambda \in [0,1]$, we have

$$\begin{aligned} \cos\alpha &= \left\langle h_0, \frac{c_n}{\|c_n\|} \right\rangle = \langle h_0, \lambda v_i + (1-\lambda)v_{i+1}\rangle \\ &= \lambda\langle h_0, v_i\rangle + (1-\lambda)\langle h_0, v_{i+1}\rangle > \frac{\sqrt{2}}{2} \Rightarrow \alpha < \frac{\pi}{4}. \end{aligned}$$

Thus, we have proved that for the choice of $\rho \in (0, 1-\sqrt{2}/2)$, we have the right estimation for the angle α. Now by (14) it holds: $\beta \geq \pi/4 > \alpha$. This implies the following property of the lengths of sides of the triangle:

$$t_n > \|c_n - t_n h_0\| > \|\xi_n - \eta_n\|.$$

Thus we proved Claim. □

So we are now able to finish the proof of Theorem 7. If $n \in \mathbb{N}$ is large enough, then following (12) and (13), by the $\tilde{\ell}$-stability and by Claim, we can write

$$\begin{aligned} 0 < \frac{A}{2} &< \frac{f^\ell(\xi_n; h_0 - h_n) - f^\ell(\eta_n; h_0 - h_n)}{t_n} \\ &= \frac{1}{\sigma}\frac{f^\ell(\xi_n; \xi_n - \eta_n) - f^\ell(\eta_n; \xi_n - \eta_n)}{t_n} \leq \frac{1}{\sigma}\frac{K\|\xi_n - \eta_n\|^2}{t_n} \\ &< \frac{1}{\sigma}K\|\xi_n - \eta_n\| = K\|h_n - h_0\|, \end{aligned} \tag{15}$$

where $\sigma := \theta_1 t_n + s_n - \theta_2 s_n > 0$ (see the proof of the claim). Now since $\|h_n - h_0\| \to 0$ as $n \to \infty$, we get a contradiction. □

4 Final remarks and questions

There exist functions which are ℓ-stable but not $\tilde{\ell}$-stable at some point. See e.g. [BP2, Ex.2]. It is not clear whether there exists a function which is $\tilde{\ell}$-stable but not ℓ–stable at a certain point. Also note, that the $C^{1,1}$ property implies in an obvious way the $\tilde{\ell}$-stability and $\tilde{\ell}$-stability. This means that Theorem 7 covers the above mentioned theorems 3 and 4 as well as [LK,Theorem 3.4].

Now it seems natural to ask the following questions.

Question 1 Can we distinguish $C^{1,1}$ functions near x and $\tilde{\ell}$-stable functions at x; or can be the $\tilde{\ell}$-stability at x a characterization of $C^{1,1}$ property near x or not?

Theorem 7 would be more elegant if one could answer the following question in the affirmative.

Question 2 Does the the $\tilde{\ell}$-stability of $f : \mathbb{R}^N \to \mathbb{R}$ at $x \in \mathbb{R}^N$ imply the continuity of f near x?

References

[BP1] Bednařík, D., Pastor, K.: *Elimination of strict convergence in optimization.* SIAM J. Control Optim. **43**, 3 (2004), 1063–1077.

[BP2] Bednařík, D., Pastor, K.: *On second-order conditions in unconstrained optimization.* Math. Programming, in print;
online: http://www.springerlink.com/content/tt7g83q36243l144/

[BP3] Bednařík, D., Pastor, K.: *Erratum to Elimination of strict convergence in optimization.* SIAM J. Control Optim. **45** (2006), 382–387.

[BZ] Ben-Tal, A., Zowe, J.: *Directional derivatives in nonsmooth optimization.* J. Optim. Theory Appl. **47** (1985), 483–490.

[CC] Cominetti, R., Correa, R.: *A generalized second-order derivative in nonsmooth optimization.*, SIAM J. Control Optim. **28** (1990), 789–809.

[GGR] Ginchev, I., Guerraggio, A., Rocca, M.: *From scalar to vector optimization.* Appl. Math. **51** (2006), 5–36.

[HSN] Hiriart-Urruty, J. B., Strodiot, J. J., Nguyen, V. H.: *Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data.* Appl. Math. Optim. **11** (1984), 43–56.

[KT] Klatte, D., Tammer, K.: *On second-order sufficient optimality conditions for $C^{1,1}$ optimization problems.* Optimization **19** (1988), 169–179.

[LK] Liu, L., Křížek, M.: *The second order optimality conditions for nonlinear mathematical programming with $C^{1,1}$ data.* Appl. Math. **42** (1997), 311–320.

[Q1] Qi, L.: *Superlinearly convergent approximate Newton methods for LC^1 optimization problem.* Math. Programming **64** (1994), 277–294.

[Q2] Qi, L.: *LC^1 functions and LC^1 optimization.* Operations Research and its applications (D.Z. Du, X.S. Zhang and K. Cheng eds.), World Publishing, Beijing, 1996, pp. 4–13.

[TR] Torre, D. L., Rocca, M.: *Remarks on second order generalized derivatives for differentiable functions with Lipschitzian jacobian.* Applied Mathematics E-Notes **3** (2003), 130–137.

[Zo] Zorich, V. A.: Mathematical Analysis. *Springer-Verlag, Berlin*, 2004.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 19–24

Varieties Satisfying the Triangular Scheme Need Not Be Congruence Distributive*

IVAN CHAJDA[1], RADOMÍR HALAŠ[2]

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

(Received October 31, 2006)

Abstract

A diagrammatic scheme characterizing congruence distributivity of congruence permutable algebras was introduced by the first author in 2001. It is known under the name Triangular Scheme. It is known that every congruence distributive algebra satisfies this scheme and an algebra satisfying the Triangular Scheme which is not congruence distributive was found by E. K. Horváth, G. Czédli and the autor in 2003. On the other hand, it was an open problem if a variety of algebras satisfying the Triangular Scheme must be congruence distributive. We get a negative solution by presenting an example.

Key words: Congruence distributivity; Triangular Scheme, variety of algebras; Jónsson terms.

2000 Mathematics Subject Classification: 08A30, 08B10

Congruence distributive varieties were characterized by B. Jónsson [7] by means of the Maltsev condition. For the reader's convenience, we can repeat this result:

Proposition 1 *A variety $\mathcal{V}$ is congruence distributive if and only if there exist ternary terms $t_0, \dots, t_n$ such that $t_0(x, y, z) = x$, $t_n(x, y, z) = z$ and*

*Supported by the Research Project of the Czech Goverment MSM 6198959214.

(a) for all $i = 0, \dots, n$ *it holds* $t_i(x, y, x) = x$

(b) for i *even,* $t_i(x, x, y) = t_{i+1}(x, x, y)$

(c) for i *odd,* $t_i(x, y, y) = t_{i+1}(x, y, y)$.

These terms $t_0, \dots, t_n$ are referred to be *Jónsson terms.*

Unfortunately, a similar characterization of congruence distributivity for a single algebra is missing. It motivated us to introduce the following concept (see [1], [4]).

Let L be a sublattice of an equivalence lattice (known also as a partition lattice) on a non-void set A. We say that L *satisfies the Triangular Scheme* if for each $\alpha, \beta, \gamma \in L$ with $\alpha \cap \beta \subseteq \gamma$ and for $x, y, z \in A$ such that $\langle x, y\rangle \in \gamma$, $\langle x, z\rangle \in \alpha$, $\langle z, y\rangle \in \beta$ we have $\langle z, y\rangle \in \gamma$.

This can be visualized as follows

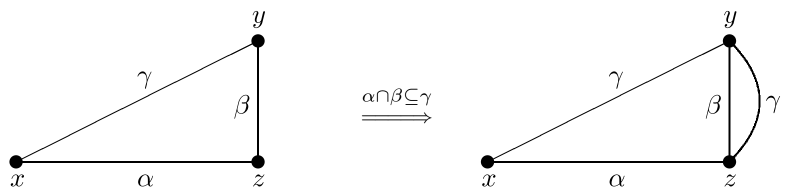

Fig. 1

We say that an algebra $\mathcal{A}$ satisfies the Triangular Scheme if the congruence lattice $\operatorname{Con}\mathcal{A}$ satisfies this condition. A variety $\mathcal{V}$ fulfils the Triangular Scheme if each $\mathcal{A} \in \mathcal{V}$ has this property.

The following was proved in [1], [4].

Proposition 2 *If an algebra is congruence distributive then it satisfies the Triangular Scheme. If an algebra is congruence permutable then it is congruence distributive if and only if it satisfies the Triangular Scheme.*

An example of algebra satisfying the Triangular Scheme but which is not congruence distributive was found in [5].

Let us note that similar schemes for congruence semidistributivity were involved in [3] and conclusions of the Triangular Scheme for n-permutable algebras were treated in [2], [3], [5]. For congruence modular algebras and varieties it was done in [5] where it is explicitely proved that for a variety, the assumption of congruence permutability of Proposition 2 can be replaced by a weaker one of congruence modularity. However, there still was an open question if a variety satisfying the Triangular Scheme is necessarily congruence distributive. To solve this question, we first characterize the Triangular Scheme for varieties by a Maltsev condition.

Theorem 1 *Let* $\mathcal{V}$ *be a variety of algebras. The following are equivalent:*

(1) For each $\mathcal{A} \in \mathcal{V}$, $\operatorname{Con}\mathcal{A}$ *satisfies the Triangular Scheme;*

(2) there exist ternary terms $t_0, \ldots, t_n$ *such that* $t_0(x,y,z) = x$, $t_n(x,y,z) = z$ *and*

(a) for i *even,* $t_i(x,y,x) = t_{i+1}(x,y,x)$, $t_i(x,x,y) = t_{i+1}(x,x,y)$,

(b) for i *odd,* $t_i(x,y,y) = t_{i+1}(x,y,y)$.

Proof Suppose $\mathcal{V}$ satisfies the Triangular Scheme, $\mathcal{F}_{\mathcal{V}}(x,y,z)$ is a free algebra of $\mathcal{V}$ with three free generators and $\alpha = \theta(x,y)$, $\beta = \theta(x,z)$ and $\gamma = (\alpha \cap \beta) \vee \theta(y,z)$. Then $\alpha \cap \beta \subseteq \gamma$ and, by Triangular Scheme, $\langle x,z\rangle \in \gamma$. Hence, there exists an integer $n \geq 0$ and ternary terms $t_0, \ldots, t_n$ such that

$$x = t_0(\alpha \cap \beta)t_1\theta(y,z)t_2(\alpha \cap \beta)t_3 \ldots t_n = z.$$

Applying the standard procedure, we easily derive that $t_0(x,y,z) = x$, $t_n(x,y,z) = z$ and $t_i(x,y,x) = t_{i+1}(x,y,x)$ and $t_i(x,x,y) = t_{i+1}(x,x,y)$ for i even, and $t_i(x,y,y) = t_{i+1}(x,y,y)$ for i odd.

Prove the converse. Let $\mathcal{A} = (A,F) \in \mathcal{V}$, $a,b,c \in A$, $\alpha,\beta,\gamma \in \operatorname{Con}\mathcal{A}$ and $\alpha \cap \beta \subseteq \gamma$. Suppose $\langle c,b\rangle \in \gamma$, $\langle a,b\rangle \in \beta$ and $\langle a,c\rangle \in \alpha$. Then

$$t_i(a,b,c)\alpha t_i(a,b,a) = t_{i+1}(a,b,a)\alpha t_{i+1}(a,b,c)$$

$$t_i(a,b,c)\beta t_i(a,a,c) = t_{i+1}(a,a,c)\beta t_{i+1}(a,b,c)$$

for i even and

$$t_i(a,b,c)\gamma t_i(a,b,b) = t_{i+1}(a,b,b)\gamma t_{i+1}(a,b,c)$$

for i odd. Altogether, we conclude

$$a = t_0(a,b,c)(\alpha \cap \beta)t_1(a,b,c)\gamma t_2(a,b,c)(\alpha \cap \beta) \ldots t_n(a,b,c) = c$$

thus

$$\langle a,c\rangle \in (\alpha \cap \beta) \circ \gamma \circ (\alpha \cap \beta) \circ \gamma \circ \ldots \subseteq \gamma \circ \gamma \circ \ldots \circ \gamma = \gamma.$$

This together with $\langle b,c\rangle \in \gamma$ yields $\langle a,b\rangle \in \gamma$. Hence, $\mathcal{A}$ and also $\mathcal{V}$ satisfies the Triangular Scheme. □

Remark 1 When comparing our terms of Theorem 2 with Jónsson terms, the difference is that we do not ask $t_i(x,y,x) = t_{i+1}(x,y,x)$ for i odd. It motivates us to suppose that this variety need not be necessarily congruence distributive. However, if $n \leq 3$ then $t_0(x,y,x) = x$ and $t_3(x,y,x) = x$ yield that also $t_1(x,y,x) = x$ and $t_2(x,y,x) = x$. To find an example of a variety which is not congruence distributive but still satisfying the Triangular Scheme, we must suppose that $n \geq 4$. We are ready to construct such an example:

Example 1 Consider a variety $\mathcal{V}$ of type $(2,1,1)$ whose operations are denoted by $\wedge$ and f, g and satisfying the identities

$$x \wedge x = x, \quad x \wedge y = y \wedge x, \quad x \wedge (y \wedge z) = (x \wedge y) \wedge z$$

(i.e. the $\wedge$-reducts of its members are semilattices) and

$$f(f(x)) = x$$

$$x \wedge g(g(x) \wedge g(y)) = x$$

$$x \wedge g(g(y)) = x \wedge f(f(x) \wedge f(y)).$$

Hence it follows also

$$x \wedge g(g(x)) = x.$$

We can take $n = 6$ and establish the following terms:

$$t_0(x, y, z) = x$$

$$t_1(x, y, z) = x \wedge g(g(y) \wedge g(z))$$

$$t_2(x, y, z) = x \wedge g(g(y)) \wedge f(f(x) \wedge f(z))$$

$$t_3(x, y, z) = x \wedge g(g(y)) \wedge f(f(y) \wedge f(z))$$

$$t_4(x, y, z) = x \wedge y \wedge z$$

$$t_5(x, y, z) = y \wedge z$$

$$t_6(x, y, z) = z.$$

Then for i even we have
$i = 0$:

$$t_0(x, x, y) = x = x \wedge g(g(x) \wedge g(y)) = t_1(x, x, y)$$

$$t_0(x, y, x) = x = x \wedge g(g(y) \wedge g(x)) = t_1(x, y, x)$$

$i = 2$:

$$t_2(x, x, y) = x \wedge g(g(x)) \wedge f(f(x) \wedge f(y)) = t_3(x, x, y)$$

$$t_2(x, y, x) = x \wedge g(g(y)) \wedge f(f(x)) = x \wedge g(g(y)) =$$

$$= x \wedge g(g(y)) \wedge f(f(y) \wedge f(x)) = t_3(x, y, x)$$

$i = 4$:

$$t_4(x, x, y) = x \wedge y = t_5(x, x, y)$$

$$t_4(x, y, x) = x \wedge y = t_5(x, y, x).$$

For i odd we have
$i = 1$:

$$t_1(x, y, y) = x \wedge g(g(y)) = x \wedge g(g(y)) \wedge f(f(x) \wedge f(y)) = t_2(x, y, y)$$

$i = 3$:

$$t_3(x, y, y) = x \wedge g(g(y)) \wedge f(f(y)) = x \wedge y \wedge g(g(y)) = x \wedge y = t_4(x, y, y)$$

$i = 5$:

$$t_5(x, y, y) = y \wedge y = y = t_6(x, y, y).$$

We have shown that our variety $\mathcal{V}$ satisfies the Triangular Scheme. Consider now a four element $\wedge$-semilattice as drawn in Fig. 2 where f and g

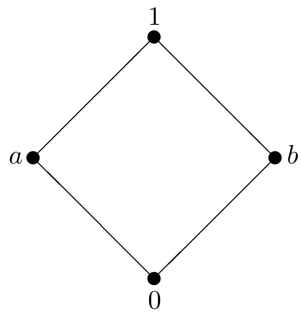

Fig. 2

are determined by the table

x	$f(x)$	$g(x)$
0	a	1
a	0	1
b	1	a
1	b	a

It is an easy exercise to check that $\mathcal{A} = (\{0, a, b, 1\}; \wedge, f, g) \in \mathcal{V}$. Consider the partitions:

$$\alpha = \{0, a\}$$

$$\beta = \{0, a\}, \{b, 1\}$$

$$\gamma = \{0, b\}, \{a, 1\}.$$

Then apparently $\operatorname{Con}\mathcal{A} = \{\omega, \alpha, \beta, \gamma, A \times A\}$ as shown in Fig. 3 thus $\mathcal{A}$ is not congruence distributive.

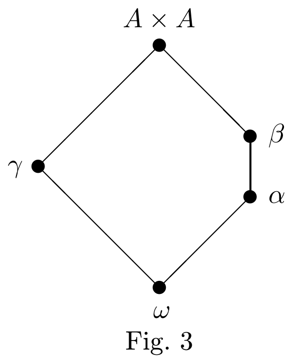

Fig. 3

References

[1] Chajda, I.: *A note on the triangular scheme.* East-West J. of Mathem. **3** (2001), 79–80.

[2] Chajda, I., Halaš, R.: *On schemes for congruence distributivity*, Central European J. of Mathem. **2**, 3 (2004), 368–376.

[3] Chajda, I., Horváth, E. K.: *A scheme for congruence semidistributivity.* Discuss. Math., General Algebra and Appl. **23** (2003), 13–18.

[4] Chajda, I., Horváth, E. K.: *A triangular scheme for congruence distributivity.* Acta Sci. Math. (Szeged) **68** (2002), 29–35.

[5] Chajda, I., Horváth, E. K., Czédli, G.: *Trapezoid Lemma and congruence distributivity.* Math. Slovaca **53** (2003), 247–253.

[6] Chajda, I., Horváth, E. K., Czédli, G.: *The Shifting Lemma and shifting lattice identities.* Algebra Universalis **50** (2003), 51–60.

[7] Jónsson, B.: *Algebras whose congruence lattice are distributive.* Math. Scand. **21** (1967), 110–121.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 25–33

Ideals, Congruences and Annihilators on Nearlattices*

Ivan CHAJDA[1], Miroslav KOLAŘÍK[2]

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

(Received January 30, 2007)

Abstract

By a nearlattice is meant a join-semilattice having the property that every principal filter is a lattice with respect to the semilattice order. We introduce the concept of (relative) annihilator of a nearlattice and characterize some properties like distributivity, modularity or 0-distributivity of nearlattices by means of certain properties of annihilators.

Key words: Nearlattice; semilattice; ideal; congruence; distributivity; modularity; 0-distributivity; annihilator.

2000 Mathematics Subject Classification: 06A12, 06D99, 06C99

1 Introduction

Algebraic structures being join-semilattices with respect to the induced order relation appear frequently in algebraic logic. For example, implication algebras, introduced by J. C. Abbott [1], describe algebraic properties of the logical connective implication in the classical propositional logic. Implication algebras have a very nice structure: with respect to the induced order, they are join-semilattices, principal filters of which are Boolean algebras. Analogously, for various logics of quantum mechanics the corresponding algebraic structures have a semilattice structure with principal filters being special lattices.

*Supported by the Research Project of the Czech Goverment MSM 6198959214.

This fact motivated us to describe $\vee$-semilattices where every principal filter is a lattice. They are called nearlattices (see e.g. [3, 5, 6, 11, 14, 15, 16, 17]).

More precisely, we studied the following structures.

Definition 1 A semilattice $\mathcal{N} = (N; \vee)$, where for each $a \in N$ the principal filter $[a) = \{x \in N; a \leq x\}$ is a lattice with respect to the induced order $\leq$ of $\mathcal{N}$, is called a *nearlattice.*

It has been shown [4, 11] that nearlattices can be considered as algebras with one ternary operation. Moreover, nearlattices considered as algebras of type (3) form an equational class: indeed, if $x, y, z \in N$ for a nearlattice $\mathcal{N}$, the element $(x \vee z) \wedge (y \vee z)$ is correctly defined since both $x \vee z, y \vee z \in [z)$ and $[z)$ is a lattice, and the following holds:

Proposition 1 ([4]) *Let $\mathcal{N} = (N; \vee)$ be a nearlattice. Define a ternary operation by $m(x, y, z) = (x \vee z) \wedge (y \vee z)$ on N. Then $m(x, y, z)$ is an everywhere defined operation and the following identities are satisfied:*

(P1) $m(x, y, x) = x$;

(P2) $m(x, x, y) = m(y, y, x)$;

(P3) $m(m(x, x, y), m(x, x, y), z) = m(x, x, m(y, y, z))$;

(P4) $m(x, y, p) = m(y, x, p)$;

(P5) $m(m(x, y, p), z, p) = m(x, m(y, z, p), p)$;

(P6) $m(x, m(y, y, x), p) = m(x, x, p)$;

(P7) $m(m(x, x, p), m(x, x, p), m(y, x, p)) = m(x, x, p)$;

(P8) $m(m(x, x, z), m(y, y, z), z) = m(x, y, z)$.

Conversely, let $\mathcal{N} = (N; m)$ be an algebra of type (3) *satisfying* (P1)–(P7)*. If we define $x \vee y = m(x, x, y)$, then $(N; \vee)$ is a join-semilattice and for each $p \in N$, $([p); \leq)$ is a lattice, where for $x, y \in [p)$ their infimum is $x \wedge y = m(x, y, p)$. Hence $(N; \vee)$ is a nearlattice. If, moreover, $\mathcal{N} = (N; m)$ satisfies also* (P8)*, then the correspondence between nearlattices and algebras $(N; m)$ satisfying* (P1)–(P8) *is one-to-one.*

Thus nearlattices similarly as lattices have two faces and we shall alternate in our investigations between them depending which one will be more convenient.

The following notions of distributivity for nearlattices have been introduced in [4]:

Definition 2 Let $\mathcal{N} = (N; m)$ be an algebra of type (3). We call $\mathcal{N}$ *distributive* if it satisfies the identity

(D1) $m(x, m(y, y, z), p) = m(m(x, y, p), m(x, y, p), m(x, z, p))$.

If $\mathcal{N}$ satisfies the identity

(D2) $m(x, x, m(y, z, p)) = m(m(x, x, y), m(x, x, z), p)$,

it is called *dually distributive.*

It is expected that both notions are related in the case of nearlattices. Indeed, one can prove the following statement:

Proposition 2 ([4]) *Let* $\mathcal{N} = (N; m)$ *be an algebra of type* (3) *satisfying* (P1)–(P7). *Then the following conditions are equivalent:*

(1) $\mathcal{N}$ *is distributive;*

(2) $\mathcal{N}$ *is dually distributive;*

(3) *in the associated semilattice, every principal filter is a distributive lattice.*

Due to the previous description of distributivity for nearlattices, we are able to get very simple arguments to prove that in a distributive nearlattice $\mathcal{N}$, every ideal of $\mathcal{N}$ is a congruence class.

2 Ideals and congruence classes on distributive nearlattices

The concept of an ideal in a distributive nearlattice was defined in [10]:

Definition 3 A subset $\emptyset \neq I \subseteq N$ of a nearlattice $\mathcal{N} = (N; m)$ is called an *ideal* if

(I1) $m(x, x, y) \in I$ for all $x, y \in I$;

(I2) $m(x, y, p) \in I$ for all $x \in I$ and $y, p \in N$ with $p \leq x$.

Note that I is an ideal of $\mathcal{N}$ if and only if it is a downset closed under suprema with respect to the induced order of $\mathcal{N}$.

Lemma 1 *A subset* $\emptyset \neq I \subseteq N$ *of a nearlattice* $\mathcal{N} = (N; \vee)$ *is an ideal if and only if it satisfies the following two conditions*

(i1) $x, y \in I \Rightarrow x \vee y \in I$;

(i2) $x \in I,\ a \leq x \Rightarrow a \in I$.

Proof It is clear. □

Example 1 Let $\mathcal{N} = (\{x, x \vee y, y, p, q, 1\}; \vee)$ be a nearlattice whose diagram is depicted in Fig. 1. The set $I = \{x, x \vee y, y\}$ is clearly an ideal on $\mathcal{N}$.

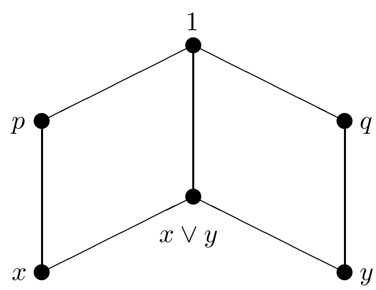

Fig. 1

By a *congruence* on a nearlattice $\mathcal{N} = (N; m)$ we mean an equivalence relation Θ on N such that for all $x_1, x_2, y_1, y_2, z_1, z_2 \in N$ we have that $\langle x_1, x_2 \rangle \in \Theta$, $\langle y_1, y_2 \rangle \in \Theta$, $\langle z_1, z_2 \rangle \in \Theta$ imply

$$\langle m(x_1, y_1, z_1), m(x_2, y_2, z_2) \rangle \in \Theta.$$

This concept can be translated for the alternative description of a nearlattice as follows:

Lemma 2 *Let $\mathcal{N} = (N; \vee)$ be a nearlattice. Then Θ is a congruence on $\mathcal{N}$ if and only if it is an equivalence relation on N which satisfies the following implication $(*)$:*
$\langle x_1, x_2 \rangle, \langle y_1, y_2 \rangle \in \Theta \Rightarrow \langle x_1 \vee y_1, x_2 \vee y_2 \rangle \in \Theta$, *and* $\langle x_1 \wedge y_1, x_2 \wedge y_2 \rangle \in \Theta$, *whenever $x_1 \wedge y_1, x_2 \wedge y_2$ are defined.*

Proof $(\Rightarrow)$: Let Θ be a congruence on $\mathcal{N}$. Let $\langle x_1, x_2 \rangle \in \Theta$ and $\langle y_1, y_2 \rangle \in \Theta$. Then, by definition of congruence on $\mathcal{N}$, $\langle m(x_1, x_1, y_1), m(x_2, x_2, y_2) \rangle \in \Theta$, i.e. $\langle x_1 \vee y_1, x_2 \vee y_2 \rangle \in \Theta$. Now, we observe the following property of Θ :

(P) If $x \leq y$, $\langle x, y \rangle \in \Theta$ and $x \wedge z$ exists, then $\langle x \wedge z, y \wedge z \rangle \in \Theta$.

Indeed, we have $\langle m(x, z, x \wedge z), m(y, z, x \wedge z) \rangle \in \Theta$, where

$$m(x, z, x \wedge z) = (x \vee (x \wedge z)) \wedge (z \vee (x \wedge z)) = x \wedge z$$

and

$$m(y, z, x \wedge z) = (y \vee (x \wedge z)) \wedge (z \vee (x \wedge z)) = y \wedge z,$$

and hence $\langle x \wedge z, y \wedge z \rangle \in \Theta$.

Now, assume that $\langle x_1, x_2 \rangle, \langle y_1, y_2 \rangle \in \Theta$, and $x_1 \wedge y_1, x_2 \wedge y_2$ exists. Then $\langle x_1, x_1 \vee x_2 \rangle \in \Theta$ and since $x_1 \wedge y_1$ exists, we have $\langle x_1 \wedge y_1, (x_1 \vee x_2) \wedge y_1 \rangle \in \Theta$ by (P). Analogously, $\langle y_1, y_1 \vee y_2 \rangle \in \Theta$ entails

$$\langle (x_1 \vee x_2) \wedge y_1, (x_1 \vee x_2) \wedge (y_1 \vee y_2) \rangle \in \Theta.$$

Therefore

$$\langle x_1 \wedge y_1, (x_1 \vee x_2) \wedge (y_1 \vee y_2) \rangle \in \Theta.$$

Similarly we can show that $\langle x_2 \wedge y_2, (x_1 \vee x_2) \wedge (y_1 \vee y_2) \rangle \in \Theta$. Consequently, due to transitivity of Θ we obtain $\langle x_1 \wedge y_1, x_2 \wedge y_2 \rangle \in \Theta$.
$(\Leftarrow)$: Let Θ be an equivalence relation on N satisfying $(*)$. Let $\langle x_1, x_2 \rangle$, $\langle y_1, y_2 \rangle$, $\langle z_1, z_2 \rangle \in \Theta$. Then $\langle x_1 \vee z_1, x_2 \vee z_2 \rangle$, $\langle y_1 \vee z_1, y_2 \vee z_2 \rangle \in \Theta$, and hence also

$$\langle (x_1 \vee z_1) \wedge (y_1 \vee z_1), (x_2 \vee z_2) \wedge (y_2 \vee z_2) \rangle \in \Theta,$$

i.e. $\langle m(x_1, y_1, z_1), m(x_2, y_2, z_2) \rangle \in \Theta$, thus Θ is a congruence on $\mathcal{N}$. □

We can show that for distributive nearlattices, the ideals are related to congruences in the same way as it is for lattices (see e.g. [9]).

Theorem 1 *Let $\mathcal{N} = (N; \vee)$ be a distributive nearlattice. Then each ideal I of $\mathcal{N}$ is a congruence class of $\Theta_I \in Con\mathcal{N}$, defined by*

$$\langle x, y\rangle \in \Theta_I \quad \textit{iff} \quad \textit{there exists } c \in I \textit{ such that } x \vee c = y \vee c.$$

Proof Of course, Θ_I is reflexive and symmetric. Suppose $\langle a, b\rangle \in \Theta_I$ and $\langle b, c\rangle \in \Theta_I$. Then $a \vee x = b \vee x$ and $b \vee y = c \vee y$ for some $x, y \in I$. Since I is an ideal, we have $x \vee y \in I$. Thus $a \vee x \vee y = b \vee x \vee y = c \vee x \vee y$, whence $\langle a, c\rangle \in \Theta_I$, i.e. Θ_I is an equivalence on N.

Let $\langle a, b\rangle \in \Theta_I$ and $c \in N$. Then there exists $x \in I$ such that $a \vee x = b \vee x$ and thus $a \vee c \vee x = b \vee c \vee x$, hence $\langle a \vee c, b \vee c\rangle \in \Theta_I$. Using transitivity of Θ_I, we easily obtain that Θ_I is compatible with the operation $\vee$.

Now, let $\langle a, b\rangle \in \Theta_I$, $\langle c, d\rangle \in \Theta_I$ and let $a \wedge c$, $b \wedge d$ are defined. Then $a \vee x = b \vee x$ and $c \vee y = d \vee y$ for some $x, y \in I$. Applying distributivity of N, we have

$$(a \vee x) \wedge (c \vee y) = (a \wedge c) \vee (x \wedge c) \vee (a \wedge y) \vee (x \wedge y) = (a \wedge c) \vee z,$$

where $z = (x \wedge (c \vee y)) \vee (y \wedge (a \vee x)) \in I$. Analogously,

$$(b \vee x) \wedge (d \vee y) = (b \wedge d) \vee (x \wedge d) \vee (b \wedge y) \vee (x \wedge y) = (b \wedge d) \vee z,$$

which gives $\langle a \wedge c, b \wedge d\rangle \in \Theta_I$, i.e. Θ_I is compatible with a partial operation $\wedge$. Applying Lemma 2, we have shown that Θ_I is a congruence on $\mathcal{N}$.

Further, suppose $a, b \in I$. By (i1), $a \vee b \in I$ and since $a \vee (a \vee b) = b \vee (a \vee b)$, we have $\langle a, b\rangle \in \Theta_I$. Conversely, let $a \in I$ and $\langle a, c\rangle \in \Theta_I$. Then there exists $x \in I$ such that $a \vee x = c \vee x$. But $a \vee x \in I$, whence $c \vee x \in I$. Since $c \leq c \vee x$, by (i2) we have $c \in I$, which yields $I = [a]_{\Theta_I}$, i.e. I is a class of Θ_I. □

Corollary 1 *Each ideal of a nearlattice $\mathcal{N} = (N; \vee)$ is a class of at least one congruence if and only if $\mathcal{N}$ is distributive.*

Proof If $\mathcal{N}$ is distributive and I is its ideal then, by Theorem 1, I is a class of the congruence Θ_I.

Conversely, let $\mathcal{N}$ be not distributive. Then, by Proposition 2, there exists a principal filter $[b)$ which is not a distributive lattice, i.e. it contains N_5 or M_3 (see Fig. 2).

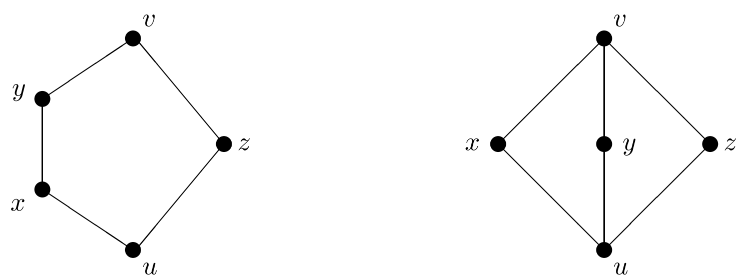

Fig. 2

In both cases, one can easily prove that $(x] = I(x) = \{a \in N; a \leq x\}$ is an ideal on nearlattice $\mathcal{N}$ which is not a class of any congruence Θ on $\mathcal{N}$. Indeed, let $(x]$ be a class of congruence Θ on $\mathcal{N}$. Since $u, x \in (x]$ we have $\langle u, x\rangle \in \Theta$ (see Fig. 2). So $\langle u \vee z, x \vee z\rangle \in \Theta$, i.e. $\langle z, v\rangle \in \Theta$. Further, $\langle z \wedge y, v \wedge y\rangle \in \Theta$ because $z \wedge y$ and $v \wedge y$ exists in $[u)$. Hence $\langle u, y\rangle \in \Theta$, which yields $y \in (x]$, a contradiction. □

3 Annihilators on nearlattices

The aim of this section is to show that annihilators can be used for a characterization of distributivity or modularity of nearlattices in the way similar to that for lattices, see e.g. [7, 12, 13]. However, the concept of a relative annihilator must be defined in a slightly different way from that for lattices [2, 8, 12].

Definition 4 Let $\mathcal{N} = (N; \vee)$ be a nearlattice and $a, b, x, z \in N$. By a *relative annihilator of* a *with respect to* b we mean the set $\langle a, b\rangle = \{z \in N; z \leq x$ where $a \wedge x$ exists and $a \wedge x \leq b\}$.

Remark 1 It means that our relative annihilator in a nearlattice is in fact a downset of a relative annihilator as defined in [7, 12, 13]. The reason is that e.g. for $\langle q, y\rangle$ of the nearlattice from Example 1 we have $(x \vee y) \wedge q \leq y$ thus $x \vee y \in \langle q, y\rangle$ but $x \leq x \vee y$ and $x \wedge q$ is not defined. Hence, we must extend the original concept into a downset.

Theorem 2 *Let $\mathcal{N} = (N; \vee)$ be a nearlattice. The following conditions are equivalent:*

(i) *$\mathcal{N}$ is distributive;*

(ii) *$\langle a, b\rangle$ is an ideal of $\mathcal{N}$ for all $a, b \in N$;*

(iii) *$\langle a, b\rangle$ is an ideal of $\mathcal{N}$ for each $b \leq a$.*

Proof (i)⇒(ii): Let $\mathcal{N}$ be distributive and $a, b \in N$. Suppose $z \in \langle a, b\rangle$ and $y \leq z$. Then obviously $y \in \langle a, b\rangle$. If $z, y \in \langle a, b\rangle$ then $z \leq x_1$ with $a \wedge x_1 \leq b$ and $y \leq x_2$ with $a \wedge x_2 \leq b$ (for some $x_1, x_2 \in N$). Thus $z \vee y \leq x_1 \vee x_2$. It is evident that all considered meets exist and due to distributivity of $\mathcal{N}$,

$$(x_1 \vee x_2) \wedge a = (x_1 \wedge a) \vee (x_2 \wedge a) \leq b.$$

Hence $x_1 \vee x_2 \in \langle a, b\rangle$ and thus also $z \vee y \in \langle a, b\rangle$, i.e. $\langle a, b\rangle$ is an ideal of $\mathcal{N}$.

(ii)⇒(iii) is trivial. Prove (iii)⇒(i). Let $a \in N$ and $x, y, z \in [a)$. Then $y \wedge x, z \wedge x$ exist and $(y \wedge x) \vee (z \wedge x) \leq x$. Hence, by (iii), $\langle x, (y \wedge x) \vee (z \wedge x)\rangle$ is an ideal I of $\mathcal{N}$. Since $x \wedge y \leq (y \wedge x) \vee (z \wedge x)$, we have $y \in I$. Analogously, $x \wedge z \leq (y \wedge x) \vee (z \wedge x)$, thus $z \in I$ and hence also $y \vee z \in I$, i.e. $(y \vee z) \wedge x \leq (y \wedge x) \vee (z \wedge x)$. We have shown that $[a)$ is a distributive lattice thus the nearlattice $\mathcal{N}$ is distributive. □

Example 2 A nearlattice $\mathcal{N}$ depicted in Fig. 3 is not distributive and hence the relative annihilator $\langle a, b\rangle = \{p, q, b, x, y\}$ is not an ideal of $\mathcal{N}$ because $x, b \in \langle a, b\rangle$ but $1 = x \vee b \notin \langle a, b\rangle$.

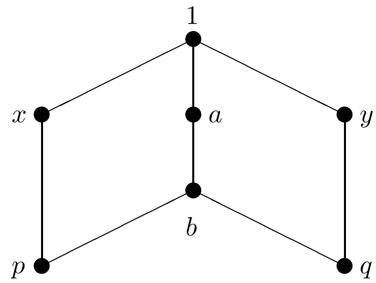

Fig. 3

We say that a nearlattice $\mathcal{N} = (N; \vee)$ is *modular* if each its principal filter is a modular lattice with respect to the induced order $\leq$.

The following result is a generalization of that from [8] for nearlattices:

Theorem 3 *Let $\mathcal{N} = (N; \vee)$ be a nearlattice. The following conditions are equivalent:*

(i) *$\mathcal{N}$ is modular;*

(ii) *$x \vee y \in \langle a, b\rangle$ for each $b \leq a$ and all $x \in (b]$, $y \in \langle a, b\rangle$.*

Proof (i)⇒(ii): Let $y \in \langle a, b\rangle$ for $b \leq a$ and $x \in (b]$, i.e. $x \leq b \leq a$, thus $x, b, a, x \vee y \in [x)$ and, due to modularity of the lattice $[x)$,

$$a \wedge (x \vee y) = (a \wedge y) \vee x \leq b$$

whence $x \vee y \in \langle a, b\rangle$.
(ii)⇒(i): Let $x, y, z \in [a)$ for some $a \in N$ with $x \leq z$. Then $z \wedge y$ exists in $[a)$ and $x \vee (z \wedge y) \leq z$ and $z \wedge x = x \leq x \vee (z \wedge y)$, therefore $x \in \langle z, x \vee (z \wedge y)\rangle$. Further, $z \wedge y \leq x \vee (z \wedge y)$ thus $y \in \langle z, x \vee (z \wedge y)\rangle$. By (ii) we have $x \vee y \in \langle z, x \vee (z \wedge y)\rangle$, i.e. $(x \vee y) \wedge z \leq x \vee (y \wedge z)$ and hence $[a)$ as well as $\mathcal{N}$ is modular. □

Example 3 One can easily see that the nearlattice $\mathcal{N}$ in Fig. 3 is not modular. For $b \leq a$ and for $p \in (b]$, $y \in \langle a, b\rangle$ we have $1 = p \vee y \notin \langle a, b\rangle$.

Let $\mathcal{N} = (N; \vee)$ be a nearlattice and $\emptyset \neq A \subseteq N$. A is called a *sublattice* of $\mathcal{N}$ if it is a lattice with respect to the induced order $\leq$ of $\mathcal{N}$ and $\vee$ and $\wedge$ coincide with the corresponding operations of $\mathcal{N}$.

A sublattice M of a nearlattice $\mathcal{N}$ is called *maximal* if M is not a proper sublattice of another sublattice of $\mathcal{N}$.

From now on, we will suppose that every maximal sublattice M_γ of a nearlattice $\mathcal{N}$ has a least element 0_γ.

We define a nearlattice $(N;\vee)$ to be 0-*distributive* if for all $x,y,z\in M_\gamma$, if $x\wedge y$, $x\wedge z$ are defined and

$$x\wedge y=0_\gamma=x\wedge z \quad\text{ then }\quad x\wedge(y\vee z)=0_\gamma.$$

Definition 5 Let $\mathcal{N}$ be a nearlattice such that each of its maximal sublattices M_γ has a least element 0_γ. For $a\in N$, define $\langle a\rangle_\gamma=\{y\in M_\gamma; a\wedge y=0_\gamma\}$, the so-called *annihilator of* a.

Remark 2 It is an easy observation that if a nearlattice $\mathcal{N}$ has a least element 0 (and hence it is a lattice M_1) then we have $\langle a\rangle_1=\langle a,0\rangle$ for each $a\in N$. Moreover, in every nearlattice $\mathcal{N}$ where each maximal sublattice M_γ has a least element 0_γ we have $\langle a\rangle_\gamma=\langle a,0_\gamma\rangle\cap M_\gamma$ for each $a\in M_\gamma$.

Theorem 4 *Let $\mathcal{N}$ be a nearlattice such that each of its maximal sublattices M_γ has a least element 0_γ. The following conditions are equivalent:*

(i) *every M_γ is 0-distributive;*

(ii) *$\langle a\rangle_\gamma$ is an ideal in M_γ for each $a\in N$ whenever $\langle a\rangle_\gamma\neq\emptyset$.*

Proof (i)⇒(ii): Let $x,y,z\in M_\gamma$ and assume $x\wedge z=0_\gamma$, $y\wedge z=0_\gamma$. Due to 0-distributivity of M_γ, also $(x\vee y)\wedge z=0_\gamma$ and hence $x\vee y\in\langle z\rangle_\gamma$. Of course, if $t\in\langle z\rangle_\gamma$ and $u\leq t$ for $u\in M_\gamma$ then $z\wedge u\leq z\wedge t=0_\gamma$ whence $u\in\langle z\rangle_\gamma$. Thus $\langle z\rangle_\gamma$ is an ideal of M_γ.

(ii)⇒(i): Let $a,b,c\in M_\gamma$ and $a\wedge c=0_\gamma$, $b\wedge c=0_\gamma$. Then $a,b\in\langle c\rangle_\gamma$ and, by (ii), also $a\vee b\in\langle c\rangle_\gamma$, i.e. $(a\vee b)\wedge c=0_\gamma$ thus M_γ is 0-distributive. □

Example 4 (a) Consider the nearlattice $\mathcal{N}=(N;\vee)$ depicted in Fig. 4.

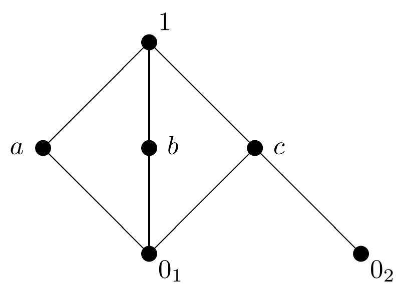

Fig. 4

Clearly, $M_1=\{0_1,a,b,c,1\}$ and $M_2=\{0_2,c,1\}$ are the only maximal sublattices of $\mathcal{N}$. We have $a\wedge b=b\wedge c=0_1$, but $b\wedge(a\vee c)=b\wedge 1=b\neq 0_1$, so M_1 is not 0-distributive, i.e. $\mathcal{N}$ is not 0-distributive. Let us note that for $x\in\{a,b,c\}$, the set $\langle x\rangle_1$ is not an ideal in M_1. On the contrary, M_2 is 0-distributive and for each its element $y\in M_2$, the set $\langle y\rangle_2$ is an ideal in M_2.

(b) It is easy to check that for each $a\in N$ of the nearlattice $\mathcal{N}$ from Example 2 (see Fig. 3), if $\langle a\rangle_\gamma\neq\emptyset$ then it is an ideal in M_γ ($\gamma=1,2$ and $0_1=p$, $0_2=q$). Hence $\mathcal{N}$ is 0-distributive.

References

[1] Abbott, J. C.: *Semi-boolean algebra.* Mat. Vestnik **4** (1967), 177–198.

[2] Beazer, R.: *Hierarchies of distributive lattices satisfying annihilator conditions.* J. London Math. Soc. **11** (1975), 216–222.

[3] Chajda, I., Kolařík, M.: *A decomposition of homomorphic images of nearlattices.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **45** (2006), 43–52.

[4] Chajda, I., Kolařík, M.: *Nearlattices.* Discrete Math., to appear.

[5] Cornish, W. H.: *The free implicative BCK-extension of a distributive nearlattice.* Math. Japonica **27**, 3 (1982), 279–286.

[6] Cornish, W. H., Noor, A. S. A.: *Standard elements in a nearlattice.* Bull. Austral. Math. Soc. **26**, 2 (1982), 185–213.

[7] Davey, B.: *Some annihilator conditions on distributive lattices.* Algebra Universalis **4** (1974), 316–322.

[8] Davey, B., Nieminen, J.: *Annihilators in modular lattices.* Algebra Universalis **22** (1986), 154–158.

[9] Grätzer, G.: General Lattice Theory. *Birkhäuser Verlag*, Basel, 1978.

[10] Halaš, R.: *Subdirectly irreducible distributive nearlattices.* Math. Notes **7**, 2 (2006), 141–146.

[11] Hickman, R.: *Join algebras.* Communications in Algebra **8** (1980), 1653–1685.

[12] Mandelker, M.: *Relative annihilators in lattices.* Duke Math. J. **40** (1970), 377–386.

[13] Nieminen, J.: *The Jordan–Hölder chain condition and annihilators in finite lattices.* Tsukuba J. Math. **14** (1990), 405–411.

[14] Noor, A. S. A., Cornish, W. H.: *Multipliers on a nearlattices.* Commentationes Mathematicae Universitatis Carolinae (1986), 815–827.

[15] Scholander, M.: *Trees, lattices, order and betweenness.* Proc. Amer. Math. Soc. **3** (1952), 369–381.

[16] Scholander, M.: *Medians and betweenness.* Proc. Amer. Math. Soc. **5** (1954), 801–807.

[17] Scholander, M.: *Medians, lattices and trees.* Proc. Amer. Math. Soc. **5** (1954), 808–812.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 35–42

Equipping Distributions for Linear Distribution*

Marina F. GREBENYUK[1], Josef MIKEŠ[2]

[1] *Department of Economical Cybernetics, Aviation National University, Komarova 1, Kiev, Ukraine*
e-mail: ahha@i.com.ua

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

(Received November 27, 2006)

Abstract

In this paper there are discussed the three-component distributions of affine space A_{n+1}. Functions $\{\mathcal{M}^\sigma\}$, which are introduced in the neighborhood of the second order, determine the normal of the first kind of $\mathcal{H}$-distribution in every center of $\mathcal{H}$-distribution.

There are discussed too normals $\{\mathcal{Z}^\sigma\}$ and quasi-tensor of the second order $\{\mathcal{S}^\sigma\}$. In the same way bunches of the projective normals of the first kind of the $\mathcal{M}$-distributions were determined in the differential neighborhood of the second and third order.

Key words: Equipping distributions; linear distribution; affine space.

2000 Mathematics Subject Classification: 53B05

1 Introduction

The given paper applies to differential geometry of a multi-dimensional affine space A_{n+1}. The three-component distributions of an affine space are discussed.

*Supported by grant No. 201/05/2707 of The Czech Science Foundation and by the Council of the Czech Government MSM 6 198 959 214.

Functions $\{\mathcal{M}^\sigma\}$ are introduced in the neighborhood of the second order. They determine the normal of the first kind of a $\mathcal{H}$-distribution in every center of a $\mathcal{H}$-distribution. The normal $\{\mathcal{M}^\sigma\}$ is a generalization of Miheylesku normal of the first kind for a hyperplane distribution of an affine space. The field of the normals $\{\mathcal{Z}^\sigma\}$ was constructed by an inner invariant method in the third differential neighborhood of the forming element of the $\mathcal{H}$-distribution. The object $\{Z^\sigma\}$ determines the projective normal – analog of Fubini normal for the $\mathcal{H}$-distribution in every center of the forming element of the $\mathcal{H}$-distribution. The quasi-tensor of the second order $\{S^\sigma\}$ determines the projective normal of the first kind of the $\mathcal{H}$-distribution. Projective normals of the first kind $\{\mathcal{M}^\sigma\}$, $\{\mathcal{Z}^\sigma\}$, $\{S^\sigma\}$ determine bunches of the projective normals of the first kind of the $\mathcal{H}$-distribution in the differential neighborhood of the second and third orders. In the same way bunches of the projective normals of the first kind of the M-distribution were determined in the differential neighborhood of the second and third orders. We use results, which we have got in [2, 3].

2 Definition of the three-component distribution

Let us consider an $(n+1)$-dimensional affine space A_{n+1}, which is taken to a movable frame $R = \{A, \bar{e}_\mathcal{I}\}$. Differential equations of an infinitesimal transference of the frame R look as follows: $dA = \omega^\mathcal{I} \bar{e}_\mathcal{I}$, $\quad d\bar{e}_\mathcal{I} = \omega_\mathcal{I}^\mathcal{K} \bar{e}_\mathcal{K}$, where $\omega_\mathcal{I}^\mathcal{K}$, $\omega^\mathcal{I}$ are invariant forms of an affine group, which satisfy equations of the structure:

$$d\omega^\mathcal{I} = \omega^\mathcal{K} \wedge \omega_\mathcal{K}^\mathcal{I}, \qquad d\omega_\mathcal{I}^\mathcal{K} = \omega_\mathcal{I}^\mathcal{J} \wedge \omega_\mathcal{J}^\mathcal{K}.$$

Structural forms of a current point $X = A + x^\mathcal{I} \bar{e}_\mathcal{I}$ of a space A_{n+1} look as follows:

$$\triangle X^\mathcal{I} \equiv dx^\mathcal{I} + x^\mathcal{K} \omega_\mathcal{K}^\mathcal{I} + \omega^\mathcal{I}.$$

The combination of the current point X and point of the frame A leads to the following equation:

$$\triangle X^\mathcal{I} = \omega^\mathcal{I}.$$

An immobility condition of the point A is written down as follows: $\omega^\mathcal{I} = 0$.

Let the frame chosen by this way be called the frame $\tilde{R}$. Let Π_r is an r-dimensional plane in A_{n+1} be given by the following way: $\Pi_r = [A, \bar{L}_p]$, where $\bar{L}_p = \bar{e}_p + \Lambda_p^{\hat{u}} \bar{e}_{\hat{u}}$. Let m-dimensional plane Π_m be set by the following way: $\Pi_m = [A, \bar{M}_a]$, where $\bar{M}_a = \bar{e}_a + M_a^{\hat{\alpha}} \bar{e}_{\hat{\alpha}}$. A hyperplane Π_n is a set $\Pi_n = [A, \bar{T}_\sigma]$, where $\bar{T}_\sigma = \bar{e}_\sigma + H_\sigma^{n+1} \bar{e}_{n+1}$.

Definition 1 The $(n+1)$-dimensional manifolds in spaces of notion $\{\triangle \Lambda_p^{\hat{u}}, \omega^\mathcal{I}\}$, $\{\triangle M_a^{\hat{\alpha}}, \omega^\mathcal{I}\}$, $\{\triangle H_\sigma^{n+1}, \omega^\mathcal{I}\}$ which are determined by differential equations

$$\triangle \Lambda_p^{\hat{u}} = \Lambda_{p\mathcal{K}}^{\hat{u}} \omega^\mathcal{K}, \quad \triangle M_a^{\hat{\alpha}} = M_{a\mathcal{K}}^{\hat{\alpha}} \omega^\mathcal{K}, \quad \triangle H_\sigma^{n+1} = H_{\sigma\mathcal{K}}^{n+1} \omega^\mathcal{K}, \tag{1}$$

are called *distributions of the first kind accordingly of: r-dimensional linear elements (Λ-distribution), m-dimensional linear elements (M-distribution)* and *hyperplanes (H-distribution).*

Equations of the system (1) to each point A (center of distribution) are the set according to planes Π_r, Π_m, Π_n.

Let consider that manifolds (1) are distributions of tangent elements: center A belongs to planes Π_r, Π_m, Π_n. We demand, that in some area of the space A_{n+1} for any center A the following condition take place: $A \in \Pi_r \subset \Pi_m \subset \Pi_n$.

Definition 2 The three of distributions of the affine space A_{n+1}, consisting of basic distribution of the first kind r-dimensional linear elements $\Pi_r \equiv \Lambda$ (Λ-*distribution*), equipping distribution of the first kind of m-dimensional linear elements $\Pi_m \equiv M$ (M-*distribution*) and equipping distribution of the first kind of hyperplane elements $\Pi_r \equiv H$ $(r < m < n)$ (H-*distribution*) with relation of an incidence of their corresponding elements in a common center A of the following view: $A \in \Lambda \subset M \subset H$ are called $\mathcal{H}$-*distribution.*

Let us make the following canonization of the frame $\tilde{R}$: we will place vectors $\bar{e}_p$ in the plane Π_r, vectors $\bar{e}_i$ – in plane Π_m, and vectors $\bar{e}_\sigma$ – in plane Π_n. Such frame will be called the frame of the null order R^0. This definition leads to the following equations:

$$\Lambda_p^{\hat{u}} = 0, \quad M_a^{\hat{\alpha}} = 0, \quad H_\sigma^{n+1} = 0.$$

In the frame R^0 the $\mathcal{H}$-distribution is defined by the differential equations:

$$\omega_p^{\hat{u}} = \Lambda_{p\mathcal{K}}^{\hat{u}} \omega^{\mathcal{K}}, \quad \omega_i^{\hat{\alpha}} = M_{i\mathcal{K}}^{\hat{\alpha}} \omega^{\mathcal{K}}, \quad \omega_\alpha^{n+1} = H_{\alpha\mathcal{K}}^{n+1} \omega^{\mathcal{K}}.$$

According to N. Ostianu lemma it is possible partial the zero-order frame R^0 canonization, where $M_{iq}^{n+1} = 0$, $H_{\alpha q}^{n+1} = 0$. We will call it frame of the first order R^1.

In the chosen frame R^1 the manifold $\mathcal{H}$ is determined by the following system of differential equations:

$$\omega_p^{\hat{u}} = \Lambda_{p\mathcal{K}}^{\hat{u}} \omega^{\mathcal{K}}, \quad \omega_i^{n+1} = M_{i\hat{u}}^{n+1} \omega^{\hat{u}},$$

$$\omega_i^\alpha = M_{i\mathcal{K}}^\alpha \omega^{\mathcal{K}}, \quad \omega_\alpha^{n+1} = H_{\alpha\hat{u}}^{n+1} \omega^{\hat{u}}, \quad \omega_u^p = A_{u\mathcal{K}}^p \omega^{\mathcal{K}}.$$

3 Tensor of inholonomicity of $\mathcal{H}$-distribution

It's easy to show, that geometry of three-component distributions can be used for studying geometry of regular and degenerate hyperzones, zones, hyperzone distributions, surfaces of full and not full range, tangent equipped surfaces in affine spaces. For example, we will suppose, that the $\mathcal{H}$-distribution is holonomic, that is the basic distribution Λ is holonomic. System of differential equations $\omega_0^{\hat{u}} = \Lambda_p^{\hat{u}} \omega_0^p$, which is associated with basic Λ-distribution, is quite integrable if and only if, when the tensor of the first order

$$r_{pq}^{\hat{u}} = \frac{1}{2}(\Lambda_{pq}^{\hat{u}} - \Lambda_{qp}^{\hat{u}})$$

turns into zero.

Tensor $\{r^{\hat{u}}_{pq}\}$ will be called tensor of the inholonomicity of the $\mathcal{H}$-distribution. The basic Λ-distribution determines $(n-r+1)$-parametric assemblage of r-dimensional surfaces V_r (planes Λ are rounded by r-dimensional surfaces of $(n-r+1)$-parametric assemblage).

In the time of displacement of center A along a fixed surface V_r, differential equations, which determine the $\mathcal{H}$-distribution relatively the frame $\tilde{R}$

$$\omega^{\hat{u}}_0 = \Lambda^{\hat{u}}_q \omega^q, \qquad \triangle\Lambda^{\hat{u}}_p = (\Lambda^{\hat{u}}_{pq} + \Lambda^{\hat{u}}_{p\hat{v}}\Lambda^{\hat{v}}_q)\omega^q,$$
$$\triangle M^{\hat{\alpha}}_i = (M^{\hat{\alpha}}_{iq} + M^{\hat{\alpha}}_{i\hat{v}}\Lambda^{\hat{v}}_q)\omega^q, \qquad \triangle H^{n+1}_{\alpha} = (H^{n+1}_{\alpha q} + H^{n+1}_{\alpha\hat{v}}\Lambda^{\hat{v}}_q)\omega^q$$

are differential equations of r-dimensional zone $V_{r(m)}$ of the order m [7, 8] equipped by a field of hyperplanes H. A geometrical object $\{H^{n+1}_r\}$ (object H) is the fundamental equipping object of a zone $V_{r(m)}$.

Following G. F. Laptev [5], the zone $V_{r(m)}$, on which the field of the fundamental equipping object H is set, we will call an equipped zone $V_{r(m)}$ and we will disignate as $V_{r(m)}(H)$.

Let note, that relatively of the frame R^0, which is adapted the fields of the planes Λ, M, H, differential equations of the manifold $V_{r(m)}(H)$ have more simple form:

$$\omega^{\hat{u}}_0 = 0, \quad \omega^{\hat{u}}_p = \Lambda^{\hat{u}}_{pq}\omega^q, \quad \Lambda^{\hat{u}}_{pq} = \Lambda^{\hat{u}}_{qp}, \tag{2}$$

$$\omega^{\hat{\alpha}}_i = M^{\hat{\alpha}}_{iq}\omega^q, \tag{3}$$

$$\omega^{n+1}_{\alpha} = H^{n+1}_{\alpha q}\omega^q, \tag{4}$$

where equations (2), (3) are analogous to equations of the zone $V_{r(m)}$, which are discussed in the work of M.M. Pohila [7]. Equations (4) characterize the equipment of the zone $V_{r(m)}$ by the field of hyperzones H.

Thus, a transformation of a tensor $\{r^{\hat{u}}_{pq}\}$ to zero is the condition, where the space A_{n+1} desintegrates to $(n-r+1)$-parametric assemblage of equipped zones $V_{r(m)}(H)$. So plane $\Lambda(A)$ in its center A is the tangent plane of the surface V_r (V_r is basic surface of equipped zone $V_{r(m)}(H)$), plane $(M(A))$ is the tangent m-plane of the basic surface in the center A. The hyperplane $H(A)$ is the equipping plane of the zone $V_{r(m)}(H)$. At that time we suppose, that the condition of the incidence of planes Λ, M, H is executed .

On the other hand, equations (2), (4) determine in the frame R^0 the hyperplane H_r [9], and equations (3) characterize an equipment of the hyperzone H_r by field of planes M. This field of planes M is determined by the field of the geometrical object $\{M^{\hat{\alpha}}_a\}$ – field of the fundamental equipping object of the hyperzone H_r. We will disignate the hyperzone H_r, which is equipped by the field of planes M, as $H_r(M)$. Thus, the theory of the three-component $\mathcal{H}$-distribution is a generalization of theories of the regular hyperzone H_r and the zone $V_{r(m)}(H)$ of the affine space.

4 Tensor of inholonomicity of equipping distributions

Let consider the system of differential equations

$$\omega_0^{\hat{\alpha}} = M_a^{\hat{\alpha}} \omega^a, \tag{5}$$

which is associated with the M-distribution. This system is fully integrable if and only if, when the tensor of inholonomicity $\{r_{ab}^{\hat{\alpha}}\}$ of the equipping M-distribution

$$r_{ab}^{\hat{\alpha}} = \frac{1}{2}(M_{ab}^{\hat{\alpha}} - M_{ba}^{\hat{\alpha}})$$

equals to zero.

At $r_{ab}^{\hat{\alpha}} = 0$ the system (5) determines $(n-m+1)$-parametric assemblage of the m-dimensional surfaces V_m – m-dimensional integral manifolds. One and only one such manifold passes across each point of the area of such manifolds (planes M are rounded by m-dimensional surfaces V_m of $(n-m+1)$-parametric assemblage).

In the time of displacement of the center A along the fixed surface V_m equations, that determine the $\mathcal{H}$-distribution, define the tangent r-equipped surface $V_{m(r)}$ [4], which is equipped by the field of tangent hyperplanes H. Actually, from system, which consists from differential equations (5) and equations, which determines the $\mathcal{H}$-distribution, we can pick out a subsystem

$$\omega^{\hat{\alpha}} = M_b^{\hat{\alpha}} \omega^b, \quad \triangle M_a^{\hat{\alpha}} = M_{ab}^{\hat{\alpha}} \omega^b, \qquad \triangle \Lambda_p^i = \Lambda_{pb}^i \omega^b, \quad M_{[ab]}^{\hat{\alpha}} = 0.$$

This subsystem determines the tangent r-equipped surface $V_{m,r}$ [4]. In this case the geometrical object $\{H_\tau^{n+1}\}$ (object H) is the fundamental equipping object of the tangent r-equipped surface $V_{m,r}$. Such tangent r-equipped surfaces $V_{m,r}$, which are equipped by the field of tangent hyperplanes , we will disignate as $V_{m,r}(H)$. Thus, if the tensor of the inholonomicity $\{r_{ab}^{\hat{\alpha}}\}$ of the equipping M-distribution equals to zero, so the space A_{n+1} disintegrates to $(n-m+1)$-parametric assemblage of manifolds look as follows $V_{m,r}(H)$.

On the other hand, the $\mathcal{H}$-distribution for which $r_{ab}^{\hat{\alpha}} = 0$ can be interpreted like the hyperzone H_m, which is equipped by the field of tangent planes Λ. Hence, geometry of the $\mathcal{H}$-distribution of the affine space, naturally, is richer than geometry of tangent r-equipped surfaces and geometry of hyperzones H_m of the affine space, because it consists of a constructions, which don't have any sense for the latter. Also, geometry of the $\mathcal{H}$-distribution can be used for studying of degenerated hyperzones [6] and surfaces [1].

The system of differential equations

$$\omega^{n+1} = H_\tau^{n+1} \omega^\tau, \tag{6}$$

which is associated with the equipping distribution of hyperplanes H (H-distribution), is fully integrable if and only if, when the tensor of the first order

$$r_{\tau\sigma}^{n+1} = \frac{1}{2}(H_{\tau\sigma}^{n+1} - H_{\sigma\tau}^{n+1})$$

turns into zero.

On the condition, that the tensor of the inholonomicity $\{r^{n+1}_{\tau\sigma}\}$ of the equipping H-distribution equals to zero, the system (6) determines one-parametric assemblage of hypersurfaces V_n (planes H are rounded by hypersurfaces V_n of one-parametric assemblage).

In the time of a displacement of the center A along the fixed surface V_n equations, which determine the $\mathcal{H}$-distribution, represent equations of the hypersurface, which is equipped by fields of geometrical objects $\{\Lambda^{\hat{u}}_p\}$ and $\{M^{\hat{\alpha}}_a\}$ (fields of planes Λ and M, where $\Lambda \subset M$). Hence, the theory of the three-component $\mathcal{H}$-distribution is also the generalization of the theory of hypersurfaces of the affine space.

5 Normals of the equipping distributions

Quasi-tensors were constructed in the second differential neighborhood:

$$B^p = -\frac{1}{r+2}a^{pq}B_q, \quad B^i = -\frac{1}{m-r+2}a^{ji}B_j - \frac{1}{m-r+2}\Lambda_{pk}a^{ki}B^p,$$

$$B^\alpha = -\frac{1}{m-r+2}(H^{\gamma\alpha}B_\gamma + \Lambda_{p\gamma}H^{\gamma\alpha}B^p + M_{i\gamma}H^{\gamma\alpha}B^i),$$

$$\nabla B^p - B^p\omega^{n+1}_{n+1} + \omega^p_{n+1} = B^p_{\mathcal{K}}\omega^{\mathcal{K}},$$

$$\nabla B^i - B^i\omega^{n+1}_{n+1} + \omega^i_{n+1} = B^i_{\mathcal{K}}\omega^{\mathcal{K}}, \qquad \nabla B^\alpha - B^\alpha\omega^{n+1}_{n+1} + \omega^\alpha_{n+1} = B^\alpha_{\mathcal{K}}\omega^{\mathcal{K}}.$$

The geometrical object $\{B^\sigma\}$ determines the normal of the first kind of the $\mathcal{H}$-distribution by an inner invariant method. The normal B coincides with the Blaschke normal in case of the hyperplane distribution. Affine normals of the first kind B_{n-r+1}, B_{n-m+1} of the Λ-distribution and of the M-distribution accordingly are determined in the same way.

Quasi-tensors were constructed in the differential neighborhood of the second order:

$$\gamma^p = -\frac{1}{r+2}\Lambda^{pq}\gamma_q, \quad \gamma^i = -\frac{1}{m-r+2}M^{ji}\gamma_j + \frac{m-r-2}{m-r+2}\Lambda_{pk}M^{ki}\gamma^p,$$

$$\gamma^\alpha = -\frac{1}{n-m+2}(H^{\alpha\beta}\gamma_\beta + \frac{n-m-2}{n-m+2}(\Lambda_{p\gamma}H^{\alpha\gamma}\gamma^p + M_{i\gamma}H^{\alpha\gamma}\gamma^i),$$

$$\nabla\gamma^p - \gamma^p\omega^{n+1}_{n+1} + \omega^p_{n+1} = \gamma^p_{\mathcal{K}}\omega^{\mathcal{K}},$$

$$\nabla\gamma^i - \gamma^i\omega^{n+1}_{n+1} + \omega^i_{n+1} = \gamma^i_{\mathcal{K}}\omega^{\mathcal{K}}, \quad \nabla\gamma^\alpha - \gamma^\alpha\omega^{n+1}_{n+1} + \omega^\alpha_{n+1} = \gamma^\alpha_{\mathcal{K}}\omega^{\mathcal{K}}.$$

Fields of the geometrical objects $\{\gamma^a\}, \{\gamma^\sigma\}$ determine fields of the normals of the first kind of the equipping M-distribution, of the equipping H-distribution accordingly.

The quasi-tensor $\{\mathcal{M}^\sigma\}$:

$$\mathcal{M}^\sigma = \frac{1}{2}(\mathcal{L}^\sigma + \gamma^\sigma), \quad \nabla\mathcal{M}^\sigma - \mathcal{M}^\sigma\omega^{n+1}_{n+1} + \omega^\sigma_{n+1} = \mathcal{M}^\sigma_{\mathcal{K}}\omega_{\mathcal{K}},$$

determines the normal of the first kind of the $\mathcal{H}$-distribution in the differential neighborhood of the second order, which is invariant relatively of the projective group of the transformations.

The normal $\{\mathcal{M}^\sigma\}$ is the Mihajlesku normal of the first kind of the hyperplane distribution of the affine space.

The field of the affine normal of the first kind of the H-distribution is determined by the object $\{\hat{B}^\tau\}$ in the differential neighborhood of the third order:

$$\hat{B}^\tau = H^{\rho\tau}\hat{B}_\rho, \quad \nabla\hat{B}^\tau - \hat{B}^\tau\omega^{n+1}_{n+1} + \omega^\tau_{n+1} = \hat{B}^\tau_{\mathcal{K}}\omega^{\mathcal{K}}.$$

The quasi-tensor $\{\mathcal{Z}^\sigma\}$ of the third order:

$$\mathcal{Z}^\sigma = \hat{B}^\sigma + \hat{h}^\sigma, \quad \nabla\mathcal{Z}^\sigma - \mathcal{Z}^\sigma\omega^{n+1}_{n+1} + \omega^\sigma_{n+1} = \mathcal{Z}^\sigma_{\mathcal{K}}\omega^{\mathcal{K}},$$

determines the projective normal—analog of the Fubiny's normal for the H-distribution in each center of the forming element of the $\mathcal{H}$-distribution.

The object $\{\mathcal{Z}^a\}$ determines the projective normal of the first kind of the M-distribution.

The object $\{S^a\}$, where

$$S^\sigma = -\frac{1}{2}(H_{\rho n+1} + \frac{1}{n+2}p_\rho)H^{\rho\sigma}, \quad \nabla S^\sigma - S^\sigma\omega^{n+1}_{n+1} + \omega^\sigma_{n+1} = S^\sigma_{\mathcal{K}}\omega_{\mathcal{K}},$$

determines the projective normal of the first kind of the M-distribution.

Theorem 1 *The projective normals of the first kind $\mathcal{M}$, $\mathcal{Z}$, S determine bunches of the projective normals of the first kind of the $\mathcal{H}$-distribution:*
a) in the differential neighborhood of the second order

$$\tilde{M}^\sigma(\mathcal{E}) = \mathcal{M}^\sigma - \mathcal{E}(\mathcal{M}^\sigma - S^\sigma);$$

b) in the differential neighborhood of the third order

$$\hat{\Phi}^\sigma(\mathcal{E}) = \mathcal{Z}^\sigma - \mathcal{E}(\mathcal{Z}^\sigma - \mathcal{M}^\sigma), \quad \hat{\mathcal{Z}}^\sigma(\mathcal{E}) = \mathcal{Z}^\sigma - \mathcal{E}(\mathcal{Z}^\sigma - S^\sigma),$$

where $\mathcal{E}$ – absolute invariant.

These normals determine bunches of the projective normals of the first kind of the equipping M-distribution:
a) in the differential neighborhood of the second order

$$\tilde{M}^a(\mathcal{E}) = \mathcal{M}^i - \mathcal{E}(\mathcal{M}^a - S^a).$$

b) in the differential neighborhood of the third order $\{\hat{\Phi}^a(\mathcal{E})\}$, $\{\hat{\mathcal{M}}^a(\mathcal{E})\}$.

References

[1] Amisheva, N. V.: *Some questions of affine geometry of the tangential degenerated surface.* Kemerov Univ., VINITI, 3826-80, 1980, 17 pp. (in Russian).

[2] Grebenjuk, M. F.: *For geometry of $H(M(\Lambda))$-distribution of affine space.* Kaliningrad Univ., Kaliningrad, VINITI, 8204-1388, 1988, 17 pp.

[3] Grebenjuk, M. F.: *Fields of geometrical objects of three-component distribution of affine space A_{n+1}.* Diff. Geometry of Manifolds of Figures: Inter-Univ. subject collection of scientific works, Kaliningrad Univ., 1987, Issue 18, 21–24.

[4] Dombrovskyj, P. F.: *To geometry of tangent equipped surfaces in P_n.* Works of Geometrical Seminar, VINITI, 1975, v. 6, 171–188.

[5] Laptev, G. F.: *Differential geometry of immersed manifolds: Theoretical and group method of differential-geometrical researches.* Works of Moscow Mathematical Society, 1953, Vol. 2, 275–382.

[6] Popov, U. I.: *Inner equipment of degenerated m-dimensional hyperstripe H_m^r of range r of many-dimensional projective space.* Diff. Geometry of Manifolds of Figures, Issue 6, Kaliningrad, 1975, 102–142.

[7] Pohila, M. M.: *Geometrical images, which are associated with many-dimensional stripe of projective space.* Abstr. of Rep. of 5th Baltic Geom. Conf., Druskininkaj, 1978, p. 70.

[8] Pohila, M. M.: *Generalized many-dimensional stripes.* Abstr. of Rep. of 6th Conf. of Sov. Union on Modern Problems of Geometry. Vilnius, 1975, 198–199.

[9] Stoljarov, A. B.: *About fundamental objects of regular hyperstripe.* News of Univ. Math., 1975, a 10, 97–99.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 43–50

Estimation of the First Order Parameters in the Twoepoch Linear Model

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 25, 2007)

Abstract

The linear regression model, where the mean value parameters are divided into stable and nonstable part in each of both epochs of measurement, is considered in this paper. Then, equivalent formulas of the best linear unbiased estimators of this parameters in both epochs using partitioned matrix inverse are derived.

Key words: Twoepoch regression model; best linear unbiased estimators of the first order parameters.

2000 Mathematics Subject Classification: 62J05

1 Motivation

Many real problems, convenient for linear statistical modelling, typically from the field of geodesy, are investigated in more than one epoch to obtain better estimations of the unknown mean value parameters (see [5], [6], [7] and also [3], [4]). From the principle of concrete situation we suppose that some mean value (first order) parameters does not change their values during epochs (i.e. *stable* parameters) contrary to so called *nonstable* parameters. There are also many other problems, convenient for linear modelling, where the role of stable and nonstable parameters is not exactly given. Dividing the first order parameter to the stable and nonstable part we achieve better fit of the corresponding linear model to the concrete situation. As an example we choose the problem from [1], p. 90. A dependence of petrol consumption in litres to the rate in kilometers per hour by a certain car marque was investigated. The quadratical trend was chosen as the most convenient to describe the dependence. Let us have a new car and an older car. Then it is clear that the petrol consumption will be generally greater by the old car but the quadratical dependence to the rate by both cars will stay approximately unchanged. If we adopt this situation as an example of twoepoch measurement (the most occured in simpler problems), we can select the linear and quadratical term parameters to be stable and absolute term parameter to be nonstable first order parameters. Estimation of the stable and nonstable parameters using both epochs together give us better information about the dependence and increasing consumption than estimations in single epochs separately.

Let us formalize the performed considerations. The results of the measurement could be described as

$$Y_{1i} = \beta_1 x_{1i}^2 + \beta_2 x_{1i} + \gamma_1 + \varepsilon_{1i}, \quad i = 1, \dots, n_1$$

in the first epoch and

$$Y_{2i} = \beta_1 x_{2i}^2 + \beta_2 x_{2i} + \gamma_2 + \varepsilon_{2i}, \quad i = 1, \dots, n_2$$

in the second epoch of measurement. Let us consider the $n_1 + n_2$ dimensional observation vector $\mathbf{Y} = (\mathbf{Y}_1', \mathbf{Y}_2')'$ after the second epoch of the measurent. The model described above could be generally rewritten in the form

$$\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{W}_1 & \mathbf{0} \\ \mathbf{X}_2 & \mathbf{0} & \mathbf{W}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \gamma_1 \\ \gamma_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}. \tag{1}$$

The (design) matrices $\mathbf{X}_1, \mathbf{X}_2, \mathbf{W}_1, \mathbf{W}_2$ are known; $\boldsymbol{\beta} \in \mathbb{R}^r$ is a vector of the useful stable parameters, the same in both epochs; $(\gamma_1', \gamma_2')' \in \mathbb{R}^{s_1+s_2}$ is a vector of nonstable parameters in the first and the second epoch of measurement.

With respect to formerly mentioned, let us consider the linear model (1), called the twoepoch model with the stable and nonstable parameters. We suppose that

1. $\mathrm{E}(\mathbf{Y}_1) = \mathbf{X}_1\boldsymbol{\beta} + \mathbf{W}_1\boldsymbol{\gamma}_1$, $\mathrm{E}(\mathbf{Y}_2) = \mathbf{X}_2\boldsymbol{\beta} + \mathbf{W}_2\boldsymbol{\gamma}_2$, $\forall\boldsymbol{\beta} \in \mathbb{R}^r, \forall\boldsymbol{\gamma}_1 \in \mathbb{R}^{s_1}, \forall\boldsymbol{\gamma}_2 \in \mathbb{R}^{s_2}$;
2. $\mathrm{var}\left[\begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}\right] = \begin{pmatrix} \sigma^2\boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \sigma^2\boldsymbol{\Sigma}_2 \end{pmatrix}$, $\sigma^2 > 0$ is unknown parameter;
3. the matrix $\boldsymbol{\Sigma}_i$ is not a function of the vector $(\boldsymbol{\beta}', \boldsymbol{\gamma}_i')'$ for $i = 1, 2$.

If the matrix

$$\begin{pmatrix} \sigma^2\boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \sigma^2\boldsymbol{\Sigma}_2 \end{pmatrix}$$

is positive definite (p.d.) and rank

$$r\left[\begin{pmatrix} \mathbf{X}_1 & \mathbf{W}_1 & \mathbf{0} \\ \mathbf{X}_2 & \mathbf{0} & \mathbf{W}_2 \end{pmatrix}\right] = r + s_1 + s_2 < n_1 + n_2,$$

the model is said to be *regular* (see [5], p. 13).

The mentioned problem produces

$$\mathbf{X}_1 = \begin{pmatrix} x_{11}^2 & x_{11} \\ \vdots & \vdots \\ x_{1n_1}^2 & x_{1n_1} \end{pmatrix}, \quad \mathbf{X}_2 = \begin{pmatrix} x_{21}^2 & x_{21} \\ \vdots & \vdots \\ x_{2n_2}^2 & x_{2n_2} \end{pmatrix},$$

$$\mathbf{W}_1 = (\overbrace{1,\dots,1}^{n_1-\text{times}})' = \mathbf{1}_{n_1}, \quad \mathbf{W}_2 = (\overbrace{1,\dots,1}^{n_2-\text{times}})' = \mathbf{1}_{n_2},$$

$$\boldsymbol{\beta} = (\beta_1, \beta_2)', \ \boldsymbol{\gamma}_1 = \gamma_1, \ \boldsymbol{\gamma}_2 = \gamma_2.$$

In the case of positive definitness of the matrices $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$ is the model evidently regular.

2 Notation and auxiliary statements

Let us summarize the notation, used throughout the paper:

$\mathbb{R}^n$	the space of all n-dimensional real vectors;
$\mathbf{u}$, $\mathbf{A}$	the real column vector, the real matrix;
$\mathbf{A}', r(\mathbf{A})$	the transpose, the rank of the matrix $\mathbf{A}$;
$\mathcal{M}(\mathbf{A})$, $\mathrm{Ker}(\mathbf{A})$	the range, the null space of the matrix $\mathbf{A}$;
$\mathbf{P}_A$	the orthogonal projector onto $\mathcal{M}(\mathbf{A})$ (in Euclidean sense);
$\mathbf{M}_A = \mathbf{I} - \mathbf{P}_A$	the orthogonal projector onto $\mathcal{M}^{\perp}(\mathbf{A}) = \mathrm{Ker}(\mathbf{A}')$;
$\mathbf{I}_k$	the $k \times k$ identity matrix;
$\mathbf{0}_{m,n}$	the $m \times n$ null matrix;
$\mathbf{1}_k$	$= (1,\dots,1)' \in \mathbb{R}^k$;
$F_{m,n}$	random variable with F distribution with m and n degrees of freedom;
$F_{m,n}(1-\alpha)$	$(1-\alpha)$-quantile of this distribution.

Lemma 1 *Inverse of partitioned p.d. matrix*

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} & \mathbf{D} \\ \mathbf{B}' & \mathbf{C} & \mathbf{0} \\ \mathbf{D}' & \mathbf{0} & \mathbf{E} \end{pmatrix} \quad \textit{is equal to} \quad \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} & \mathbf{Q}_{13} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} & \mathbf{Q}_{23} \\ \mathbf{Q}_{31} & \mathbf{Q}_{32} & \mathbf{Q}_{33} \end{pmatrix},$$

where ($\mathbf{Q}_{21} = (\mathbf{Q}_{12})'$, $\mathbf{Q}_{31} = (\mathbf{Q}_{13})'$, $\mathbf{Q}_{32} = (\mathbf{Q}_{23})'$)

$$\begin{pmatrix} \mathbf{Q}_{11} & -\mathbf{Q}_{11}\mathbf{B}\mathbf{C}^{-1} & -\mathbf{Q}_{11}\mathbf{D}\mathbf{E}^{-1} \\ -\mathbf{C}^{-1}\mathbf{B}'\mathbf{Q}_{11} & \mathbf{C}^{-1} + \mathbf{C}^{-1}\mathbf{B}'\mathbf{Q}_{11}\mathbf{B}\mathbf{C}^{-1} & -\mathbf{Q}_{21}\mathbf{D}\mathbf{E}^{-1} \\ -\mathbf{E}^{-1}\mathbf{D}'\mathbf{Q}_{11} & -\mathbf{E}^{-1}\mathbf{D}'\mathbf{Q}_{12} & \mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{D}'\mathbf{Q}_{11}\mathbf{D}\mathbf{E}^{-1} \end{pmatrix},$$

with

$$\mathbf{Q}_{11} = (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}' - \mathbf{D}\mathbf{E}^{-1}\mathbf{D}')^{-1}$$

(Version I); equivalently

$$\begin{aligned}
\mathbf{Q}_{11} &= (\mathbf{A} - \mathbf{D}\mathbf{E}^{-1}\mathbf{D}')^{-1} + (\mathbf{A} - \mathbf{D}\mathbf{E}^{-1}\mathbf{D}')^{-1}\mathbf{B}\mathbf{Q}_{22}\mathbf{B}'(\mathbf{A} - \mathbf{D}\mathbf{E}^{-1}\mathbf{D}')^{-1},\\
\mathbf{Q}_{12} &= -(\mathbf{A} - \mathbf{D}\mathbf{E}^{-1}\mathbf{D}')^{-1}\mathbf{B}\mathbf{Q}_{22},\\
\mathbf{Q}_{13} &= -\mathbf{Q}_{11}\mathbf{D}\mathbf{E}^{-1},\\
\mathbf{Q}_{22} &= [\mathbf{C} - \mathbf{B}'(\mathbf{A} - \mathbf{D}\mathbf{E}^{-1}\mathbf{D}')^{-1}\mathbf{B}]^{-1},\\
\mathbf{Q}_{23} &= -\mathbf{Q}_{21}\mathbf{D}\mathbf{E}^{-1},\\
\mathbf{Q}_{33} &= \mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{D}'\mathbf{Q}_{11}\mathbf{D}\mathbf{E}^{-1},
\end{aligned}$$

(Version II) and equivalently

$$\begin{aligned}
\mathbf{Q}_{11} &= (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1} + (\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{D}\mathbf{Q}_{33}\mathbf{D}'(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1},\\
\mathbf{Q}_{12} &= -\mathbf{Q}_{11}\mathbf{B}\mathbf{C}^{-1},\\
\mathbf{Q}_{13} &= -(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{D}\mathbf{Q}_{33},\\
\mathbf{Q}_{22} &= \mathbf{C}^{-1} + \mathbf{C}^{-1}\mathbf{B}'\mathbf{Q}_{11}\mathbf{B}\mathbf{C}^{-1},\\
\mathbf{Q}_{23} &= -\mathbf{C}^{-1}\mathbf{B}'\mathbf{Q}_{13},\\
\mathbf{Q}_{33} &= [\mathbf{E} - \mathbf{D}'(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}')^{-1}\mathbf{D}]^{-1},
\end{aligned}$$

(Version III).

Proof The statement can be proved directly using [3, Theorem 1 and Remark 1–3], with proper Rohde formula (see [2, Theorem 8.5.11, p. 99]). □

It can be easily shown that there exist five versions of such inverse alltogether but only above mentioned are convenient for our later purposes.

Lemma 2 *Let us consider regular linear model*

$$(\mathbf{Y},\, \mathbf{X}\boldsymbol{\delta}, \sigma^2\mathbf{V}),\ \ \boldsymbol{\delta} \in \mathbb{R}^k,\, \sigma^2 > 0 \tag{2}$$

where $\mathbf{Y}$ *is* n*-dimensional normally distributed observation vector,* $\mathbf{X}$ $n \times k$ *design matrix (*$r(\mathbf{X}) = k < n$*) and* $\boldsymbol{\Sigma}$ *is known p.d. variance matrix of the type* $n \times n$*. Then the best linear unbiased estimator (BLUE) of the vector* $\boldsymbol{\delta}$ *equals*

$$\widehat{\boldsymbol{\delta}} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} \tag{3}$$

with the variance matrix

$$\mathrm{var}(\widehat{\boldsymbol{\delta}}) = \sigma^2(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}. \tag{4}$$

Unbiased and invariant estimator of σ^2 *is*

$$\widehat{\sigma^2} = (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\delta}})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\delta}})/(n-k). \tag{5}$$

Let null hypothesis about parameter $\boldsymbol{\delta}$ *is*

$$H_0 : \mathbf{H}\boldsymbol{\delta} + \mathbf{h} = \mathbf{0}, \tag{6}$$

where $\mathbf{H}$ *is* $q \times k$ *matrix with* $r(\mathbf{H}) = q < k$, *and alternative hypothesis*

$$H_a : \mathbf{H}\boldsymbol{\delta} + \mathbf{h} \neq \mathbf{0}. \tag{7}$$

Then, in premise of validity of H_0, *the random variable* F

$$F = \frac{(\mathbf{H}\widehat{\boldsymbol{\delta}} + \mathbf{h})'[\mathbf{H}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{H}']^{-1}(\mathbf{H}\widehat{\boldsymbol{\delta}} + \mathbf{h})}{\widehat{\sigma^2}} \tag{8}$$

is $F_{q,n-k}$ *distributed.*

Proof see [5, p. 13, Theorem 1.1.1 and p. 54, Theorem 1.8.9]. □

3 Best linear unbiased estimators

Let the twoepoch linear regression model (1) be given. This model arises by sequential realizations of the linear partial regression models,

$$\mathbf{Y}_1 = (\mathbf{X}_1, \mathbf{W}_1)\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma}_1 \end{pmatrix} + \boldsymbol{\varepsilon}_1,\ \mathrm{var}(\mathbf{Y}_1) = \sigma^2\boldsymbol{\Sigma}_1 \tag{9}$$

and

$$\mathbf{Y}_2 = (\mathbf{X}_2, \mathbf{W}_2)\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma}_2 \end{pmatrix} + \boldsymbol{\varepsilon}_2,\ \mathrm{var}(\mathbf{Y}_2) = \sigma^2\boldsymbol{\Sigma}_2, \tag{10}$$

representing the model of the measurement within the first and second epoch, respectively. Let us remark that the parameter σ^2 is supposed to be the same in both epochs. Althougt this condition could be too restricting in some cases we adopt it to make the computations easily. Moreover there are many situations, mainly in simpler problems, where this condition is acceptable. The further derived formulas are more complicated that in general case but they show the structure of the twoepoch model what is useful in many applications. In addition, thanks to Lemma 1 the formulas in the twoepoch model will be derived in a friendly methodical way. The next theorem can be used to verify the stableness of the first order parameters.

Theorem 1 *The BLUE of the parameters* $\boldsymbol{\beta}$, $\boldsymbol{\gamma}_1$ *and* $\boldsymbol{\gamma}_2$ *in the single first and second epoch (models (9) and (10)) are for* $i = 1, 2$

$$\begin{aligned} \widehat{\boldsymbol{\beta}}^{(i)} &= (\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{M}_{W_i}^{\Sigma_i^{-1}}\mathbf{X}_i)^{-1}\mathbf{X}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{M}_{W_i}^{\Sigma_i^{-1}}\mathbf{Y}_i, \\ \widehat{\boldsymbol{\gamma}_i}^{(i)} &= (\mathbf{W}_i'\boldsymbol{\Sigma}_i^{-1}\mathbf{W}_i)^{-1}\mathbf{W}_i'\boldsymbol{\Sigma}_i^{-1}(\mathbf{Y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}^{(i)}), \end{aligned}$$

equivalently

$$\begin{aligned}\widehat{\boldsymbol\beta}^{(i)} &= (\mathbf{X}_i'\boldsymbol\Sigma_i^{-1}\mathbf{X}_i)^{-1}\mathbf{X}_i'\boldsymbol\Sigma_i^{-1}(\mathbf{Y}_i - \mathbf{W}_i\widehat{\gamma}_i^{(i)}),\\ \widehat{\gamma}_i^{(i)} &= (\mathbf{W}_i'\boldsymbol\Sigma_i^{-1}\mathbf{M}_{X_i}^{\Sigma_i^{-1}}\mathbf{W}_i)^{-1}\mathbf{W}_i'\boldsymbol\Sigma_i^{-1}\mathbf{M}_{X_i}^{\Sigma_i^{-1}}\mathbf{Y}_i,\end{aligned}$$

where $\boldsymbol\Sigma_i^{-1}\mathbf{M}_{W_i}^{\Sigma_i^{-1}} = \boldsymbol\Sigma_i^{-1} - \boldsymbol\Sigma_i^{-1}\mathbf{W}_i(\mathbf{W}_i'\boldsymbol\Sigma_i^{-1}\mathbf{W}_i)^{-1}\mathbf{W}_i'\boldsymbol\Sigma_i^{-1}$.

Proof see [5, p. 369, Theorem 9.1.2]. □

Theorem 2 *In the regular twoepoch linear model (1), the BLUE of the parameters* $\boldsymbol\beta, \boldsymbol\gamma_1, \boldsymbol\gamma_2$ *in both epochs equals*

$$\begin{aligned}\widehat{\boldsymbol\beta} &= (\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{M}_{W_1}^{\Sigma_1^{-1}}\mathbf{X}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{M}_{W_2}^{\Sigma_2^{-1}}\mathbf{X}_2)^{-1}\\ &\quad\times(\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{M}_{W_1}^{\Sigma_1^{-1}}\mathbf{Y}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{M}_{W_2}^{\Sigma_2^{-1}}\mathbf{Y}_2),\\ \widehat{\gamma_1} &= (\mathbf{W}_1'\boldsymbol\Sigma_1^{-1}\mathbf{W}_1)^{-1}\mathbf{W}_1'\boldsymbol\Sigma_1^{-1}(\mathbf{Y}_1 - \mathbf{X}_1\widehat{\boldsymbol\beta}),\\ \widehat{\gamma_2} &= (\mathbf{W}_2'\boldsymbol\Sigma_2^{-1}\mathbf{W}_2)^{-1}\mathbf{W}_2'\boldsymbol\Sigma_2^{-1}(\mathbf{Y}_2 - \mathbf{X}_2\widehat{\boldsymbol\beta}),\end{aligned}$$

(Version I); equivalently

$$\begin{aligned}\widehat{\boldsymbol\beta} &= (\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{X}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{M}_{W_2}^{\Sigma_2^{-1}}\mathbf{X}_2)^{-1}\\ &\quad\times[\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}(\mathbf{Y}_1 - \mathbf{W}_1\widehat{\gamma_1}) + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{M}_{W_2}^{\Sigma_2^{-1}}\mathbf{Y}_2],\\ \widehat{\gamma_1} &= [\mathbf{W}_1'\boldsymbol\Sigma_1^{-1}\mathbf{W}_1 - \mathbf{W}_1'\boldsymbol\Sigma_1^{-1}\mathbf{X}_1(\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{X}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{M}_{W_2}^{\Sigma_2^{-1}}\mathbf{X}_2)^{-1}\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{W}_1]^{-1}\\ &\quad\times[\mathbf{W}_1'\boldsymbol\Sigma_1^{-1}\mathbf{Y}_1 - \mathbf{W}_1'\boldsymbol\Sigma_1^{-1}\mathbf{X}_1(\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{X}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{M}_{W_2}^{\Sigma_2^{-1}}\mathbf{X}_2)^{-1}\\ &\quad\times(\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{Y}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{M}_{W_2}^{\Sigma_2^{-1}}\mathbf{Y}_2)],\\ \widehat{\gamma_2} &= (\mathbf{W}_2'\boldsymbol\Sigma_2^{-1}\mathbf{W}_2)^{-1}\mathbf{W}_2'\boldsymbol\Sigma_2^{-1}(\mathbf{Y}_2 - \mathbf{X}_2\widehat{\boldsymbol\beta}),\end{aligned}$$

(Version II); equivalently

$$\begin{aligned}\widehat{\boldsymbol\beta} &= (\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{M}_{W_1}^{\Sigma_1^{-1}}\mathbf{X}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{X}_2)^{-1}\\ &\quad\times[\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{M}_{W_1}^{\Sigma_1^{-1}}\mathbf{Y}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}(\mathbf{Y}_2 - \mathbf{W}_2\widehat{\gamma_2})],\\ \widehat{\gamma_1} &= (\mathbf{W}_1'\boldsymbol\Sigma_1^{-1}\mathbf{W}_1)^{-1}\mathbf{W}_1'\boldsymbol\Sigma_1^{-1}(\mathbf{Y}_1 - \mathbf{X}_1\widehat{\boldsymbol\beta}),\\ \widehat{\gamma_2} &= [\mathbf{W}_2'\boldsymbol\Sigma_2^{-1}\mathbf{W}_2 - \mathbf{W}_2'\boldsymbol\Sigma_2^{-1}\mathbf{X}_2(\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{M}_{W_1}^{\Sigma_1^{-1}}\mathbf{X}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{X}_2)^{-1}\mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{W}_2]^{-1}\\ &\quad\times[\mathbf{W}_2'\boldsymbol\Sigma_2^{-1}\mathbf{Y}_2 - \mathbf{W}_2'\boldsymbol\Sigma_2^{-1}\mathbf{X}_2(\mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{X}_2 + \mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{M}_{W_1}^{\Sigma_1^{-1}}\mathbf{X}_1)^{-1}\\ &\quad\times(\mathbf{X}_1'\boldsymbol\Sigma_1^{-1}\mathbf{M}_{W_1}^{\Sigma_1^{-1}}\mathbf{Y}_1 + \mathbf{X}_2'\boldsymbol\Sigma_2^{-1}\mathbf{Y}_2)],\end{aligned}$$

(Version III).

Proof Lemma 2 is ready to help us in proving this theorem. Here

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{W}_1 & \mathbf{0} \\ \mathbf{X}_2 & \mathbf{0} & \mathbf{W}_2 \end{pmatrix}, \ \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 \end{pmatrix}, \ \boldsymbol{\delta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}_1', \boldsymbol{\gamma}_2')',$$

so that

$$\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X} = \begin{pmatrix} \mathbf{X}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{X}_1 + \mathbf{X}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{X}_2 & \mathbf{X}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{W}_1 & \mathbf{X}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{W}_2 \\ \mathbf{W}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{X}_1 & \mathbf{W}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{W}_1 & \mathbf{0} \\ \mathbf{W}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{X}_2 & \mathbf{0} & \mathbf{W}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{W}_2 \end{pmatrix}. \tag{11}$$

Using Lemma 1, Versions I–III, and obvious relations

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= \mathbf{Q}_{11}(\mathbf{X}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{Y}_1 + \mathbf{X}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{Y}_2) + \mathbf{Q}_{12}\mathbf{W}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{Y}_1 + \mathbf{Q}_{13}\mathbf{W}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{Y}_2, \\ \widehat{\boldsymbol{\gamma}_1} &= \mathbf{Q}_{21}(\mathbf{X}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{Y}_1 + \mathbf{X}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{Y}_2) + \mathbf{Q}_{22}\mathbf{W}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{Y}_1 + \mathbf{Q}_{23}\mathbf{W}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{Y}_2, \\ \widehat{\boldsymbol{\gamma}_2} &= \mathbf{Q}_{31}(\mathbf{X}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{Y}_1 + \mathbf{X}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{Y}_2) + \mathbf{Q}_{32}\mathbf{W}_1'\boldsymbol{\Sigma}_1^{-1}\mathbf{Y}_1 + \mathbf{Q}_{33}\mathbf{W}_2'\boldsymbol{\Sigma}_2^{-1}\mathbf{Y}_2 \end{aligned}$$

we get the results—Versions I–III (in this order). □

Let us remark that functional dependence of above derived estimators to the other estimator(s) coheres with the functional dependence of diagonal blocks of inverse matrices in Lemma 1 to their other diagonal block(s). So we obtained such estimators of each of the parameters $\boldsymbol{\beta}, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2$ that are not a function of any other parameter's estimator. For instance, in Version I the estimator $\boldsymbol{\beta}$ is not a function of $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2$, in Version II $\boldsymbol{\gamma}_1$ is not a function of $\boldsymbol{\beta}, \boldsymbol{\gamma}_2$ and analogously in Version III. This result is important mainly from the theoretic point of view and as an effective tool for checking of numerical results. In the practice, Version I seems to be the most convenient for computing the estimators.

Example 1 Let us test the hypothesis $H_0 : \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2$ in the regular model (1) with $\mathbf{X}_1 = \mathbf{X}_2 = \mathbf{X}$, $\mathbf{W}_1 = \mathbf{W}_2 = \mathbf{W}$ and $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$ and without evaluating the corresponding estimators of the first order parameters. We will follow Lemma 2. Here $\mathbf{H} = (\mathbf{0}, \mathbf{I}, -\mathbf{I})$, so the F statistics (8) equals

$$\frac{(\widehat{\boldsymbol{\gamma}_1} - \widehat{\boldsymbol{\gamma}_2})'(\mathbf{Q}_{22} - \mathbf{Q}_{23} - \mathbf{Q}_{32} + \mathbf{Q}_{33})^{-1}(\widehat{\boldsymbol{\gamma}_1} - \widehat{\boldsymbol{\gamma}_2})}{\widehat{\sigma^2}},$$

where $\mathbf{Q}_{22}, \mathbf{Q}_{23}, \mathbf{Q}_{32}, \mathbf{Q}_{33}$ are given by inverse of (11) using Lemma 1. Then $\widehat{\boldsymbol{\gamma}_1} - \widehat{\boldsymbol{\gamma}_2}$ using Version I from Theorem 1 for $\widehat{\boldsymbol{\gamma}_1}$ and $\widehat{\boldsymbol{\gamma}_2}$ is of the form

$$(\mathbf{W}'\boldsymbol{\Sigma}^{-1}\mathbf{W})^{-1}\mathbf{W}'\boldsymbol{\Sigma}^{-1}(\mathbf{Y}_1 - \mathbf{Y}_2).$$

The same term $\widehat{\boldsymbol{\gamma}_1} - \widehat{\boldsymbol{\gamma}_2}$ with $\widehat{\boldsymbol{\gamma}_1}$ from Version II and $\widehat{\boldsymbol{\gamma}_2}$ from Version III is

$$\begin{aligned} &[\mathbf{W}'\boldsymbol{\Sigma}^{-1}\mathbf{W} - \mathbf{W}'\boldsymbol{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{M}_W^{\Sigma^{-1}}\mathbf{X} + \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{W}]^{-1} \\ &\times\mathbf{W}'\boldsymbol{\Sigma}^{-1}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{M}_W^{\Sigma^{-1}}\mathbf{X} + \mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}_W^{\Sigma^{-1}}](\mathbf{Y}_1 - \mathbf{Y}_2). \end{aligned}$$

Moreover $s_1 = s_2 = s$ and $n_1 = n_2 = N$, so we reject H_0 on the level α, if $F \geq F_{s,2N-r-2s}(1-\alpha)$.

Acknowledgement Author would like to thank to Mr. Martin Petera for help in preparation of this paper.

References

[1] Anděl, J.: *Základy matematické statistiky.* MATFYZPRESS, Praha, 2005 (in Czech).

[2] Harville, D. A.: *Matrix Algebra From a Statistican's Perspective.* Springer-Verlag, New York, 1999.

[3] Hron, K.: *Inversion of* 3×3 *Partitioned Matrices in Investigation of the Twoepoch Linear Model with the Nuisance Parameters.* Acta Univ. Palacki. Olomuc. Fac. rer. nat. Math., **45** (2006), 67–80.

[4] Hron, K.: *On One Twoepoch Linear Model with the Nuisance Parameters.* Mathematica Slovaca, to appear.

[5] Kubáček, L., Kubáčková, L., Volaufová, J.: *Statistical Models with Linear Structures.* Veda, Publishing House of the Slovak Academy of Sciences, Bratislava, 1995.

[6] Kubáček, L., Kubáčková, L.: *A Multiepoch Regression Model used in Geodesy.* Austrian Journal of Statistics **28** (1999), Nr. 4, 203–214.

[7] Kubáčková, L., Kubáček, L.: *Estimation in Multiepoch Regression Models with Different Structures for Studying Recent Crustal Movements.* Acta Univ. Palacki. Olomuc. Fac. rer. nat. Math. **35** (1996), 83–102.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 51–64

Eliminating Transformations for Nuisance Parameters in Linear Regression Models with Type I Constraints

Pavla KUNDEROVÁ

Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: kunderov@inf.upol.cz

(Received January 22, 2007)

Abstract

The linear regression model in which the vector of the first order parameter is divided into two parts: to the vector of the useful parameters and to the vector of the nuisance parameters is considered. The type I constraints are given on the useful parameters. We examine eliminating transformations which eliminate the nuisance parameters without loss of information on the useful parameters.

Key words: Regular linear regression model; nuisance parameters; BLUE; constraints.

2000 Mathematics Subject Classification: 62J05

1 Introduction, notations

Transformations for nuisance parameters in linear regression models with nuisance parameters are studied for instance in [3], [4], [6]. This paper deals with similar problems in models to which type I constraints are added.

The following notation will be used throughout the paper:

R^n	the space of all n-dimensional real vectors;
u_p, $\mathsf{A}_{m,n}$	the real column p-dimensional vector, the real $m \times n$ matrix;
$\mathsf{A}', r(\mathsf{A})$	the transpose, the rank of the matrix A;
$\mathscr{M}(\mathsf{A})$, $\mathrm{Ker}(\mathsf{A})$	the range, the null space of the matrix A;
A^-	a generalized inverse of a matrix A (satisfying $\mathsf{AA^-A} = \mathsf{A}$);
A^+	the Moore–Penrose generalized inverse of a matrix A (satisfying $\mathsf{AA^+A} = \mathsf{A}$, $\mathsf{A^+AA^+} = \mathsf{A}^+$, $(\mathsf{AA}^+)' = \mathsf{AA}^+$, $(\mathsf{A^+A})' = \mathsf{A^+A}$);
P_A	the orthogonal projector in the Euclidean norm onto $\mathscr{M}(\mathsf{A})$;
$\mathsf{M}_A = \mathsf{I} - \mathsf{P}_A$	the orthogonal projector in the Euclidean norm onto $\mathscr{M}^{\perp}(\mathsf{A})$;
I_k	the $k \times k$ identity matrix;
$\mathsf{O}_{m,n}$	the $m \times n$ null matrix;
o	the null vector.

If $\mathscr{M}(\mathsf{A}) \subset \mathscr{M}(\mathsf{U})$, U p.s.d., then the symbol $\mathsf{P}_A^{U^-}$ denotes the projector projecting vectors in $\mathscr{M}(\mathsf{U})$ onto $\mathscr{M}(\mathsf{A})$ along $\mathscr{M}(\mathsf{UA}^{\perp})$. A general representation of all such projectors $\mathsf{P}_A^{U^-}$ is given by $\mathsf{A}(\mathsf{A}'\mathsf{U}^-\mathsf{A})^-\mathsf{A}'\mathsf{U}^- + \mathsf{B}(\mathsf{I} - \mathsf{UU}^-)$, where B is arbitrary, (see [7], (2.14)). $\mathsf{M}_A^{U^-} = \mathsf{I} - \mathsf{P}_A^{U^-}$.

Let $\mathsf{N}_{n,n}$ is p.d. (p.s.d.) matrix and $\mathsf{A}_{m,n}$ an arbitrary matrix, then the symbol $\mathsf{A}^-_{m(N)}$ denotes the matrix satisfying $\mathsf{AA}^-_{m(N)}\mathsf{A} = \mathsf{A}$ and $\mathsf{NA}^-_{m(N)}\mathsf{A} = (\mathsf{NA}^-_{m(N)}\mathsf{A})'$. [$\mathsf{A}^-_{m(N)}\mathsf{y}$ is any solution of the consistent system $\mathsf{Ax} = \mathsf{y}$ whose N-seminorm is minimal]. In general $\mathsf{A}^-_{m(N)} = (\mathsf{N} + \mathsf{A}'\mathsf{A})^-\mathsf{A}'[\mathsf{A}(\mathsf{N} + \mathsf{AA}^-)^-\mathsf{A}']^-$. If the condition $\mathscr{M}(\mathsf{A}') \subset \mathscr{M}(\mathsf{N})$ is fulfilled, then $\mathsf{A}^-_{m(N)} = \mathsf{N}^-\mathsf{A}'(\mathsf{AN}^-\mathsf{A}')^-$, (see [2], pp. 14–15).

Assertion 1 (see [3], Lemma 10.1.35) *Let* X *be any* $n \times k$ *matrix and* Σ *an* $n \times n$ *p.s.d. matrix.*

(i) If Σ *is p.d., then*

$$(\mathsf{M}_X\Sigma\mathsf{M}_X)^+ = \Sigma^{-1} - \Sigma^{-1}\mathsf{X}(\mathsf{X}'\Sigma^{-1}\mathsf{X})^-\mathsf{X}'\Sigma^{-1} = \Sigma^{-1}\mathsf{M}_X^{\Sigma^{-1}}.$$

(ii) If Σ *is not p.d. however* $\mathscr{M}(\mathsf{X}) \subset \mathscr{M}(\Sigma)$, *then*

$$(\mathsf{M}_X\Sigma\mathsf{M}_X)^+ = \Sigma^+ - \Sigma^+\mathsf{X}(\mathsf{X}'\Sigma^-\mathsf{X})^-\mathsf{X}'\Sigma^+.$$

(iii) In general case

$$(\mathsf{M}_X\Sigma\mathsf{M}_X)^+ = (\Sigma + \mathsf{XX}')^+ - (\Sigma + \mathsf{XX}')^+\mathsf{X}[\mathsf{X}'(\Sigma + \mathsf{XX}')^-\mathsf{X}]^-\mathsf{X}'(\Sigma + \mathsf{XX}')^+.$$

(iv)

$$(\mathsf{M}_X\Sigma\mathsf{M}_X)^+ = \mathsf{M}_X(\mathsf{M}_X\Sigma\mathsf{M}_X)^+ = (\mathsf{M}_X\Sigma\mathsf{M}_X)^+\mathsf{M}_X = \mathsf{M}_X(\mathsf{M}_X\Sigma\mathsf{M}_X)^+\mathsf{M}_X.$$

Assertion 2 *Let* $\mathsf{D} = \begin{pmatrix} \mathsf{A}, & \mathsf{B} \\ \mathsf{B}', & \mathsf{C} \end{pmatrix}$ *be symmetric and positive semidefinite matrix. If* $\mathscr{M}(\mathsf{B}') \subset \mathscr{M}(\mathsf{C} - \mathsf{B}'\mathsf{A}^+\mathsf{B})$, *then*

$$\mathsf{D}^+ = \begin{pmatrix} \mathsf{A}, & \mathsf{B} \\ \mathsf{B}', & \mathsf{C} \end{pmatrix}^+ = \begin{pmatrix} \mathsf{A}^+ + \mathsf{A}^+\mathsf{B}(\mathsf{C} - \mathsf{B}'\mathsf{A}^+\mathsf{B})^+\mathsf{B}'\mathsf{A}^+, & -\mathsf{A}^+\mathsf{B}(\mathsf{C} - \mathsf{B}'\mathsf{A}^+\mathsf{B})^+ \\ -(\mathsf{C} - \mathsf{B}'\mathsf{A}^+\mathsf{B})^+\mathsf{B}'\mathsf{A}^+, & (\mathsf{C} - \mathsf{B}'\mathsf{A}^+\mathsf{B})^+ \end{pmatrix}.$$

If $\mathcal{M}(\mathsf{B}) \subset \mathcal{M}(\mathsf{A} - \mathsf{BC}^+\mathsf{B}')$, *then*

$$\mathsf{D}^+ = \begin{pmatrix} \mathsf{A}, & \mathsf{B} \\ \mathsf{B}', & \mathsf{C} \end{pmatrix}^+ = \begin{pmatrix} (\mathsf{A} - \mathsf{BC}^+\mathsf{B}')^+, & -(\mathsf{A} - \mathsf{BC}^+\mathsf{B}')^+\mathsf{BC}^+ \\ -\mathsf{C}^+\mathsf{B}'(\mathsf{A} - \mathsf{BC}^+\mathsf{B}')^+, & \mathsf{C}^+ + \mathsf{C}^+\mathsf{B}'(\mathsf{A} - \mathsf{BC}^+\mathsf{B}')^+\mathsf{BC}^+ \end{pmatrix}.$$

Proof Assertions can be proved directly. As D is p.s.d. matrix, there exists block matrix $\binom{\mathsf{J}}{\mathsf{K}}$ such that

$$\begin{pmatrix} \mathsf{A}, & \mathsf{B} \\ \mathsf{B}', & \mathsf{C} \end{pmatrix} = \begin{pmatrix} \mathsf{J} \\ \mathsf{K} \end{pmatrix}(\mathsf{J}', \mathsf{K}') = \begin{pmatrix} \mathsf{JJ}', & \mathsf{JK}' \\ \mathsf{KJ}', & \mathsf{KK}' \end{pmatrix} \Rightarrow \mathcal{M}(\mathsf{B}) = \mathcal{M}(\mathsf{JK}') \subset \mathcal{M}(\mathsf{J}) = \mathcal{M}(\mathsf{A}),$$

analogously $\mathcal{M}(\mathsf{B}') \subset \mathcal{M}(\mathsf{C})$. It implies that $\mathsf{AA}^+\mathsf{B} = \mathsf{B}$, $\mathsf{B}'\mathsf{A}^+\mathsf{A} = \mathsf{B}'$, $\mathsf{CC}^+\mathsf{B}' = \mathsf{B}'$, $\mathsf{BC}^+\mathsf{C} = \mathsf{B}$. These matrices don't depend on the choice of g-inverses. We can easily prove, that relations $\mathsf{DD}^+\mathsf{D} = \mathsf{D}$, $\mathsf{D}^+\mathsf{DD}^+ = \mathsf{D}^+$ are valid for both formulas. Matrices $\mathsf{D}^+\mathsf{D}$, DD^+ are symmetric, if conditions $\mathcal{M}(\mathsf{B}') \subset \mathcal{M}(\mathsf{C} - \mathsf{B}'\mathsf{A}^+\mathsf{B})$ and $\mathcal{M}(\mathsf{B}) \subset \mathcal{M}(\mathsf{A} - \mathsf{BC}^+\mathsf{B}')$ are satisfied. It is to be remarked that these conditions are valid if $r(\mathsf{D}) = r(\mathsf{A}) + r(\mathsf{C})$. □

Let us consider following linear model with nuisance parameters

$$\mathsf{Y} \sim [\,(\mathsf{X}, \mathsf{S}) \begin{pmatrix} \beta \\ \kappa \end{pmatrix}, \Sigma_\vartheta\,], \quad \Sigma_\vartheta \;\text{ known matrix,} \tag{1}$$

where $\mathsf{Y} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)'$ is a random observation vector; $\beta \in R^k$ is a vector of the useful parameters; $\kappa \in R^l$ is a vector of the nuisance parameters; $\mathsf{X}_{n,k}$ is a design matrix belonging to the vector β; $\mathsf{S}_{n,l}$ is a design matrix belonging to the vector κ.

We suppose that

1. $E(\mathsf{Y}) = \mathsf{X}\beta + \mathsf{S}\kappa$, $\forall \beta \in R^k$, $\forall \kappa \in R^l$,
2. $\mathrm{var}(\mathsf{Y}) = \Sigma_\vartheta = \sum_{i=1}^p \vartheta_i \mathsf{V}_i$, $\vartheta = (\vartheta_1, \ldots, \vartheta_p)' \in \underline{\vartheta} \subset R^p$, $\underline{\vartheta}$ is supposed to be with nonempty topological interior.

In this paper we consider that the given matrices $\mathsf{V}_1, \ldots, \mathsf{V}_p$ are p.s.d. and that the variance components $\vartheta_1, \ldots, \vartheta_p$ are positive (mixed linear model, see [1], Chapter 4).

3. Σ_ϑ is not a function of the vector $(\beta', \kappa')'$.

If matrix Σ_ϑ is positive definite and $r(\mathsf{X}, \mathsf{S}) = k + l < n$, the model is said to be *regular*, (see [3], p.13).

Parametric function $\mathsf{f}'\beta$ is unbiasedly estimable in model (1) iff $\mathsf{f} \in \mathcal{M}(\mathsf{X}'\mathsf{M}_S)$, see [6], Remark 2.

There are situations in the practice that auxiliary information on the vector of useful regression coefficients β is known, it means that the parametric space for β is not R^k but its subset only,

$$\beta \;\in\; \{\mathsf{u} \in R^k : \mathsf{b} + \mathsf{B}\mathsf{u} = o\}, \tag{2}$$

where B is a $q \times k$ known matrix. Since no assumption on the $r(\mathsf{B})$ is considered, it must be assumed that a given q-dimensional vector b satisfies $\mathsf{b} \in \mathcal{M}(\mathsf{B})$. This constraints on the useful parameters will be called *type I constraints*.

Lemma 1 *The class of unbiasedly estimable functions of the useful parameters in model (1) with constraints (2) is created by all functions* $\mathsf{h}'\beta$ *possesing*

$$\mathsf{h} \in \mathscr{M}(\mathsf{X}'\mathsf{M}_S, \mathsf{B}').$$

Proof Function $\mathsf{h}'\beta + a$, $\mathsf{h} \in R^k$, $a \in R$ is in model (1) with constraints (2) unbiasedly estimable iff there exists statistic $\mathsf{g}'\mathsf{Y} + c$, $\mathsf{g} \in R^n$, $c \in R$ such that

$$E(\mathsf{g}'\mathsf{Y} + \mathsf{c}) = \mathsf{g}'[\mathsf{X}\beta + \mathsf{S}\kappa] + c = \mathsf{h}'\beta + a, \quad \forall\beta,\ \forall\kappa$$

$$\Leftrightarrow (\mathsf{g}'\mathsf{X} - \mathsf{h}')\beta + c - a = 0 \wedge \mathsf{g}'\mathsf{S} = \mathsf{o}',\ \forall\beta$$

$$\Leftrightarrow (\mathsf{u}'\mathsf{M}_S\mathsf{X} - \mathsf{h}')\beta + c - a = 0,\ \forall\beta,\ \mathsf{u} \in R^n$$

$\Leftrightarrow$ there exists vector $\mathsf{k} \in R^q$ such that $\mathsf{k}'\mathsf{B} = \mathsf{u}'\mathsf{M}_S\mathsf{X} - \mathsf{h}' \wedge \mathsf{k}'\mathsf{b} = c - a$.

Because c can be chosen arbitrarily, the necessary and sufficient condition for unbiasedly estimable function is

$$\mathsf{u}'\mathsf{M}_S\mathsf{X} - \mathsf{k}'\mathsf{B} = \mathsf{h}' \ \Leftrightarrow\ \mathsf{h} = \mathsf{X}'\mathsf{M}_S\mathsf{u} - \mathsf{B}'\mathsf{k} \ \Leftrightarrow\ \mathsf{h} \in \mathscr{M}(\mathsf{X}'\mathsf{M}_S, \mathsf{B}'). \qquad \square$$

Remark 1 The BLUE (best linear unbiased estimator) of the vector function $\mathsf{M}_S\mathsf{X}\beta$ in the singular model (1) with constraints (2) is

$$\widehat{\mathsf{M}_S\mathsf{X}\beta} = \mathsf{M}_S\mathsf{P}_{(XM_{B'},S)}^{[\Sigma_\vartheta + XM_{B'}X' + SS']^+}\mathsf{Y} - \mathsf{M}_S\mathsf{M}_{(XM_{B'},S)}^{[\Sigma_\vartheta + XM_{B'}X' + SS']^+}\mathsf{X}\mathsf{B}'(\mathsf{B}\mathsf{B}')^-\mathsf{b}.$$

It is proved in [1], 2.10.2. and enables us to get BLUE of the unbiasedly estimable functions $\mathsf{h}'\beta, \mathsf{h} \in \mathscr{M}(\mathsf{X}'\mathsf{M}_S)$ in singular model (1) with constraints (2).

In the regular model (1) with constraints (2) the BLUE of the parameter β is given by

$$\hat{\beta} = [\mathsf{I} - \mathsf{C}^{-1}\mathsf{B}'(\mathsf{B}\mathsf{C}^{-1}\mathsf{B}')^{-1}\mathsf{B}]\beta^* - \mathsf{C}^{-1}\mathsf{B}'(\mathsf{B}\mathsf{C}^{-1}\mathsf{B}')^{-1}\mathsf{b},$$

where

$$\mathsf{C} = \mathsf{X}'(\mathsf{M}_S\Sigma_\vartheta\mathsf{M}_S)^+\mathsf{X},$$

and where

$$\beta^* = [\mathsf{X}'(\mathsf{M}_S\Sigma_\vartheta\mathsf{M}_S)^+\mathsf{X}]^{-1}\mathsf{X}'(\mathsf{M}_S\Sigma_\vartheta\mathsf{M}_S)^+\mathsf{Y},$$

(estimator in the regular model (1) without constraints).

The variance matrix of the estimator $\hat{\beta}$ in regular model (1) with constraints (2) is given by

$$\operatorname{var}(\hat{\beta}) = (\mathsf{M}_{B'}\mathsf{C}\mathsf{M}_{B'})^+.$$

These assertions are proved in [5], Theorem 1, Theorem 2.

In the literature there are investigated properties of estimators of the parameters β, κ in model (1) under constraints (2), see for example [1], [5]. In cases when we are interested on useful parameters only it is possible to simplify model (1) by the propriate eliminating transformation, see [3], [4], [6].

In this paper we join both of the procedures mentioned. Firstly we use eliminating transformation and then we add constraints to the transformed model.

2 Type I constraints in the transformed model

Our task will be to eliminate the matrix S belonging to the vector of nuisance parameters, i.e. we consider the following class of eliminating matrices

$$\mathscr{T} = \{\mathsf{T} : \mathsf{TS} = \mathsf{O}\},$$

where T is matrix of the proper dimension, say of the type $r \times n$.

That leads us to linear models

$$\mathsf{TY} \sim [\mathsf{TX}\beta, \mathsf{T}\Sigma_\vartheta \mathsf{T}']. \tag{3}$$

If we now add constraints (2) to the model (3), we get model

$$\begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix} \sim \left[\begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix} \beta, \begin{pmatrix} \mathsf{T}\Sigma_\vartheta \mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} \right]. \tag{4}$$

Lemma 2 *Linear function* $\mathsf{f}'\beta + a, \mathsf{f} \in R^k, a \in R$ *is unbiasedly estimable in model (4), iff*

$$\mathsf{f} \in \mathscr{M}(\mathsf{X}'\mathsf{T}', \mathsf{B}').$$

Proof The assertion can be proved in the same way as in Lemma 1. □

In the following text we consider only transformation matrices T with the property

$$\mathscr{M}(\mathsf{X}'\mathsf{T}') = \mathscr{M}(\mathsf{X}'\mathsf{M}_S),$$

it means that transformations do not cause a loss of information on the parameter β.

Theorem 1 *For the BLUE of the function of the parameter* β *in the model (4) holds*

$$\widehat{\mathsf{TX}\beta} = \mathsf{P}_{TXM_{B'}}^{[T(\Sigma_\vartheta + XM_{B'}X')T']^+}\mathsf{TY} - \mathsf{M}_{TXM_{B'}}^{[T(\Sigma_\vartheta + XM_{B'}X')T']^+}\mathsf{TXB}'(\mathsf{BB}')^-\mathsf{b}.$$

Proof According to Theorem 3.1.3. in [3]

$$\widehat{\begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix}\beta} = \begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix} \left[(\mathsf{X}'\mathsf{T}', \mathsf{B}')^-_{m\left(\begin{smallmatrix} T\Sigma_\vartheta T', & 0 \\ 0, & 0 \end{smallmatrix}\right)} \right]' \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}$$

$$= \begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix} \left[(\mathsf{X}'\mathsf{T}', \mathsf{B}') \left\{ \begin{pmatrix} \mathsf{T}\Sigma_\vartheta \mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} + \begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix} (\mathsf{X}'\mathsf{T}', \mathsf{B}') \right\}^- \begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix} \right]^-$$

$$\times (\mathsf{X}'\mathsf{T}', \mathsf{B}') \left\{ \begin{pmatrix} \mathsf{T}\Sigma_\vartheta \mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} + \begin{pmatrix} \mathsf{TXX}'\mathsf{T}', & \mathsf{TXB}' \\ \mathsf{BX}'\mathsf{T}', & \mathsf{BB}' \end{pmatrix} \right\}^- \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}.$$

By the help of the Rohde's formula for g-inverse of the p.s.d. partitioned matrix (see [3], Theorem 10.1.40) we can write

$$\begin{pmatrix} \mathsf{T}[\Sigma_\vartheta + \mathsf{XX}']\mathsf{T}', & \mathsf{TXB}' \\ \mathsf{BX}'\mathsf{T}', & \mathsf{BB}' \end{pmatrix}^- = \begin{pmatrix} \boxed{11}, & \boxed{12} \\ \boxed{21}, & \boxed{22} \end{pmatrix},$$

where

$$\boxed{11} = [\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^-,$$
$$\boxed{12} = -[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^-\mathsf{TXB}'(\mathsf{BB}')^-$$
$$\boxed{21} = -(\mathsf{BB}')^-\mathsf{BX}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^-$$
$$\boxed{22} = (\mathsf{BB}')^- + (\mathsf{BB}')^-\mathsf{BX}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^-\mathsf{TXB}'(\mathsf{BB}')^-.$$

Then (we use Moore–Penrose g-inverse matrix for the sake of simplicity)

$$\left[(\mathsf{X}'\mathsf{T}', \mathsf{B}')\begin{pmatrix}\boxed{11}, & \boxed{12}\\ \boxed{21}, & \boxed{22}\end{pmatrix}\begin{pmatrix}\mathsf{TX}\\ \mathsf{B}\end{pmatrix}\right]^+$$
$$= [\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'(\mathsf{T}[\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}']\mathsf{T}')^-\mathsf{TXM}_{B'} + \mathsf{P}_{B'}]^+,$$

thus

$$\widehat{\begin{pmatrix}\mathsf{TX}\\ \mathsf{B}\end{pmatrix}}\beta =$$
$$= \begin{pmatrix}\mathsf{TX}\\ \mathsf{B}\end{pmatrix}\{[\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'(\mathsf{T}[\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}']\mathsf{T}')^-\mathsf{TXM}_{B'}]^+ + \mathsf{P}_{B'}\}(\mathsf{X}'\mathsf{T}', \mathsf{B}')$$
$$\times \begin{pmatrix}\boxed{11}, & \boxed{12}\\ \boxed{21}, & \boxed{22}\end{pmatrix}\begin{pmatrix}\mathsf{TY}\\ -\mathsf{b}\end{pmatrix}.$$

After some calculations we get

$$\widehat{\begin{pmatrix}\mathsf{TX}\\ \mathsf{B}\end{pmatrix}}\beta = \begin{pmatrix}\mathsf{P}_{TXM_{B'}}^{[T(\Sigma_\vartheta+XM_{B'}X')T']^+}\mathsf{TY} - \mathsf{M}_{TXM_{B'}}^{[T(\Sigma_\vartheta+XM_{B'}X')T']^+}\mathsf{TXB}'(\mathsf{BB}')^-\mathsf{b}\\ -\mathsf{b}\end{pmatrix}.$$

In the course of the proof following assertion has been used

$$\mathsf{A}'\mathsf{B} = \mathsf{O} \ \wedge \ \mathsf{BA}' = \mathsf{O} \ \Rightarrow (\mathsf{A}+\mathsf{B})^+ = \mathsf{A}^+ + \mathsf{B}^+. \qquad \square$$

Theorem 2 *The covariance matrix of the estimator* $\widehat{\mathsf{TX}\beta}$ *in model (4) is*

$$\mathrm{var}[\widehat{\mathsf{TX}\beta}] = \mathsf{TX}\{[\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'(\mathsf{T}[\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}']\mathsf{T}')^-\mathsf{TXM}_{B'}]^+ - \mathsf{M}_{B'}\}\mathsf{X}'\mathsf{T}'.$$

Proof

$$\mathrm{var}[\widehat{\mathsf{TX}\beta}] =$$
$$= \mathsf{P}_{TXM_{B'}}^{[T(\Sigma_\vartheta+XM_{B'}X')T']^+}[\mathsf{T}\Sigma_\vartheta\mathsf{T}' + \mathsf{TXM}_{B'}\mathsf{X}'\mathsf{T}' - \mathsf{TXM}_{B'}\mathsf{X}'\mathsf{T}'](\mathsf{P}_{TXM_{B'}}^{[T(\Sigma_\vartheta+XM_{B'}X')T']^+})'$$
$$= \mathsf{TXM}_{B'}\left(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXM}_{B'}\right)^+\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'$$
$$\times [\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}'][\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+$$
$$\times \mathsf{TXM}_{B'}\left(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{X}'\mathsf{T}'\mathsf{M}_{B'}\right)^+\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'$$

$$
\begin{aligned}
&- \mathsf{TXM}_{B'} \left(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXM}_{B'}\right)^+ \mathsf{M}_{B'}\mathsf{X}'\mathsf{T}' \\
&\times [\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXM}_{B'}\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+ \\
&\times \mathsf{TXM}_{B'} \left(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXM}_{B'}\right)^+ \mathsf{M}_{B'}\mathsf{X}'\mathsf{T}' \\
&= \mathsf{TXM}_{B'} \left(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXM}_{B'}\right)^+ \mathsf{M}_{B'}\mathsf{X}'\mathsf{T}' - \mathsf{TXM}_{B'}\mathsf{X}'\mathsf{T}' \\
&= \mathsf{TX}\left\{\left(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXM}_{B'}\right)^+ - \mathsf{M}_{B'}\right\}\mathsf{X}'\mathsf{T}'.
\end{aligned}
$$

In the course of the proof we have used Assertion 1, (ii) and following statement

$$\mathscr{M}(\mathsf{B}') \subset \mathscr{M}(\mathsf{A}') \;\Leftrightarrow\; \mathsf{BA}^-\mathsf{A} = \mathsf{B},$$

for matrices $\mathsf{A} = \mathsf{T}(\Sigma_\vartheta + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}'$ and $\mathsf{B} = \mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'$. □

Theorem 3 *Let the transformed model (4) where* $\Sigma_\vartheta = \sum_{i=1}^p \vartheta_i \mathsf{V}_i$, V_i *p.s.d.,* $\vartheta_i > 0$, $\forall i = 1,\dots,p$, *(mixed linear model) be under consideration. Let* $\Sigma_0 = \sum_{i=1}^p \vartheta_i^0 \mathsf{V}_i$, *where* $\vartheta^0 = (\vartheta_1^0,\dots,\vartheta_p^0)'$ *is as near to the actual value* ϑ^* *of the parameter as possible. The linear function* $\mathsf{g}'\vartheta, \vartheta \in \underline{\vartheta}$ *can be estimated by MINQUE (minimum norm quadratic unbiased estimator) iff*

$$\mathsf{g} \in \mathscr{M}[\mathsf{S}_{\left(M_{\binom{TX}{B}}\binom{T\Sigma_0 T',\ 0}{0,\ \ 0}M_{\binom{TX}{B}}\right)^+}], \tag{5}$$

where the (i,j)*-th element of the matrix* $\mathsf{S}_{\left(M_{\binom{TX}{B}}\binom{T\Sigma_0 T',\ 0}{0,\ \ 0}M_{\binom{TX}{B}}\right)^+}$ *is*

$$
\begin{aligned}
&\{\mathsf{S}_{\left(M_{\binom{TX}{B}}\binom{T\Sigma_0 T',\ 0}{0,\ \ 0}M_{\binom{TX}{B}}\right)^+}\}_{i,j} = \\
&= \mathrm{Tr}[(\mathsf{M}_{TXM_{B'}}\mathsf{T}\Sigma_0\mathsf{T}'\mathsf{M}_{TXM_{B'}})^+\mathsf{TV}_i\mathsf{T}'(\mathsf{M}_{TXM_{B'}}\mathsf{T}\Sigma_0\mathsf{T}'\mathsf{M}_{TXM_{B'}})^+\mathsf{TV}_j\mathsf{T}'], \\
&\qquad\qquad i,j = 1,\dots,p.
\end{aligned}
$$

If the condition (5) is satisfied, then the ϑ^0*-MINQUE is*

$$\widehat{\mathsf{g}'\vartheta} = \sum_{i=1}^p \lambda_i \begin{pmatrix}\mathsf{TY}\\ -\mathsf{b}\end{pmatrix}'$$

$$\times \begin{pmatrix} \mathsf{ZTV}_i\mathsf{T}'\mathsf{Z}; & -\mathsf{ZTV}_i\mathsf{T}'\mathsf{ZTXB}'(\mathsf{BB}')^- \\ -(\mathsf{BB}')^-\mathsf{BX}'\mathsf{T}'\mathsf{Z}'\mathsf{TV}_i\mathsf{T}'\mathsf{Z}'; & (\mathsf{BB}')^-\mathsf{BX}'\mathsf{T}'\mathsf{ZTV}_i\mathsf{T}'\mathsf{Z}'\mathsf{TXB}'(\mathsf{BB}')^- \end{pmatrix} \begin{pmatrix}\mathsf{TY}\\ -\mathsf{b}\end{pmatrix},$$

where $\mathsf{Z} = [\mathsf{M}_{TXM_{B'}}\mathsf{T}\Sigma_0\mathsf{T}'\mathsf{M}_{TXM_{B'}}]^+$, *and where the vector* $\lambda = (\lambda_1,\dots,\lambda_p)'$ *is a solution of the equation*

$$\mathsf{S}_{\left(M_{\binom{TX}{B}}\binom{T\Sigma_0 T',\ 0}{0,\ \ 0}M_{\binom{TX}{B}}\right)^+}\lambda = \mathsf{g}\,.$$

Proof We use following statement (see [4], p. 101) valid for the linear model $\mathsf{Y} \sim [\mathsf{X}\beta, \Sigma_\vartheta]$ where $\beta \in R^k$, $\Sigma_\vartheta = \sum_{i=1}^p \vartheta_i \mathsf{V}_i$, $\vartheta = (\vartheta_1, \dots, \vartheta_p)' \in \underline{\vartheta} \subset R^p$, $\vartheta_i > 0$, $\forall i = 1, \dots, p$, $\mathsf{V}_1, \dots, \mathsf{V}_p$ p.s.d. matrices (mixed linear model):

a) Let $\Sigma_0 = \sum_{i=1}^p \vartheta_i^0 \mathsf{V}_i$. The function $\mathsf{g}'\vartheta = \sum_{i=1}^p g_i \vartheta_i$, $\vartheta \in \underline{\vartheta}$, can be unbiasedly quadratically and invariantly estimated [i.e. the estimator has the form $\mathsf{Y}'\mathsf{A}\mathsf{Y}$, where $\mathsf{A}_{n,n}$ is symmetric matrix, the estimator is invariant with respect to the change of the vector β] if and only if $\mathsf{g} \in \mathscr{M}(\mathsf{S}_{(M_X\Sigma_0M_X)^+})$, where

$$\{\mathsf{S}_{(M_X\Sigma_0M_X)^+}\}_{i,j} = \mathrm{Tr}[(\mathsf{M}_X\Sigma_0\mathsf{M}_X)^+\mathsf{V}_i(\mathsf{M}_X\Sigma_0\mathsf{M}_X)^+\mathsf{V}_j],$$

$i, j = 1, \dots, p$.

b) If the function $\mathsf{g}'\vartheta$ satisfies the condition from a), then the ϑ^0-MINQUE of $\mathsf{g}'\vartheta$ is given as

$$\widehat{\mathsf{g}'\vartheta} = \sum_{i=1}^p \lambda_i \mathsf{Y}'(\mathsf{M}_X\Sigma_0\mathsf{M}_X)^+\mathsf{V}_i(\mathsf{M}_X\Sigma_0\mathsf{M}_X)^+\mathsf{Y},$$

where the vector $\lambda = (\lambda_1, \dots, \lambda_p)'$ is a solution of the system of equations

$$\mathsf{S}_{(M_X\Sigma_0M_X)^+}\lambda = \mathsf{g}.$$

We use this statement for the model (4) by following substitutions

$$\mathsf{Y} \to \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix} \quad \mathsf{X} \to \begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix} \quad \Sigma_0 \to \begin{pmatrix} \mathsf{T}\Sigma_0\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^p \vartheta_i^0 \mathsf{TV}_i\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix}.$$

Thus

$$\begin{aligned}
&\{\mathsf{S}_{\left(M_{\binom{TX}{B}}\begin{pmatrix} T\Sigma_0T', & 0 \\ 0, & 0\end{pmatrix}M_{\binom{TX}{B}}\right)^+}\}_{i,j} \\
&= \mathrm{Tr}\left\{\left[\mathsf{M}_{\binom{TX}{B}}\begin{pmatrix} \mathsf{T}\Sigma_0\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\mathsf{M}_{\binom{TX}{B}}\right]^+\begin{pmatrix} \mathsf{TV}_i\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\right. \\
&\quad \left.\times\left[\mathsf{M}_{\binom{TX}{B}}\begin{pmatrix} \mathsf{T}\Sigma_0\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\mathsf{M}_{\binom{TX}{B}}\right]^+\begin{pmatrix} \mathsf{TV}_j\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\right\}.
\end{aligned}$$

Let us denote

$$\left[\begin{pmatrix} \mathsf{T}\Sigma_0\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O}\end{pmatrix} + \begin{pmatrix} \mathsf{TX} \\ \mathsf{B}\end{pmatrix}(\mathsf{X}'\mathsf{T}', \mathsf{B}')\right]^+ = \begin{pmatrix} \boxed{\mathrm{aa}}, & \boxed{\mathrm{ab}} \\ \boxed{\mathrm{ba}}, & \boxed{\mathrm{bb}}\end{pmatrix},$$

where (see Assertion 2)

$$\begin{aligned}
\boxed{\mathrm{aa}} &= [\mathsf{T}(\Sigma_0 + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+ \\
\boxed{\mathrm{ab}} &= -[\mathsf{T}(\Sigma_0 + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXB}'(\mathsf{BB}')^+, \\
\boxed{\mathrm{ba}} &= -(\mathsf{BB}')^+\mathsf{BX}'\mathsf{T}'[\mathsf{T}(\Sigma_0 + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+, \\
\boxed{\mathrm{bb}} &= (\mathsf{BB}')^+ + (\mathsf{BB}')^+\mathsf{BX}'\mathsf{T}'[\mathsf{T}(\Sigma_0 + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXB}'(\mathsf{BB}')^+.
\end{aligned}$$

By Assertion 1,(ii) (the Moore–Penrose matrices are used because of uniqueness of matrix expressions)

$$\left[\mathsf{M}_{\binom{TX}{B}}\begin{pmatrix}\mathsf{T}\Sigma_0\mathsf{T}', \mathsf{O}\\ \mathsf{O} \quad ,\mathsf{O}\end{pmatrix}\mathsf{M}_{\binom{TX}{B}}\right]^+ = \begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix}$$

$$-\begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix}\begin{pmatrix}\mathsf{TX}\\ \mathsf{B}\end{pmatrix}\left\{(\mathsf{X}'\mathsf{T}',\mathsf{B}')\begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix}\begin{pmatrix}\mathsf{TX}\\ \mathsf{B}\end{pmatrix}\right\}^+$$

$$\times(\mathsf{X}'\mathsf{T}',\mathsf{B}')\begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix} = \begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix}$$

$$-\begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix}\begin{pmatrix}\mathsf{TX}\\ \mathsf{B}\end{pmatrix}\{\mathsf{P}_{B'} + \mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'[\mathsf{T}(\Sigma_0 + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+\mathsf{TXM}_{B'}\}^+$$

$$\times(\mathsf{X}'\mathsf{T}',\mathsf{B}')\begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix} = \begin{pmatrix}\boxed{\text{aa}}, & \boxed{\text{ab}}\\ \boxed{\text{ba}}, & \boxed{\text{bb}}\end{pmatrix} - \begin{pmatrix}\boxed{\text{I}}, & \boxed{\text{II}}\\ \boxed{\text{III}}, & \boxed{\text{IV}}\end{pmatrix},$$

where by notation

$$\mathsf{U} = [\mathsf{T}(\Sigma_0 + \mathsf{XM}_{B'}\mathsf{X}')\mathsf{T}']^+,$$

$$\begin{aligned}
\boxed{\text{I}} &= \mathsf{UTX}(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'\mathsf{UTXM}_{B'})^+\mathsf{X}'\mathsf{T}'\mathsf{U},\\
\boxed{\text{II}} &= -\mathsf{UTX}(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'\mathsf{UTXM}_{B'})^+\mathsf{X}'\mathsf{T}'\mathsf{UTXB}'(\mathsf{BB}')^+ = \boxed{\text{III}}\,',\\
\boxed{\text{IV}} &= (\mathsf{BB}')^+\mathsf{BX}'\mathsf{T}'\mathsf{UTX}(\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'\mathsf{UTXM}_{B'})^+\mathsf{X}'\mathsf{T}'\mathsf{UTXB}'(\mathsf{BB}')^+ + (\mathsf{BB}')^+.
\end{aligned}$$

After some calculations using notation

$$\mathsf{Z} = (\mathsf{M}_{TXM_{B'}}\mathsf{T}\Sigma_0\mathsf{T}'\mathsf{M}_{TXM_{B'}})^+,$$

we get

$$\left[\mathsf{M}_{\binom{TX}{B}}\begin{pmatrix}\mathsf{T}\Sigma_0\mathsf{T}', \mathsf{O}\\ \mathsf{O} \quad ,\mathsf{O}\end{pmatrix}\mathsf{M}_{\binom{TX}{B}}\right]^+$$

$$= \begin{pmatrix}\mathsf{Z}, & -\mathsf{ZTXB}'(\mathsf{BB}')^+\\ -(\mathsf{BB}')^+\mathsf{BX}'\mathsf{T}'\mathsf{Z}', & (\mathsf{BB}')^+\mathsf{BX}'\mathsf{T}'\mathsf{ZTXB}'(\mathsf{BB}')^+\end{pmatrix} = \begin{pmatrix}\mathsf{E}, & \mathsf{F}\\ \mathsf{F}', & \mathsf{G}\end{pmatrix}.$$

Thus

$$\{\mathsf{S}_{\left(M_{\binom{TX}{B}}\binom{T\Sigma_0T',\ 0}{0,\quad 0}M_{\binom{TX}{B}}\right)^+}\}_{i,j}$$

$$= \mathrm{Tr}\left[\begin{pmatrix}\mathsf{E}, & \mathsf{F}\\ \mathsf{F}', & \mathsf{G}\end{pmatrix}\begin{pmatrix}\mathsf{TV}_i\mathsf{T}', & \mathsf{O}\\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\begin{pmatrix}\mathsf{E}, & \mathsf{F}\\ \mathsf{F}', & \mathsf{G}\end{pmatrix}\begin{pmatrix}\mathsf{TV}_j\mathsf{T}', & \mathsf{O}\\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\right] = \mathrm{Tr}[\mathsf{ETV}_i\mathsf{T}'\mathsf{ETV}_j\mathsf{T}']$$

$$= \mathrm{Tr}[(\mathsf{M}_{TXM_{B'}}\mathsf{T}\Sigma_0\mathsf{T}'\mathsf{M}_{TXM_{B'}})^+\mathsf{TV}_i\mathsf{T}'(\mathsf{M}_{TXM_{B'}}\mathsf{T}\Sigma_0\mathsf{T}'\mathsf{M}_{TXM_{B'}})^+\mathsf{TV}_j\mathsf{T}'],$$

$i, j = 1, \dots, p.$

If

$$\mathsf{g} \in \mathscr{M}(\mathsf{S}_{\left(M_{\binom{TX}{B}}\binom{T\Sigma_0 T',\ 0}{0,\ \ 0}M_{\binom{TX}{B}}\right)^+}),$$

then under the model (4)

$$\widehat{\mathsf{g}'\vartheta} = \sum_{i=1}^{p} \lambda_i \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}' \left[\mathsf{M}_{\binom{TX}{B}} \begin{pmatrix} \mathsf{T}\Sigma_0\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} \mathsf{M}_{\binom{TX}{B}} \right]^+ \begin{pmatrix} \mathsf{TV}_i\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix}$$

$$\times \left[\mathsf{M}_{\binom{TX}{B}} \begin{pmatrix} \mathsf{T}\Sigma_0\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} M_{\binom{TX}{B}} \right]^+ \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}$$

$$= \sum_{i=1}^{p} \lambda_i \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}' \begin{pmatrix} \mathsf{ETV}_i\mathsf{T}'\mathsf{E}, & \mathsf{ETV}_i\mathsf{T}'\mathsf{F} \\ \mathsf{F}'\mathsf{TV}_i\mathsf{T}'\mathsf{E}, & \mathsf{F}'\mathsf{TV}_i\mathsf{T}'\mathsf{F} \end{pmatrix} \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix} = \sum_{i=1}^{p} \lambda_i \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}'$$

$$\times \begin{pmatrix} \mathsf{ZTV}_i\mathsf{T}'\mathsf{Z}, & -\mathsf{ZTV}_i\mathsf{T}'\mathsf{ZTXB}'(\mathsf{BB}')^+ \\ -(\mathsf{BB}')^+\mathsf{BX}'\mathsf{T}'\mathsf{Z}'\mathsf{TV}_i\mathsf{T}'\mathsf{Z}', & (\mathsf{BB}')^+\mathsf{BX}'\mathsf{T}'\mathsf{Z}'\mathsf{TV}_i\mathsf{T}'\mathsf{ZTXB}'(\mathsf{BB}')^+ \end{pmatrix} \begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix},$$

where $\mathsf{Z} = (\mathsf{M}_{TXM_{B'}}\mathsf{T}\Sigma_0\mathsf{T}'\mathsf{M}_{TXM_{B'}})^+$. □

Theorem 4 *Function* $\mathsf{g}_1'\mathsf{TY}$ *is the best unbiased estimator of its mean value in the model (4) iff*

$$\mathsf{g}_1 \in \mathscr{M}[\mathsf{M}_{(T\Sigma_\vartheta T'[I-TX(X'T'TX+BB')^- X'T],T\Sigma_\vartheta T'TX(X'T'TX+BB')^- B')}].$$

Proof Function $\mathsf{g}'\binom{\mathsf{TY}}{-\mathsf{b}}$, $\mathsf{g} = \binom{\mathsf{g}_1}{\mathsf{g}_2}$, $\mathsf{g}_1 \in R^r$, $\mathsf{g}_2 \in R^q$, is in the model (4) the best unbiased estimator of its mean value iff

$$\operatorname{cov}\left\{ \mathsf{g}'\begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}, \tau_0\left[\begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}\right] \right\} = 0,$$

where $\tau_0\left[\binom{\mathsf{TY}}{-\mathsf{b}}\right]$ is arbitrary unbiased estimator of the null function $\mathsf{g}_0(\beta,\vartheta) = 0$, (see [4], p. 84). Any unbiased estimator of this function is of the form

$$\tau_0\left[\begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}\right] = \mathsf{f}'\begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}, \quad \mathsf{f} = \begin{pmatrix} \mathsf{f}_1 \\ \mathsf{f}_2 \end{pmatrix}, \quad \mathsf{f} \in \mathscr{M}\left(\mathsf{M}_{\binom{TX}{B}}\right),$$

as

$$E[\mathsf{f}_1'\mathsf{TY} + \mathsf{f}_2'(-\mathsf{b})] = \mathsf{f}_1'\mathsf{TX}\beta + \mathsf{f}_2'(-\mathsf{b}) = (\mathsf{f}_1'\mathsf{TX} + \mathsf{f}_2'\mathsf{B})\beta = 0, \quad \forall\beta,$$

$$\Leftrightarrow \quad (\mathsf{f}_1', \mathsf{f}_2')\begin{pmatrix} \mathsf{TX} \\ \mathsf{B} \end{pmatrix} = \mathsf{o}', \quad \Leftrightarrow \quad \mathsf{f} \in \mathscr{M}\left(\mathsf{M}_{\binom{TX}{B}}\right).$$

Let $\mathsf{u} = \binom{\mathsf{u}_1}{\mathsf{u}_2}$, $\mathsf{u}_1 \in R^r$, $\mathsf{u}_2 \in R^q$, be arbitrary. Then the covariance

$$\operatorname{cov}\left(\mathsf{g}'\begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix}, \mathsf{u}'\mathsf{M}_{\binom{TX}{B}}\begin{pmatrix} \mathsf{TY} \\ -\mathsf{b} \end{pmatrix} \right) = \mathsf{g}'\begin{pmatrix} \mathsf{T}\Sigma_\vartheta\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} \mathsf{M}_{\binom{TX}{B}}\mathsf{u}$$

$$= \mathsf{g}' \begin{pmatrix} \mathsf{T}\Sigma_\vartheta\mathsf{T}', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} \begin{pmatrix} \mathsf{I}_r - \mathsf{TX}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{X}'\mathsf{T}', & -\mathsf{TX}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{B}' \\ -\mathsf{B}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{X}'\mathsf{T}', & \mathsf{I}_q - \mathsf{B}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{B}' \end{pmatrix} \mathsf{u}$$

$$= (\mathsf{g}_1', \mathsf{g}_2') \begin{pmatrix} \mathsf{T}\Sigma_\vartheta\mathsf{T}'(\mathsf{I}_r - \mathsf{TX}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{X}'\mathsf{T}'), & -\mathsf{T}\Sigma_\vartheta\mathsf{T}'\mathsf{TX}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{B}' \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} \mathsf{u}$$

$$= 0.$$

$$\Leftrightarrow\ \mathsf{g}_1'(\mathsf{T}\Sigma_\vartheta\mathsf{T}'[\mathsf{I}_r - \mathsf{TX}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{X}'\mathsf{T}'], -\mathsf{T}\Sigma_\vartheta\mathsf{T}'\mathsf{TX}(\mathsf{X}'\mathsf{T}'\mathsf{TX}+\mathsf{B}'\mathsf{B})^-\mathsf{B}') = \mathsf{o}'.$$

Thus $\mathsf{g}_1'\mathsf{TY}$ is the best unbiased estimator of its mean value iff

$$\mathsf{g}_1 \in \mathscr{M}[\mathsf{M}_{(T\Sigma_\vartheta T'[I_r - TX(X'T'TX+B'B)^-X'T'], T\Sigma_\vartheta T'TX(X'T'TX+B'B)^-B')}]. \quad \square$$

Remark 2 If we change the ordering of the procedures described at the beginning of this section, we get the same model. Indeed by joining linear model (1) with constraints (2), we can write

$$\begin{pmatrix} \mathsf{Y} \\ -\mathsf{b} \end{pmatrix} \sim \left[\begin{pmatrix} \mathsf{X}, & \mathsf{S} \\ \mathsf{B}, & \mathsf{O} \end{pmatrix} \begin{pmatrix} \beta \\ \kappa \end{pmatrix}, \begin{pmatrix} \Sigma, & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} \right].$$

The transformation by the matrix $\begin{pmatrix} \mathsf{T}, & \mathsf{O} \\ \mathsf{O}, & \mathsf{I} \end{pmatrix}$, such that $\mathsf{TS} = \mathsf{O}$, leeds to the model (4).

3 Examples of the transformation matrices

The general solution of the matrix eguation $\mathsf{TS} = \mathsf{O}$ is of the form

$$\mathsf{T} = \mathsf{A}(\mathsf{I} - \mathsf{SS}^-),$$

where A is an arbitrary matrix of the corresponding type, S^- is some version of generalized inverse of the matrix S.

If we choose $\mathsf{S}^- = (\mathsf{S}^-\mathsf{WS})^-\mathsf{S}'\mathsf{W}$, where W is an arbitrary p.s.d. matrix such that

$$\mathscr{M}(\mathsf{S}') = \mathscr{M}(\mathsf{S}'\mathsf{WS}), \tag{6}$$

then $\mathsf{T} = \mathsf{AM}_S^W$, where M_S^W is given uniquely.

First we confine us to the transformation matrix

a) $\mathsf{T} = \mathsf{M}_S^W$,

i.e. we consider transformed linear model

$$\mathsf{M}_S^W\mathsf{Y} \sim [\mathsf{M}_S^W\mathsf{X}\beta, \mathsf{M}_S^W\Sigma(\mathsf{M}_S^W)']. \tag{7}$$

Thus model with the type I constraints is following

$$\begin{pmatrix} \mathsf{M}_S^W\mathsf{Y} \\ -\mathsf{b} \end{pmatrix} \sim \left[\begin{pmatrix} \mathsf{M}_S^W\mathsf{X} \\ \mathsf{B} \end{pmatrix} \beta, \begin{pmatrix} \mathsf{M}_S^W\Sigma_\vartheta(\mathsf{M}_S^W)', & \mathsf{O} \\ \mathsf{O}, & \mathsf{O} \end{pmatrix} \right]. \tag{8}$$

It can be proved (see [6], chapter 3) that

$$\mathscr{M}(\mathsf{M}_S) = \mathscr{M}((\mathsf{M}_S^W)'),$$

thus

$$\mathscr{M}(\mathsf{X}'\mathsf{M}_S, \mathsf{B}') = \mathscr{M}(\mathsf{X}'(\mathsf{M}_S^W)', \mathsf{B}'),$$

i.e. the classes of unbiasedly estimable functions $\mathsf{g}'\beta$ in model (1) with constraints (2) and in model (8) are identical.

According to Theorem 1 and Theorem 2

$$\widehat{\mathsf{M}_S^W\mathsf{X}\beta} = \mathsf{P}_{M_S^W X M_{B'}}^{[M_S^W(\Sigma+XM_{B'}X')(M_S^W)']^+}\mathsf{M}_S^W[\mathsf{Y}+\mathsf{XB}'(\mathsf{BB}')^-\mathsf{b}] - \mathsf{M}_S^W\mathsf{XB}'(\mathsf{BB}')^-\mathsf{b}$$

$$= \mathsf{M}_S^W\mathsf{XM}_{B'}\left[\mathsf{M}_{B'}\mathsf{X}'(\mathsf{M}_S^W)'\left(\mathsf{M}_S^W(\Sigma+\mathsf{XM}_{B'}\mathsf{X}')(\mathsf{M}_S^W)'\right)^+\mathsf{M}_S^W\mathsf{XM}_{B'}\right]^+\mathsf{M}_{B'}$$

$$\times\mathsf{X}'(\mathsf{M}_S^W)'\left(\mathsf{M}_S^W(\Sigma+\mathsf{XM}_{B'}\mathsf{X}')(\mathsf{M}_S^W)'\right)^+\mathsf{M}_S^W[\mathsf{Y}+\mathsf{XB}'(\mathsf{BB}')^-\mathsf{b}]-\mathsf{M}_S^W\mathsf{XB}'(\mathsf{BB}')^-\mathsf{b}.$$

$$\mathrm{var}[\widehat{\mathsf{M}_S^W\mathsf{X}\beta}] =$$

$$= \mathsf{M}_S^W\mathsf{X}\left\{\left[\mathsf{M}_{B'}\mathsf{X}'(\mathsf{M}_S^W)'\left(\mathsf{M}_S^W[\Sigma+\mathsf{XM}_{B'}\mathsf{X}'](\mathsf{M}_S^W)'\right)^-\mathsf{M}_S^W\mathsf{XM}_{B'}\right]^+ - \mathsf{M}_{B'}\right\}\mathsf{X}'(\mathsf{M}_S^W)'.$$

Remark 3 If the matrix $\Sigma+\mathsf{XM}_{B'}\mathsf{X}'$ is regular or if

$$\mathscr{M}(\mathsf{S}) \subset \mathscr{M}(\Sigma+\mathsf{XM}_{B'}\mathsf{X}'),$$

it can be proved that (see [6], Lemma 1)

$$(\mathsf{M}_S^W)'\left[\mathsf{M}_S^W(\Sigma+\mathsf{XM}_{B'}\mathsf{X}')(\mathsf{M}_S^W)'\right]^+\mathsf{M}_S^W = \left[\mathsf{M}_S(\Sigma+\mathsf{XM}_{B'}\mathsf{X}')\mathsf{M}_S\right]^+.$$

Then

$$\widehat{\mathsf{M}_S^W\mathsf{X}\beta} = \mathsf{M}_S^W\mathsf{X}\left(\mathsf{M}_{B'}\mathsf{X}'[\mathsf{M}_S(\Sigma+\mathsf{XM}_{B'}\mathsf{X}')\mathsf{M}_S]^+\mathsf{XM}_{B'}\right)^+\mathsf{X}'[\mathsf{M}_S(\Sigma+\mathsf{XM}_{B'}\mathsf{X}')\mathsf{M}_S]^+$$
$$\times(\mathsf{Y}+\mathsf{XB}'(\mathsf{BB}')^-\mathsf{b}) - \mathsf{M}_S^W\mathsf{XB}'(\mathsf{BB}')^-\mathsf{b}.$$

$$\mathrm{var}[\widehat{\mathsf{M}_S^W\mathsf{X}\beta}] = \mathsf{M}_S^W\mathsf{X}\left\{\left(\mathsf{M}_{B'}\mathsf{X}'[\mathsf{M}_S(\Sigma+\mathsf{XM}_{B'}\mathsf{X}')\mathsf{M}_S]^+\mathsf{XM}_{B'}\right)^+ - \mathsf{M}_B'\right\}\mathsf{X}'(\mathsf{M}_S^W)'.$$

When we choose transformation matrix

b) $\mathsf{T} = \mathsf{M}_S^{(M_X\Sigma M_X)^+}$,

we get the model with type I constraints with the same design matrix belonging to the vector β,

$$\begin{pmatrix}\mathsf{M}_S^{(M_X\Sigma M_X)^+}\mathsf{Y}\\ -\mathsf{b}\end{pmatrix} \sim \left[\begin{pmatrix}\mathsf{X}\\ \mathsf{B}\end{pmatrix}\beta, \begin{pmatrix}\mathsf{M}_S^{(M_X\Sigma M_X)^+}\Sigma(\mathsf{M}_S^{(M_X\Sigma M_X)^+})', & \mathsf{O}\\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\right],$$

because it is
$$\mathsf{M}_S^{(M_X\Sigma M_X)^+}\mathsf{S}=\mathsf{O},\qquad \mathsf{M}_S^{(M_X\Sigma M_X)^+}\mathsf{X}=\mathsf{X}.$$
According to assumption (6) it should be
$$\mathscr{M}(\mathsf{S}')=\mathscr{M}(\mathsf{S}'[\mathsf{M}_X\Sigma\mathsf{M}_X]^+\mathsf{S}).$$
It is valid if the model (1) is regular (see [3], page 189).

In this model
$$\widehat{\mathsf{X}\beta}=\mathsf{P}_{XM_{B'}}^{[M_S^{(M_X\Sigma M_X)^+}\Sigma(M_S^{(M_X\Sigma M_X)^+})'+XM_{B'}X']^+}\mathsf{M}_S^{(M_X\Sigma M_X)^+}\mathsf{Y}$$
$$-\mathsf{M}_{XM_{B'}}^{[M_S^{(M_X\Sigma M_X)^+}\Sigma(M_S^{(M_X\Sigma M_X)^+})'+XM_{B'}X']^+}\mathsf{X}\mathsf{B}'(\mathsf{B}\mathsf{B}')^-\mathsf{b},$$
$$\mathrm{var}[\widehat{\mathsf{X}\beta}]=$$
$$=\mathsf{X}\Big\{\Big[\mathsf{M}_{B'}\mathsf{X}'\Big(\mathsf{M}_S^{(M_X\Sigma M_X)^+}\Sigma(\mathsf{M}_S^{(M_X\Sigma M_X)^+})'+\mathsf{X}\mathsf{M}_{B'}\mathsf{X}'\Big)^-\mathsf{X}\mathsf{M}_{B'}\Big]^+-\mathsf{M}_{B'}\Big\}\mathsf{X}'.$$
If we suppose, that
$$\mathscr{M}(\mathsf{X}')\ \subset\ \mathscr{M}(\mathsf{X}'[\mathsf{M}_S\Sigma\mathsf{M}_S]^+\mathsf{X}),\tag{9}$$
we can use transformation matrix

c) $\mathsf{T}=\mathsf{P}_X^{(M_S\Sigma M_S)^+}$

that leads to the model
$$\begin{pmatrix}\mathsf{P}_X^{(M_S\Sigma M_S)^+}\mathsf{Y}\\ -\mathsf{b}\end{pmatrix}\sim\left[\begin{pmatrix}\mathsf{X}\\ \mathsf{B}\end{pmatrix}\beta,\begin{pmatrix}\mathsf{X}(\mathsf{X}'[\mathsf{M}_S\Sigma\mathsf{M}_S]^+\mathsf{X})^-\mathsf{X}', & \mathsf{O}\\ \mathsf{O}, & \mathsf{O}\end{pmatrix}\right],$$
because under assumption (9) it is
$$\mathsf{P}_X^{(M_S\Sigma M_S)^+}\mathsf{X}=\mathsf{X},\qquad \mathsf{P}_X^{(M_S\Sigma M_S)^+}\mathsf{S}=\mathsf{O},$$
$$\mathsf{P}_X^{(M_S\Sigma M_S)^+}\Sigma\Big(\mathsf{P}_X^{(M_S\Sigma M_S)^+}\Big)'=\mathsf{X}(\mathsf{X}'[\mathsf{M}_S\Sigma\mathsf{M}_S]^+\mathsf{X})^-\mathsf{X}'.$$
$$\widehat{\mathsf{X}\beta}=\mathsf{P}_{XM_{B'}}^{[X(X'[M_S\Sigma M_S]^+X)^-X'+XM_{B'}X']^+}\mathsf{P}_X^{(M_S\Sigma M_S)^+}\mathsf{Y}$$
$$-\mathsf{M}_{XM_{B'}}^{[X(X'[M_S\Sigma M_S]^+X)^-X'+XM_{B'}X']^+}\mathsf{X}\mathsf{B}'(\mathsf{B}\mathsf{B}')^-\mathsf{b}$$
$$=\Big\{\mathsf{X}\mathsf{M}_{B'}\Big[\mathsf{M}_{B'}\mathsf{X}'\big(\mathsf{X}[\mathsf{X}'(\mathsf{M}_S\Sigma\mathsf{M}_S)^+\mathsf{X}]^-\mathsf{X}'+\mathsf{X}\mathsf{M}_{B'}\mathsf{X}'\big)^+\mathsf{X}\mathsf{M}_{B'}\Big]^-\mathsf{M}_{B'}\mathsf{X}'-\mathsf{X}\mathsf{M}_{B'}\mathsf{X}'\Big\}$$
$$\times(\mathsf{M}_S\Sigma\mathsf{M}_S)^+[\mathsf{Y}+\mathsf{X}\mathsf{B}'(\mathsf{B}\mathsf{B}')^-\mathsf{b}]-\mathsf{X}\mathsf{B}'(\mathsf{B}\mathsf{B}')^-\mathsf{b}.$$
$$\mathrm{var}[\widehat{\mathsf{X}\beta}]=\mathsf{X}\Big\{\Big[\mathsf{M}_{B'}\mathsf{X}'\big(\mathsf{X}[\mathsf{X}'(\mathsf{M}_S\Sigma\mathsf{M}_S)^+\mathsf{X}]^-\mathsf{X}'+\mathsf{X}\mathsf{M}_{B'}\mathsf{X}'\big)^-\mathsf{X}\mathsf{M}_{B'}\Big]^+-\mathsf{M}_{B'}\Big\}\mathsf{X}'.$$

Remark 4 In the practice we have to decide, whether to use transformation or not. We should compute variance matrices of the estimators in the original model and in the transformed model and decide according to the accuracy of the estimates. We can use following formulas:

a) if the model (1) is regular, then under condition (2) without transformation (see Remark 1)

$$\mathrm{var}(\widehat{\mathsf{X}\beta}) = \mathsf{X}[\mathsf{M}_{B'}\mathsf{X}'(\mathsf{M}_S\Sigma_\vartheta\mathsf{M}_S)^+\mathsf{X}\mathsf{M}_{B'}]^+\mathsf{X}',$$

b) in the singular model (1) with constraints (2) without transformation (see Remark 1)

$$\mathrm{var}[\widehat{\mathsf{X}\beta}] = \mathsf{P}_{(XM_{B'},S)}^{[\Sigma_\vartheta+XM_{B'}X'+SS']^+}\Sigma_\vartheta\left(\mathsf{P}_{(XM_{B'},S)}^{[\Sigma_\vartheta+XM_{B'}X'+SS']^+}\right)',$$

c) in the transformed singular model (4) (see Theorem 2)

$$\mathrm{var}[\widehat{\mathsf{TX}\beta}] = \mathsf{TX}\{[\mathsf{M}_{B'}\mathsf{X}'\mathsf{T}'(\mathsf{T}[\Sigma+\mathsf{X}\mathsf{M}_{B'}\mathsf{X}']\mathsf{T}')^-\mathsf{TX}\mathsf{M}_{B'}]^+ - \mathsf{M}_{B'}\}\mathsf{X}'\mathsf{T}'.$$

References

[1] Fišerová, E., Kubáček, L., Kunderová, P.: Linear Statistical Models: Regularity and Singularities. *Academia, Praha*, in preparation.

[2] Kubáček, L.: Foundations of estimation theory. *Elsevier, Amsterdam, Oxford, New York, Tokyo*, 1988.

[3] Kubáček, L., Kubáčková, L., Volaufová, J.: Statistical Models with Linear Structures. *Veda, Publishing House of the Slovak Academy of Sciences, Bratislava*, 1995.

[4] Kubáček, L., Kubáčková, L.: Statistika a metrologie. *Vydavatelství UP, Olomouc*, 2000.

[5] Kunderová, P.: *Regular linear model with the nuisance parameters with constraints of the type I.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **40** (2001), 151–159.

[6] Kunderová, P.: *Eliminating transformations for nuisance parameters in linear model.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **42** (2003), 59–68.

[7] Nordström, K., Fellman, J.: *Characterizations and dispersion-matrix robustness of efficiently estimable parametric functionals in linear models with nuisance parameters.* Linear Algebra and its Applications **127** (1990), 341–361.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 65–73

Ultimate Boundedness Results for a Certain Third Order Nonlinear Matrix Differential Equations

M. O. OMEIKE

Department of Mathematical Sciences, University of Agriculture,
Abeokuta, Nigeria
e-mail: moomeike@yahoo.com

(Received April 25, 2006)

Abstract

This paper extends some known results on the boundedness of solutions and the existence of periodic solutions of certain vector equations to matrix equations.

Key words: Matrix differential equation; Lyapunov function; boundedness.

2000 Mathematics Subject Classification: 34D40, 34D20.

1 Introduction

Let $\mathcal{M}$ denote the space of all real $n \times n$ matrices, $\mathbb{R}^n$ the real n-dimensional Euclidean space and $\mathbb{R}$ the real line $-\infty < t < \infty$. We shall be concerned here with certain properties of solutions of differential equations of the form

$$\dddot{X} + A\ddot{X} + B\dot{X} + H(X) = P(t, X, \dot{X}, \ddot{X}) \tag{1.1}$$

where $X : \mathbb{R} \to \mathcal{M}$ is the unknown, $A, B \in \mathcal{M}$ are constants, $H : \mathcal{M} \to \mathcal{M}$ and $P : \mathbb{R} \times \mathcal{M} \times \mathcal{M} \times \mathcal{M} \to \mathcal{M}$. The specific properties we shall be interested in are the ultimate boundedness of all solutions and the existence of periodic solutions when P is periodic in t.

In [8], Tejumola establishes conditions under which all solutions of the matrix differential equation,

$$\ddot{X} + A\dot{X} + H(X) = P(t, X, \dot{X}), \tag{1.2}$$

are stable, bounded and periodic (depending on the choice of P). These results are extended to the equation (1.1).

For the special case in which (1.1) is an n-vector equation (so that $X : \mathbb{R} \to \mathbb{R}^n$, $H : \mathbb{R}^n \to \mathbb{R}^n$ and $P : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$) a number of boundedness, stability and existence of periodic solutions results have been established by Ezeilo and Tejumola [4], Afuwape [1] , Meng [5] and others for a number of various vector third order differential equations. The conditions obtained in each of these previous investigations are generalizations of the well-known Routh–Hurwitz conditions

$$a > 0, \quad c > 0, \quad ab - c > 0 \tag{1.3}$$

for the stability of the trivial solution of the linear differential equation

$$\dddot{x} + a\ddot{x} + b\dot{x} + cx = 0 \tag{1.4}$$

with constant coefficients. Our present investigations are akin to those of Tejumola [8], Meng [5], Afuwape [1] and we shall provide extensions of their results to matrix differential equations of the form (1.1).

2 Notations and definitions

Some standard matrix notation will be used. For any $X \in \mathcal{M}$, X^T and x_{ij}, $i, j = 1, 2, \ldots, n$ denote the transpose and the elements of X respectively while $(x_{ij})(y_{ij})$ will sometimes denote the product matrix XY of the matrices $X, Y \in \mathcal{M}$. $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ and $X^j = (x_{1j}, x_{2j}, \ldots, x_{nj})$ stand for the i-th row and j-th column of X respectively and $\underline{X} = (X_1, X_2, \ldots, X_n)$ is the n^2 column vector consisting of the n rows of X.

We shall denote by $JH(X)$ the $n^2 \times n^2$ generalised Jacobian matrix associated with the function $H : \mathcal{M} \to \mathcal{M}$ and evaluated at X : that is, $JH(X)$ is the matrix associated with the Jacobian determinant $\frac{\partial(H_1,H_2,\ldots,H_n)}{\partial(X_1,X_2,\ldots,X_n)}$. Corresponding to the constant matrix $A \in \mathcal{M}$ we define an $n^2 \times n^2$ matrix $\tilde{A}$ consisting of n^2 diagonal $n \times n$ matrix$(a_{ij}I_n)(I_n$ being the unit $n \times n$ matrix) and such that $(a_{ij}I_n)$ belongs to the i-th n row and j-th n column (that is, counting n at a time) of $\tilde{A}$. In the special case $n = 2$, $\tilde{A}$ is the 4×4 matrix

$$\begin{pmatrix} a_{11}I_2 & a_{12}I_2 \\ a_{21}I_2 & a_{22}I_2 \end{pmatrix}.$$

Next we introduce an inner product $\langle ., . \rangle$ and a norm $\| \cdot \|$ on $\mathcal{M}$ as follows. For arbitrary $X, Y \in \mathcal{M}$, $\langle X, Y \rangle = \text{trace}\, XY^T$. It is easy to check that $\langle X, Y \rangle = \langle Y, X \rangle$ and that $\|X - Y\|^2 = \langle X - Y, X - Y \rangle$ defines a norm of $\mathcal{M}$. Indeed,

$\|X\| = |\underline{X}|_{n^2}$ where $|\cdot|_{n^2}$ denotes the usual Euclidean norm in $\mathbb{R}^{n^2}$ and $\underline{X} \in \mathbb{R}^{n^2}$ is as defined above.

Lastly the symbol δ, with or without subscripts, denote finite positive constants whose magnitudes depend only on A, B, H and P. Any δ, with a subscript, retains a fixed identity throughout while the unnumbered ones are not necessarily the same each time they occur.

3 Statement of results

It will be assumed throughout the sequel that $H \in C'(\mathcal{M})$ and that $P \in C(\mathbb{R} \times \mathcal{M} \times \mathcal{M} \times \mathcal{M})$. Further, H and P satisfy conditions for the existence of solutions of (1.1) for any set of preassigned initial conditions.

Theorem 1 *Let $H(0) = 0$ and suppose that*

(i) the Jacobian matrix $JH(X)$ of $H(X)$ is symmetric and furthermore that the eigenvalues $\lambda_i(JH(X))$ of $JH(X)$, $(i = 1, 2, \ldots, n^2)$ satisfy for $X \in \mathcal{M}$,

$$0 < \delta_h \leq \lambda_i(JH(X)) \leq \Delta_h \tag{3.1}$$

where δ_h, Δ_h are finite constants;

(ii) the matrices $\tilde{A}, \tilde{B}, JH(X)$ are associative and commute pairwise. The eigenvalues $\lambda_i(\tilde{A})$ of $\tilde{A}$ and $\lambda_i(\tilde{B})$ of $\tilde{B}$ $(i = 1, 2, \ldots, n^2)$ satisfy

$$0 < \delta_a \leq \lambda_i(\tilde{A}) \leq \Delta_a \tag{3.2}$$

$$0 < \delta_b < \lambda_i(\tilde{B}) \leq \Delta_b \tag{3.3}$$

where $\delta_a, \delta_b, \Delta_a, \Delta_b$ are finite constants. Furthermore,

$$\Delta_h \leq k\delta_a\delta_b, \tag{3.4}$$

where

$$k = \min\left\{\frac{\alpha(1-\beta)\delta_b}{\delta_a(\alpha+\Delta_a)^2}; \frac{\alpha(1-\beta)\delta_a}{2(\delta_a+2\alpha)^2}\right\} \tag{3.5}$$

$\alpha > 0$, $0 < \beta < 1$ are some constants,

(iii) P satisfies

$$\|P(t, X, Y, Z)\| \leq \delta_0 + \delta_1(\|X\| + \|Y\| + \|Z\|) \tag{3.6}$$

for all arbitrary $X, Y, Z \in \mathcal{M}$, where $\delta_0 \geq 0$, $\delta_1 \geq 0$ are constants and δ_1 is sufficiently small.

Then every solution $X(t)$ of (1.1) satisfies

$$\|X(t)\| \leq \Delta_1, \quad \|\dot{X}(t)\| \leq \Delta_1, \quad \|\ddot{X}(t)\| \leq \Delta_1 \tag{3.7}$$

for all t sufficiently large, where Δ_1 is a constant the magnitude of which depends only on δ_0, δ_1, A, B, H and P.

This result provides an extension of a result of Afuwape [1], and Meng [5] for an n-vector.

Theorem 2 *Suppose, further to the conditions of Theorem 1, that P satisfies $P(t, X, Y, Z) = P(t + \omega, X, Y, Z)$ uniformly for all $X, Y, Z \in \mathcal{M}$. Then (1.1) admits of at least one periodic solution with period ω.*

4 Some preliminaries

The following results will be basic to the proofs of Theorems 1 and 2.

Lemma 1 [8] *Let $H(0) = 0$ and assume that the matrices $\tilde{A}$ and $JH(X)$ are symmetric and commute for all $X \in \mathcal{M}$. Then*

$$\langle H(X), AX\rangle = \int_0^1 \underline{X}^T \tilde{A} JH(\sigma X)\underline{X} d\sigma.$$

Lemma 2 [1] *Let D be a real symmetric $\ell \times \ell$ matrix, then for any $X \in \mathbb{R}^\ell$ we have*

$$\delta_d \|X\|^2 \leq \langle DX, X\rangle \leq \Delta_d \|X\|^2,$$

where δ_d, Δ_d are the least and greatest eigenvalues of D, respectively.

Lemma 3 [1] *Let Q, D be any two real $\ell \times \ell$ commuting symmetric matrices. Then*

(i) the eigenvalues $\lambda_i(QD)$ $(i = 1, 2, \ldots, \ell)$ of the product matrix QD are all real and satisfy

$$\max_{i \leq j,k \leq \ell} \lambda_j(Q)\lambda_k(D) \geq \lambda_i(QD) \geq \min_{1 \leq j,k \leq \ell} \lambda_j(Q)\lambda_k(D);$$

(ii) the eigenvalues $\lambda_i(Q + D)$ $(i = 1, 2, \ldots, \ell)$ of the sum of matrices Q and D are real and satisfy

$$\left\{\max_{i \leq j \leq \ell} \lambda_j(Q) + \max_{1 \leq k \leq \ell} \lambda_k(D)\right\} \geq \lambda_i(Q+D) \geq \left\{\min_{1 \leq j \leq \ell} \lambda_j(Q) + \min_{1 \leq k \leq \ell} \lambda_k(D)\right\}.$$

Proof of Theorem 1 Let us for convenience, replace Eq.(1.1) by the equivalent system form

$$\begin{aligned} \dot{X} &= Y, \\ \dot{Y} &= Z, \\ \dot{Z} &= -AZ - BY - H(X) + P(t, X, Y, Z). \end{aligned} \tag{4.1}$$

Our main tool in the proof is the scalar Lyapunov function

$$V : \mathcal{M} \times \mathcal{M} \times \mathcal{M} \to \mathbb{R}$$

adapted from [5] and defined for any function $X, Y, Z \in \mathcal{M}$ by

$$\begin{aligned}2V = &\big\{\langle \beta(1-\beta)BX, BX\rangle + \langle 2\alpha A^{-1}BY, Y\rangle + \langle \beta BY, Y\rangle \\ &+ \langle \alpha A^{-1}Z, Z\rangle + \langle \alpha(Z+AY), Y + A^{-1}Z\rangle \\ &\langle Z + AY + (1-\beta)BX, Z + AY + (1-\beta)BX\rangle\big\} \end{aligned} \tag{4.2}$$

where $\alpha > 0$, $0 < \beta < 1$ are some constants. For each term of this function it is clear that

$$\beta(1-\beta)\delta_b\|X\|^2 \le \langle \beta(1-\beta)BX, BX\rangle = \beta(1-\beta)\sum_{i=1}^{n}|BX^i|_n^2 \le \beta(1-\beta)\Delta_b\|X\|^2, \tag{4.3a}$$

$$2\alpha\Delta_a^{-1}\delta_b\|Y\|^2 \le \langle 2\alpha A^{-1}BY, Y\rangle = 2\alpha\sum_{i=1}^{n}|A^{-1}BY^i|_n^2 \le 2\alpha\delta_a^{-1}\Delta_b\|Y\|^2. \tag{4.3b}$$

In a similar manner,

$$\beta\delta_b\|Y\| \le \langle \beta BY, Y\rangle = \beta\sum_{i=1}^{n}|BY^i|_n^2 \le \beta\Delta_b\|Y\|^2, \tag{4.3c}$$

$$\alpha\Delta_a^{-1}\|Z\|^2 \le \langle \alpha A^{-1}Z, Z\rangle \le \alpha\delta_a^{-1}\|Z\|^2, \tag{4.3d}$$

$$0 \le \langle \alpha(Z+AY), Y + A^{-1}Z\rangle \le \nu(\|Y\|^2 + \|Z\|^2), \tag{4.3e}$$

and

$$\begin{aligned}&0 \le \langle Z + AY + (1-\beta)BX, Z + AY + (1-\beta)BX\rangle \\ &= \sum_{i=1}^{n}|Z^i + AY^i + (1-\beta)BX^i|_n^2 \le \mu(\|Z\|^2 + \|Y\|^2 + \|X\|^2),\end{aligned} \tag{4.3f}$$

for some positive constants ν, μ. The estimates above are valid since

$$\sum_{i=1}^{n}|X^i|_n^2 = \sum_{i=1}^{n}|X_i|_n^2 = |\underline{X}|_{n^2}^2 \qquad \text{for any } X \in \mathcal{M}.$$

Combining these estimates (4.3a–4.3f) in (4.2) we obtain that

$$\delta_2(\|X\|^2 + \|Y\|^2 + \|Z\|^2) \le 2V \le \delta_3(\|X\|^2 + \|Y\|^2 + \|Z\|^2), \tag{4.4}$$

$$\delta_2 = \min\{\beta(1-\beta)\delta_b; 2\alpha\Delta_a^{-1}\delta_b + \beta\delta_b; \alpha\Delta_a^{-1}\}$$

and

$$\delta_3 = \max\{\beta(1-\beta)\Delta_b + \mu;\ 2\alpha\delta_a^{-1}\Delta_b + \beta\Delta_b + \nu + \mu;\ \alpha\delta_a^{-1} + \nu + \mu\}.$$

From (4.4), we have that $V(X, Y, Z) \to \infty$ as $\|X\|^2 + \|Y\|^2 + \|Z\|^2 \to \infty$.

To prove our result, it suffices to prove that there exists a constant $\Delta_1 \geq 1$ such that

$$\|X\|^2 + \|Y\|^2 + \|Z\|^2 \leq \Delta_1, \quad \text{for } t \geq T(X_0, Y_0, Z_0), \tag{4.5}$$

for any solution (X, Y, Z) for (4.1), $(X_0 = X(0),\ Y_0 = Y(0),\ Z_0 = Z(0))$.

Let (X, Y, Z) be any solution of (4.1), then the total derivative of V with respect to t along this solution path is

$$\dot{V} = -U_1 - U_2 - U_3 + U_4 \tag{4.6}$$

where

$$U_1 = \left\langle \frac{1-\beta}{2} BX, H(X) \right\rangle + \langle \beta ABY, Y\rangle + \left\langle \frac{\alpha}{2} Z, Z \right\rangle$$

$$U_2 = \left\langle \frac{1-\beta}{2} BY, H(X) \right\rangle + \langle \alpha BY, Y\rangle + \langle (A + \alpha I)Y, H(X)\rangle$$

$$U_3 = \left\langle \frac{1-\beta}{4} BX, H(X) \right\rangle + \left\langle \frac{\alpha}{2} Z, Z \right\rangle + \langle (I + 2\alpha A^{-1})Z, H(X)\rangle$$

$$U_4 = \left\langle (1-\beta)BX + (A + \alpha I)Y + (I + 2\alpha A^{-1})Z, P(t, X, Y, Z) \right\rangle.$$

To arrive at (4.5), we first prove the following:

Lemma 4 *Subject to a conveniently chosen value for k in (3.5), we have for all X, Y, Z*

$$U_j \geq 0, \quad (j = 2, 3).$$

Proof For strictly positive constants k_1, k_2 conveniently chosen later, we have

$$\begin{aligned}\langle (\alpha I + A)Y, H(X)\rangle &= \left\| k_1 (\alpha I + A)^{1/2} Y + 2^{-1} k_1^{-1} (\alpha I + A)^{1/2} H(X) \right\|^2 \\ &- \langle k_1^2 (\alpha I + A)Y, Y\rangle - 4^{-1} k_1^{-2} \langle (\alpha I + A)H(X), H(X)\rangle \end{aligned} \tag{4.7a}$$

and

$$\begin{aligned} &\langle (I + 2\alpha A^{-1})Z, H(X)\rangle = \\ &= \left\| k_2 \left(I + 2\alpha A^{-1}\right)^{1/2} Z + 2^{-1} k_2^{-1} \left(I + 2\alpha A^{-1}\right)^{1/2} H(X) \right\|^2 \\ &- \langle k_2^2 (I + 2\alpha A^{-1})Z, Z\rangle - \left\langle 4^{-1} k_2^{-2} (I + 2\alpha A^{-1})H(X), H(X) \right\rangle, \end{aligned} \tag{4.7b}$$

thus,

$$U_2 = \|k_1(\alpha I + A)^{1/2} Y + 2^{-1} k_1^{-1} (\alpha I + A)^{1/2} H(X)\|^2$$
$$\langle 4^{-1}(1-\beta)BX - 4^{-1} k_1^{-1} (\alpha I + A)H(X), H(X)\rangle + \langle (\alpha B - k_1^2 (\alpha I + A)Y, Y\rangle$$

and

$$U_3 = \|k_2(I+2\alpha A^{-1})^{1/2}Z + 2^{-1}k_2^{-1}(I+2\alpha A^{-1})^{1/2}H(X)\|^2$$
$$+ \langle (1-\beta)4^{-1}BX - 4^{-1}k_2^{-2}(I+2\alpha A^{-1})H(X), H(X)\rangle$$
$$+ \left\langle \left[\frac{\alpha}{2}I - k_2^2(I+2\alpha A^{-1})\right] Z, Z \right\rangle .$$

By Lemmas 1,2 and 3, we obtain

$$U_2 \geq \left\{ \int_0^1 \sigma \int_0^1 \underline{X}^T \left[\frac{1-\beta}{4}\tilde{B} - \frac{1}{4k_1^2}\left(\alpha\tilde{I} + \tilde{A}\right) JH(\sigma X)\right] JH(\tau\sigma X)\underline{X} d\tau d\sigma \right.$$
$$\left. + \underline{Y}^T \left[\alpha\tilde{B} - k_1^2(\alpha\tilde{I} + \tilde{A})\right] \underline{Y} \right\}, \tag{4.8a}$$

and

$$U_3 \geq \left\{ \int_0^1 \sigma \int_0^1 \underline{X}^T \left[\frac{1-\beta}{4}\tilde{B} - \frac{1}{4k_2^2}\left(\alpha\tilde{I} + 2\alpha\tilde{A}^{-1}\right) JH(\sigma X)\right] JH(\tau\sigma X)\underline{X} d\tau d\sigma \right.$$
$$\left. + \underline{Z}^T \left[\frac{\alpha}{2}\tilde{I} - k_2^2(\tilde{I} + 2\alpha\tilde{A}^{-1})\right] \underline{Z} \right\}. \tag{4.8b}$$

Furthermore, by using Lemmas 2 and 3, we obtain

$$U_2 \geq \left\{ \delta_h \left[\frac{1-\beta}{4}\delta_b - \frac{1}{4k_1^2}(\alpha+\Delta_a)\Delta_h\right] \|X\|^2 + \left[\alpha\delta_b - k_1^2(\alpha+\Delta_a)\right] \|Y\|^2 \right\}, \tag{4.8c}$$

and

$$U_3 \geq \left\{ \delta_h \left[\frac{1-\beta}{4}\delta_b - \frac{1}{4k_2^2}(1+2\alpha\delta_a^{-1})\Delta_h\right] \|X\|^2 + \left[\frac{\alpha}{2} - k_2^2(1+2\alpha\delta_a^{-1})\right] \|Z\|^2 \right\}, \tag{4.8d}$$

Thus, using (3.1), (3.2), (3.3) we obtain, for all $X, Y \in \mathcal{M}$,

$$U_2 \geq 0 \tag{4.9a}$$

if $k_1^2 \leq \frac{\alpha\delta_b}{\alpha+\Delta_a}$ with

$$\Delta_h \leq \frac{k_1^2(1-\beta)\delta_b}{(\alpha+\Delta_a)} \leq \frac{\alpha(1-\beta)\delta_b^2}{(\alpha+\Delta_a)^2} \tag{4.10a}$$

and for all X, Z in $\mathcal{M}$,

$$U_3 \geq 0 \tag{4.9b}$$

if $k_2^2 \leq \frac{\alpha\delta_a}{2(\delta+2\alpha)}$ with

$$\Delta_h \leq \frac{k_2^2(1-\beta)\delta_a\delta_b}{(2\alpha+\delta_a)} \leq \frac{\alpha(1-\beta)\delta_a^2\delta_b}{2(2\alpha+\delta_a)^2}. \tag{4.10b}$$

Combining all the inequalities in (4.9) and (4.10), we have inequalities (3.4) with (3.5) satisfied. Thus, for all $X, Y, Z \in \mathcal{M}$, $U_2 \geq 0$ and $U_3 \geq 0$. This completes the proof of Lemma 4. □

Finally, we are left with estimates for U_1 and U_4. From (4.6), we clearly have

$$U_1 = \frac{1-\beta}{2}\int_0^1 \underline{X}^T\tilde{B}JH(\sigma X)\underline{X}\, d\sigma + \beta\underline{Y}^T\tilde{A}\tilde{B}\underline{Y} + \frac{\alpha}{2}\underline{Z}^T\underline{Z}$$

$$\geq \frac{1-\beta}{2}\delta_b\delta_h\|X\|^2 + \beta\delta_a\delta_b\|Y\|^2 + \frac{\alpha}{2}\|Z\|^2 \geq \delta_4(\|X\|^2 + \|Y\|^2 + \|Z\|^2) \quad (4.11)$$

where

$$\delta_4 = \min\left\{\frac{\delta_b}{2}\delta_h(1-\beta); \beta\delta_a\delta_b; \frac{\alpha}{2}\right\}.$$

Since $P(t, X, Y, Z)$ satisfies (3.6), by Schwarz's inequality, we obtain

$$\begin{aligned}|U_4| &\leq \{(1-\beta)\Delta_b\|X\| + (\alpha+\Delta_a)\|Y\| + (1+2\alpha\delta_a^{-1})\|Z\|\}\|P(t, X, Y, Z)\| \\ &\leq \delta_5(\|X\| + \|Y\| + \|Z\|)[\delta_0 + \delta_1(\|X\| + \|Y\| + \|Z\|)] \\ &\leq 3\delta_1\delta_5(\|X\|^2 + \|Y\|^2 + \|Z\|^2) + 3^{1/2}\delta_0\delta_5(\|X\|^2 + \|Y\|^2 + \|Z\|^2)^{1/2},\end{aligned} \quad (4.12)$$

where

$$\delta_5 = \max\{(1-\beta)\Delta_b; (\alpha+\Delta_a); (1+2\alpha\delta_a^{-1})\}.$$

Combining inequalities (4.9), (4.11) and (4.12) in (4.6), we obtain

$$\dot{V} \leq -2\delta_6(\|X\|^2 + \|Y\|^2 + \|Z\|^2) + \delta_7(\|X\|^2 + \|Y\|^2 + \|Z\|^2)^{1/2}, \quad (4.13)$$

where

$$\delta_6 = \frac{1}{2}(\delta_4 - 3\delta_1\delta_5) \quad \text{and} \quad \delta_7 = 3^{1/2}\delta_0\delta_5.$$

Thus, with $\delta_1 < 3^{-1}\delta_5^{-1}\delta_4$, we have that $\delta_6 > 0$.

If we choose

$$(\|X\|^2 + \|Y\|^2 + \|Z\|^2)^{1/2} \geq \delta_8 = 2\delta_7\delta_6^{-1},$$

inequality (4.13) implies that

$$\dot{V} \leq -\delta_6(\|X\|^2 + \|Y\|^2 + \|Z\|^2). \quad (4.14)$$

Then there exists δ_9 such that

$$\dot{V} \leq -1 \quad \text{if} \quad \|X\|^2 + \|Y\|^2 + \|Z\|^2 \geq \delta_9^2.$$

The remainder of the proof of Theorem 1 may now be obtained by use of the estimates (4.4) and (4.14) and an obvious adaptation of the Yoshizawa type reasoning employed in [5].

Proof of Theorem 2 The proof of this theorem follows as in the proof of [5, Theorem 3].

Acknowledgement The author would like to express sincere thanks to the anonymous referees for their invaluable corrections, comments and suggestions.

References

[1] Afuwape, A. U.: *Ultimate boundedness results for a certain system of third order non-linear differential equation.* J. Math. Annal. Appl. **97** (1983), 140–150.

[2] Afuwape, A. U., Omeike, M. O.: *Further ultimate boundedness of solutions of some system of third order non-linear ordinary differential equations.* Acta Univ. Palacki. Olomuc., Fac. rer. nat., Math. **43** (2004), 7–20.

[3] Browder, F. E.: *On a generalization of the Schauder fixed point theorem.* Duke Math. J. **26** (1959), 291–303.

[4] Ezeilo, J. O. C., Tejumola, H. O.: *Boundedness and periodicity of solutions of a certain system of third order non-linear differential equations.* Annali Mat. Pura. Appl. **74** (1966), 283–316.

[5] Meng, F. W.: *Ultimate boundedness results for a certain system of third order nonlinear differential equations.* J. Math. Anal. Appl. **177** (1993), 496–509.

[6] Omeike, M. O.: *Qualitative Study of solutions of certain n-system of third order non-linear ordinary differential equations.* Ph.D. Thesis, University of Agriculture, Abeokuta, 2005.

[7] Reissig, R., Sansone, G., Conti, R.: Non-linear Differential Equations of higher order. *No-ordhoff International Publishing*, 1974.

[8] Tejumola, H. O.: *On a Lienard type matrix differential equation.* Atti Accad. Naz. Lincei Rendi. Cl. Sci. Fis. Mat. Natur (8) **60**, 2 (1976), 100–107.

[9] Yoshizawa, T.: Stability Theory by Liapunov's Second Method. *The Mathematical Society of Japan*, 1966.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 75–84

Singular Nonlinear Problem for Ordinary Differential Equation of the Second Order*

Irena RACHŮNKOVÁ[1], Jan TOMEČEK[2]

Department of Mathematical Analysis and Applications of Mathematics,
Faculty of Science, Palacký University,
Tomkova 40, 779 00 Olomouc, Czech Republic
e-mail: [1]*rachunko@inf.upol.cz*
[2]*jan_tomecek@seznam.cz*

(Received October 10, 2006)

Abstract

The paper deals with the singular nonlinear problem

$$u''(t) + f(t, u(t), u'(t)) = 0,$$
$$u(0) = 0, \quad u'(T) = \psi(u(T)),$$

where $f \in Car((0,T) \times D)$, $D = (0,\infty) \times \mathbb{R}$. We prove the existence of a solution to this problem which is positive on $(0,T]$ under the assumption that the function $f(t,x,y)$ is nonnegative and can have time singularities at $t = 0$, $t = T$ and space singularity at $x = 0$. The proof is based on the Schauder fixed point theorem and on the method of a priori estimates.

Key words: Singular ordinary differential equation of the second order; lower and upper functions; nonlinear boundary conditions; time singularities; phase singularity.

2000 Mathematics Subject Classification: 34B16, 34B18

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

1 Introduction

We will study a singular boundary value problem with nonlinear boundary conditions

$$u''(t) + f(t, u(t), u'(t)) = 0 \quad \text{for a.e. } t \in [0, T], \tag{1}$$

$$u(0) = 0, \quad u'(T) = \psi(u(T)), \tag{2}$$

where $[0, T] \subset \mathbb{R}$, $D = (0, \infty) \times \mathbb{R}$, f satisfies the Carathéodory conditions on $(0, T) \times D$. The function $f(t, x, y)$ is allowed to have time singularities at $t = 0$, $t = T$ and space singularity at $x = 0$, the function ψ is continuous on $[0, \infty)$.

For a given interval $[a, b] \subset \mathbb{R}$ assume that $L^1[a, b]$ denotes the set of all measurable functions defined a.e. on $[a, b]$ which are Lebesgue integrable on $[a, b]$, equipped with the norm

$$\|u\|_1 = \int_a^b |u(t)| \, \mathrm{d}t \quad \text{for each } u \in L^1[a, b];$$

$C^0[a, b]$ (or $C^1[a, b]$) denotes the set of all functions which are continuous (or have continuous first derivatives) on $[a, b]$, with the norm $\|u\|_\infty = \max\{|u(t)|$ $t \in [a, b]\}$ (or $\|u\|_{C^1[a,b]} = \|u\|_\infty + \|u'\|_\infty$); $AC^1[a, b]$ denotes the set of all functions which have absolutely continuous first derivatives on $[a, b]$. We say that $f : [a, b] \times D \to \mathbb{R}$, $D \subset \mathbb{R}^2$ satisfies the *Carathéodory conditions* on $[a, b] \times D$ if f has the following properties: (i) for each $(x, y) \in D$ the function $f(\cdot, x, y)$ is measurable on $[a, b]$; (ii) for almost each $t \in [a, b]$ the function $f(t, \cdot, \cdot)$ is continuous on D; (iii) for each compact set $K \subset D$ there exists a function $m_K \in L^1[a, b]$ such that $|f(t, x, y)| \le m_K(t)$ for a.e. $t \in [a, b]$ and all $(x, y) \in K$. For the set of functions satisfying the *Carathéodory conditions* on $[a, b] \times D$ we write $Car([a, b] \times D)$. By $f \in Car((0, T) \times D)$ we mean that $f \in Car([a, b] \times D)$ for each $[a, b] \subset (0, T)$ and $f \notin Car([0, T] \times D)$.

Singular problems have been studied by many authors (see [1]–[6] and references therein). For instance a similar problem is considered in [3], where the right-hand side function is continuous and it is allowed to change its sign. Moreover, the singularity of f is possible in space variable x. In this work, we consider the function f, which is non–negative and can have both time and space singularities. Here, we found effective necessary conditions for solvability of the problem (1), (2). The arguments are based on the ideas of the paper [5], where the non–linear singular problem with mixed boundary conditions

$$u'' + f(t, u, u') = 0, \qquad u'(0) = 0, \quad u(T) = 0$$

is investigated.

Definition 1 Let $f \in Car((0, T) \times D)$, where $D = (0, \infty) \times \mathbb{R}$. We say that f has a time singularity at $t = 0$ and/or at $t = T$, if there exists $(x_1, y_1) \in D$ and/or $(x_2, y_2) \in D$ such that

$$\int_0^\epsilon |f(t, x_1, y_1)| \, \mathrm{d}t = \infty \quad \text{and/or} \quad \int_{T-\epsilon}^T |f(t, x_2, y_2)| \, \mathrm{d}t = \infty$$

for each sufficiently small $\epsilon > 0$. The point $t = 0$ and/or $t = T$ will be called a singular point of f.

We say that f has a space singularity at $x = 0$ if

$$\limsup_{x\to 0+} |f(t,x,y)| = \infty \quad \text{for a.e. } t \in [0,T] \text{ and for some } y \in \mathbb{R}.$$

Here, we will treat with following definition of the solution of the problem (1), (2).

Definition 2 By a solution of the problem (1), (2) we understand a function $u \in AC^1[0,T]$ satisfying the differential equation (1) and the boundary conditions (2).

2 Regular problem, lower and upper function

In order to prove the main result we need the existence theorem for regular boundary value problems. Let us consider a problem

$$u'' + h(t,u,u') = 0, \qquad g_1(u(0),u'(0)) = 0, \quad g_2(u(T),u'(T)) = 0, \qquad (3)$$

where $h \in Car([0,T] \times \mathbb{R}^2)$, g_1, $g_2 : \mathbb{R}^2 \to \mathbb{R}$ are continuous functions.

Definition 3 A function $u \in AC^1[0,T]$ which satisfies the differential equation in (3) a. e. in $[0,T]$ and fulfils the boundary conditions in (3) is called a solution of the problem (3).

In the existence theorem the concept of upper and lower function will be needed.

Definition 4 A function $\sigma \in AC^1[0,T]$ is called a lower function of the problem (3) if

$$\sigma''(t) + h(t,\sigma(t),\sigma'(t)) \geq 0 \qquad \text{for a.e. } t \in [0,T]$$

and

$$g_1(\sigma(0),\sigma'(0)) \geq 0, \qquad g_2(\sigma(T),\sigma'(T)) \geq 0.$$

If these inequalities are reversed, the function σ is called an upper function of the problem (3).

For $\sigma_1, \sigma_2 \in AC^1[0,T]$ such that $\sigma_1 \leq \sigma_2$ on $[0,T]$ we can define a function $\gamma : [0,T] \times \mathbb{R} \to \mathbb{R}$ by

$$\gamma(t,x) = \max\{\sigma_1(t), \min\{x, \sigma_2(t)\}\} \quad \text{for each } t \in [0,T],\ x \in \mathbb{R}. \qquad (4)$$

Now, we introduce the following result [7, Lemma 2]. It is fundamental in the proof of Lemma 6.

Lemma 5 *For $u \in C^1[0,T]$ the two following properties hold:*

a) $\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t,u(t))$ *exists for a. e.* $t \in [0,T]$.

b) If $u_m \in C^1[0,T]$ and $u_m \to u$ in $C^1[0,T]$, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t,u_m(t)) \to \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t,u(t)) \quad \textit{for a.e. } t \in [0,T].$$

Lemma 6 *Let $h \in Car([0,T]\times\mathbb{R}^2)$, g_1, $g_2 : \mathbb{R}^2 \to \mathbb{R}$ be continuous functions and σ_1, σ_2 be lower and upper function of the problem* (3)*, respectively, such that*

$$\sigma_1(t) \le \sigma_2(t) \quad \textit{for each } t \in [0,T].$$

Further, assume that there exists $\varphi \in L^1[0,T]$ such that

$$|h(t,x,y)| \le \varphi(t)$$

for a.e. $t \in [0,T]$, each $x \in [\sigma_1(t),\sigma_2(t)]$ and each $y \in \mathbb{R}$, g_1 is nondecreasing in the second variable and g_2 is nonincreasing in the second variable. Then there exists a solution u of the problem (3) *such that*

$$\sigma_1 \le u \le \sigma_2 \quad \textit{on } [0,T]. \tag{5}$$

Proof Let us define functionals A, $B : C^1[0,T] \to \mathbb{R}$ by

$$\begin{aligned} A(u) &= \gamma(0, u(0) + g_1(u(0), u'(0))), \\ B(u) &= \gamma(T, u(T) + g_2(u(T), u'(T))) \end{aligned}$$

for each $u \in C^1[0,T]$. Lemma 5 allows us to define for each $u \in C^1[0,T]$ a function $\tilde{h}_u : [0,T] \to \mathbb{R}$ such that

$$\tilde{h}_u(t) = h(t, \gamma(t,u(t)), \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t,u(t))) \quad \text{for a.e. } t \in [0,T].$$

Obviously, there exists $\bar{h} \in L^1[0,T]$ such that

$$|\tilde{h}_u(t)| \le \bar{h}(t) \qquad \text{for a.e. } t \in [0,T] \text{ and each } u \in C^1[0,T].$$

Consider an auxiliary problem

$$\begin{cases} u''(t) = -\tilde{h}_u(t) & \text{a.e. } t \in [0,T], \\ u(0) = A(u), \\ u(T) = B(u). \end{cases} \tag{6}$$

Let us define a mapping $F : C^1[0,T] \to C^1[0,T]$ by

$$(Fu)(t) = -\int_0^T G(t,s)\tilde{h}_u(s)\,\mathrm{d}s + \frac{T-t}{T}A(u) + \frac{t}{T}B(u)$$

for each $u \in C^1[0,T]$ and $t \in [0,T]$, where

$$G(t,s) = \begin{cases} \frac{t(s-T)}{T} & \text{for } 0 \le t \le s \le T, \\ \frac{s(t-T)}{T} & \text{for } 0 \le s < t \le T. \end{cases}$$

We can check that each fixed point of the operator F is a solution of the problem (6). Using the Schauder fixed point theorem we will prove that there exists a fixed point u of the operator F satisfying the inequalities (5) and such that u is a solution of the problem (3).

It is easy to see that

$$\|Fu\|_\infty \le T\|\bar{h}\|_1 + 2(\|\sigma_1\|_\infty + \|\sigma_2\|_\infty)$$

and

$$\|(Fu)'\|_\infty \le \|\bar{h}\|_1 + \frac{2}{T}(\|\sigma_1\|_\infty + \|\sigma_2\|_\infty),$$

i. e. that there exist $K > 0$ and $\Omega = \{u \in C^1[0,T] : \|u\|_{C^1[0,T]} \le K\}$, such that $F(\Omega) \subset \Omega$. It suffices to prove that the set

$$F' = \{(Fu)' : u \in \Omega\}$$

is relatively compact in $C^0[0,T]$. Obviously, for each $\epsilon > 0$ there exists $\delta > 0$ such that for each $u \in \Omega$ and $s_1,\ s_2 \in [0,T]$, $|s_2 - s_1| < \delta$, relations

$$|(Fu)'(s_2) - (Fu)'(s_1)| = |\int_{s_1}^{s_2} \tilde{h}_u(s)\,\mathrm{d}s| \le |\int_{s_1}^{s_2} \bar{h}(s)\,\mathrm{d}s| < \epsilon$$

are valid. Now, applying Arzelà–Ascoli theorem we get that $F(\Omega)$ is relatively compact in $C^1[0,T]$. Thus, there exists a fixed point u of the operator F and $u \in AC^1[0,T]$. We will prove that relations (5) are satisfied. From boundary conditions in (6) it follows that

$$\sigma_1(0) \le u(0) \le \sigma_2(0) \quad \text{and} \quad \sigma_1(T) \le u(T) \le \sigma_2(T).$$

Assume that there exists $\tau \in (0,T)$ such that $u(\tau) < \sigma_1(\tau)$. Then there exist $\xi \in (0,T)$ and $\delta > 0$ such that

$$(u - \sigma_1)(\xi) = \min_{t \in [0,T]} (u - \sigma_1)(t) < 0$$

and

$$0 > (u - \sigma_1)(t) > (u - \sigma_1)(\xi) \quad \text{for each } t \in (\xi, \xi + \delta). \tag{7}$$

Obviously, $(u - \sigma_1)'(\xi) = 0$ and $u(t) < \sigma_1(t)$ for each $t \in (\xi, \xi + \delta)$. According to the definition of $\tilde{h}_u$, we have

$$(u - \sigma_1)'(t) \le \int_\xi^t [-\tilde{h}_u(s) + h(s, \sigma_1(s), \sigma_1'(s))]\,\mathrm{d}s = 0,$$

for each $t \in (\xi, \xi + \delta)$, which contradicts (7). Similarly, we can prove that $u \le \sigma_2$ on $[0,T]$. From (5) it follows that u satisfies the differential equation in (3). It suffices to prove that u satisfies boundary conditions in (3), i. e. according to (5) and definition of γ, to prove inequalities

$$\sigma_1(0) \le u(0) + g_1(u(0), u'(0)) \le \sigma_2(0) \tag{8}$$

and

$$\sigma_1(T) \le u(T) + g_2(u(T), u'(T)) \le \sigma_2(T).$$

Let the first inequality in (8) be not satisfied. Then according to (5) we have

$$u(0) = \sigma_1(0), \quad 0 > g_1(\sigma_1(0), u'(0)) \quad \text{and} \quad u'(0) \ge \sigma_1'(0).$$

Using the monotonicity of g_1 we have $0 > g_1(\sigma_1(0), \sigma_1'(0))$, which contradicts the definition of a lower function. The remaining inequalities can be proven in a similar way. □

3 Main result

Now, we are ready to prove the existence theorem for singular problem (1), (2).

Theorem 7 *Assume that $f \in Car((0,T) \times D)$, where $T > 0$, $D = (0,\infty) \times \mathbb{R}$, with possible time singularities at $t = 0$ and/or $t = T$ and a space singularity at $x = 0$. Further assume that there exist $\epsilon \in (0,1)$, $\nu \in (0,T)$, $c \in (\nu, \infty)$ and $\epsilon_0 \in (0,\infty)$ such that*

$$f(t, ct, c) = 0 \quad \text{for a.e. } t \in [0,T], \tag{9}$$

$$0 \le f(t,x,y) \quad \text{for a.e. } t \in [0,T],\ \text{each } x \in (0, ct],\ y \in [\min_{t\in[0,cT]} \psi(t), c], \tag{10}$$

$$\epsilon \le f(t,x,y) \quad \text{for a.e. } t \in [T-\nu, T],\ \text{each } x \in (0, ct],\ y \in (-\epsilon_0, \nu], \tag{11}$$

$$0 = \psi(0), \quad \psi(cT) \le c \tag{12}$$

hold. Then there exists a solution u of the problem (1), (2) such that

$$0 < u(t) \le ct \tag{13}$$

for each $t \in (0,T]$.

Proof Step 1. Let $k \in \mathbb{N}$, $k \ge 3/T$. We define

$$\alpha_k(t,x) = \begin{cases} c/k & \text{for } x < c/k, \\ x & \text{for } c/k \le x \le ct, \\ ct & \text{for } x > ct, \end{cases}$$

for each $t \in [1/k, T - 1/k]$, $x \in \mathbb{R}$,

$$\beta(y) = \begin{cases} \min_{t\in[0,cT]} \psi(t) & \text{for } y < \min_{t\in[0,cT]} \psi(t), \\ y & \text{for } \min_{t\in[0,cT]} \psi(t) \le y \le c, \\ c & \text{for } y > c, \end{cases}$$

and

$$\gamma(y) = \begin{cases} \epsilon & \text{for } y < \nu, \\ \epsilon \frac{c-y}{c-\nu} & \text{for } \nu \le y \le c, \\ 0 & \text{for } y > c, \end{cases}$$

for each $y \in \mathbb{R}$ and

$$f_k(t,x,y) = \begin{cases} 0 & \text{for } t \in [0,1/k), \\ f(t,\alpha_k(t,x),\beta(y)) & \text{for } t \in [1/k, T-1/k], \\ \gamma(y) & \text{for } t \in (T-1/k, T], \end{cases}$$

for each x, $y \in \mathbb{R}$. Obviously, $f_k \in Car([0,T] \times \mathbb{R}^2)$ and

$$f_k(t,x,y) \geq 0 \quad \text{for a.e. } t \in [0,T] \text{ and each } x, y \in \mathbb{R}. \tag{14}$$

Let us define regular problem

$$u'' + f_k(t,u,u') = 0, \quad u(0) = 0, \ u'(T) = \psi(u(T)). \tag{15}$$

From relations (9), (12) and (14) it follows that $\sigma_1(t) = 0$ and $\sigma_2(t) = ct$ for $t \in [0,T]$ are lower and upper functions of problems (15), respectively. From Lemma 6 we get a solution u_k of the problem (15) (where we put $h = f_k$, $g_1(x,y) = -x$, $g_2(x,y) = \psi(x) - y$) such that

$$0 \leq u_k(t) \leq ct \quad t \in [0,T]. \tag{16}$$

Obviously, it is valid

$$u_k'(0) \geq 0 \quad \text{and} \quad u_k'(0) = \lim_{t \to 0+} \frac{u_k(t)}{t} \leq c.$$

From (14) it follows that u_k' is nonincreasing on $[0,T]$. These facts, (15) and (16) imply

$$\min_{s \in [0,cT]} \psi(s) \leq \psi(u_k(T)) = u_k'(T) \leq u_k'(t) \leq u_k'(0) \leq c$$

for every $t \in [0,T]$.

Step 2. (A priori estimates) Consider a sequence $\{u_k\}$ from Step 1. We will prove the relation

$$\liminf_{k \to \infty} u_k(T) > 0. \tag{17}$$

Let (17) be not valid, i.e. $\liminf_{k\to\infty} u_k(T) = 0$. From the continuity of ψ and (12) it follows that for each arbitrarily small $\epsilon_1 > 0$ ($\epsilon_1 \leq \epsilon_0$ and $\epsilon_1 \leq \nu$) there exists $\delta > 0$ (we can choose it such that $\delta \leq \epsilon_1$) such that for every $x \in \mathbb{R}$ the implication

$$0 \leq x \leq \delta \implies |\psi(x)| < \epsilon_1$$

holds. Then there exists $l \in \mathbb{N}$ such that

$$0 \leq u_l(T) < \delta \leq \epsilon_1 \quad \text{and} \quad |u_l'(T)| = |\psi(u_l(T))| < \epsilon_1. \tag{18}$$

particularly, $-\epsilon_0 \leq -\epsilon_1 < u_l'(T) \leq u_l'(t)$ for each $t \in [0,T]$ and $u_l'(T) < \epsilon_1 < \nu$. Then there exists $t_l \in (0,T)$ such that $-\epsilon_0 \leq u_l'(t) \leq \nu$ for every $t \in (t_l, T]$.

There are two possibilities. If $t_l \leq T - \nu$, then integrating the differential equation from (15) we get

$$\begin{aligned} u_l'(T) - u_l'(t) &= \int_t^T u_l''(s)\,\mathrm{d}s \\ &= -\int_t^T f_l(s, u_l(s), u_l'(s))\,\mathrm{d}s \leq -\int_t^T \epsilon\,\mathrm{d}s = -\epsilon(T - t) \end{aligned} \tag{19}$$

for every $t \in [T - \nu, T]$. If $t_l > T - \nu$ and $u_l'(t) > \nu$ for every $t \in [T - \nu, t_l)$, then (19) is valid for each $t \in [t_l, T]$. Since $\nu \geq \epsilon(T - t)$ for $t \in [T - \nu, t_l)$, it follows that $u_l'(t) \geq \epsilon(T - t)$ for each $t \in [T - \nu, t_l)$. In both cases we have the inequality

$$u_l'(t) \geq -\epsilon_1 + \epsilon(T - t)$$

for $t \in [T - \nu, T]$. Integrating this relation over the interval $[T - \nu, T]$ we get

$$u_l(T) - u_l(T - \nu) \geq -\epsilon_1 \nu + \frac{\epsilon \nu^2}{2}$$

and according to (16) and (18) (and since $u_l(T - \nu) \geq 0$) we have

$$\frac{\epsilon \nu^2}{2} < \epsilon_1(\nu + 1).$$

Taking ϵ_1 sufficiently small we get a contradiction. Hence (17) is valid. According to the concavity of u_k and (17), there exists $\omega > 0$ such that

$$u_k(t) \geq \omega t \qquad \text{for every } t \in [0, T], \text{ a.e. } k \in \mathbb{N}. \tag{20}$$

Step 3. (Convergence of the sequence $\{u_k\}$) Let u_k be a solution of the problem (15) for each $k \in \mathbb{N}$, $k \geq 3/T$ and $[a, b] \subset (0, T)$ be a compact interval. Then (20) implies that there exists $k_0 \in \mathbb{N}$ such that for every $t \in [a, b]$ and $k \geq k_0$

$$\frac{c}{k_0} \leq u_k(t) \leq ct.$$

There exists $\varphi \in L^1[a, b]$ such that

$$|f_k(t, u_k(t), u_k'(t))| \leq \varphi(t) \quad \text{for a.e. } t \in [a, b]$$

From Arzelà–Ascoli theorem and diagonalization principle it follows that there exists $u \in C^0[0, T]$ such that u' is continuous on $(0, T)$ and a subsequence $\{u_{n_k}\}$ such that

$$\left.\begin{aligned} &u_{n_k} \to u \quad \text{uniformly on } [0, T], \\ &u_{n_k}' \to u' \quad \text{locally uniformly on } (0, T), \quad u_{n_k}'(T) \to \psi(u(T)) \end{aligned}\right\} \tag{21}$$

and $u(0) = 0$. Without any loss of generality we assume that $\{n_k\} = \{k\}$.

Step 4. (Convergence of the approximate problems) Let us take $\xi \in (0,T)$ such that $f(\xi,\cdot,\cdot)$ is continuous on $(0,\infty)\times\mathbb{R}$. Then there exists a compact interval $J^\star \subset (0,T)$ and $k^\star \in \mathbb{N}$ such that $\xi \in J^\star$ and for each $k \geq k_0$

$$u_k(\xi) > \frac{c}{k^\star}, \qquad J^\star \subset \left[\frac{1}{k}, T-\frac{1}{k}\right].$$

Then $f_k(\xi, u_k(\xi), u_k'(\xi)) = f(\xi, u_k(\xi), u_k'(\xi))$. We get assertion

$$\lim_{k\to\infty} f_k(t,u_k(t),u_k'(t)) = f(t,u(t),u'(t)) \quad \text{for a.e. } t \in (0,T). \tag{22}$$

Let $t \in (0,T)$. Then there exists a compact interval $[a,b] \subset (0,T)$ and $\varphi \in L^1[a,b]$ such that $t \in [a,b]$, $T/2 \in [a,b]$ and

$$|f_k(s,u_k(s),u_k'(s))| \leq \varphi(s) \quad \text{for a.e. } s \in [a,b]. \tag{23}$$

Obviously,

$$u_k'\left(\frac{T}{2}\right) - u_k'(t) = \int_{\frac{T}{2}}^{t} f_k(s,u_k(s),u_k'(s))\,\mathrm{d}s.$$

In view of this fact, (21), (22), (23) and Lebesgue dominated convergence theorem we have

$$u'\left(\frac{T}{2}\right) - u'(t) = \int_{\frac{T}{2}}^{t} f(s,u(s),u'(s))\,\mathrm{d}s.$$

Obviously, this inequality is valid for every $t \in (0,T)$. It means that u' is continuous on each compact subinterval of the interval $(0,T)$ and

$$u''(t) + f(t,u(t),u'(t)) = 0 \quad \text{for a.e. } t \in (0,T).$$

For $k \geq 3/T$ we have

$$\int_0^T f_k(s,u_k(s),u_k'(s))\,\mathrm{d}s = u_k'(0) - u_k'(T) = u_k'(0) - \psi(u_k(T)) \leq c - \min_{s\in[0,cT]} \psi(s)$$

From this fact, (14) and Fatou Lemma it follows that $f(\cdot,u(\cdot),u'(\cdot)) \in L^1[0,T]$ and obviously $u \in AC^1[0,T]$. It remains to prove the last boundary condition in (2). For $k \geq 3/T$ and $t \in (0,T)$ we have

$$\begin{aligned} |u_k'(t) - u_k'(T)| \leq &\int_t^T |f(s,u(s),u'(s))|\,\mathrm{d}s \\ &+ \int_t^T |f_k(s,u_k(s),u_k'(s)) - f(s,u(s),u'(s))|\,\mathrm{d}s. \end{aligned}$$

This inequality and (21) imply that for every $\epsilon > 0$ there exists $\delta > 0$ such that for every $t \in (T-\delta, T)$ there exists $k_0 = k_0(\epsilon,t) \in \mathbb{N}$ such that

$$|u'(t) - \psi(u(T))| \leq |u'(t) - u_{k_0}'(t)| + |u_{k_0}'(t) - u_{k_0}'(T)| + |u_{k_0}'(T) - \psi(u(T))| < \epsilon.$$

Thus, $u'(T) = \lim_{t\to T-} u'(t) = \psi(u(T))$. This completes the proof. □

Example 8 Let α, $\beta \in (0, \infty)$. Then, by Theorem 7 the problem

$$u'' + (u^{-\alpha} + u^{\beta} + t^2 + 1)(1 - (u')^3) = 0, \quad u(0) = 0, \quad u'(1) = -(u(1))^2$$

has a solution $u \in AC^1[0,1]$ such that

$$0 < u(t) \le t \quad \text{for each } t \in (0,1].$$

References

[1] Kiguradze, I. T.: On Some Singular Boundary Value Problems for Ordinary Differential Equations. *Tbilisi Univ. Press, Tbilisi*, 1975 (in Russian).

[2] Kiguradze, I. T., Shekhter, B. L.: *Singular boundary value problems for second order ordinary differential equations.* Itogi Nauki i Tekhniki Ser. Sovrem. Probl. Mat. Nov. Dost. **30** (1987), 105–201 (in Russian), translated in J. Soviet Math. **43** (1988), 2340–2417.

[3] O'Regan, D.: *Upper and lower solutions for singular problems arising in the theory of membrane response of a spherical cap.* Nonlinear Anal. **47** (2001), 1163–1174.

[4] O'Regan, D.: Theory of Singular Boundary Value Problems. *World Scientific, Singapore*, 1994.

[5] Rachůnková, I.: *Singular mixed boundary value problem.* J. Math. Anal. Appl. **320** (2006), 611–618.

[6] Rachůnková, I.; Staněk, S.; Tvrdý, M.: *Singularities and Laplacians in Boundary Value Problems for Nonlinear Ordinary Differential Equations.* Handbook of Differential Equations. Ordinary Differential Equations, Ed. by A. Cañada, P. Drábek, A. Fonda, Vol. 3., pp. 607–723, Elsevier, 2006.

[7] Wang, M., Cabada, A., Nieto, J. J.: *Monotone method for nonlinear second order periodic boundary value problems with Carathéodory functions.* Ann. Polon. Math. **58**, 3 (1993), 221–235.

Acta Univ. Palacki. Olomuc., Fac. rer. nat.,
Mathematica **46** (2007) 85–88

Bol-loops of Order $3 \cdot 2^n$

Daniel WAGNER [1], Stefan WOPPERER [2]

[1] *39 Telford House, Tiverton Street, SE1 6NY London, United Kingdom*
e-mail: d.wagner@imperial.ac.uk

[2] *Im Heuschlag 21, 91054 Erlangen, Gemany*
e-mail: stefan.wopperer@gmx.de

(Received July 19, 2006)

Abstract

In this article we construct proper Bol-loops of order $3 \cdot 2^n$ using a generalisation of the semidirect product of groups defined by Birkenmeier and Xiao. Moreover we classify the obtained loops up to isomorphism.

Key words: Bol-loop; loop; group; semidirect product.

2000 Mathematics Subject Classification: 20N05

1 Introduction

Burn proofs in [3] that the smallest proper Bol-loops are of order 8. But they can not be constructed as a semidirect product defined in [1]. The smallest proper Bol-loops which can be constructed using a semidirect product as defined in this article have order 12. Up to isomorphism these loops can be realised as semidirect product of the cyclic group of order 3 and the elementary abelian groups of order 4. There are no proper Bol-loops of order 9, 10 or 11. It seems that order 12 plays an interesting role in the theory of loops since the smallest proper Moufang-loop has also order 12 (cf. [5]).

A *loop* is a set L with a binary operation $\cdot$, a neutral element 1 and unique solutions of the equations $x \cdot a = b$ and $a \cdot x = b$. The loop L is a *left Bol-loop* if $((x \cdot y)z)y = x((y \cdot z)y)$ for all $x, y, z \in L$ holds. Analogously one defines a right Bol-loop by the identity $x(y(x \cdot z)) = (x(y \cdot x))z$.

In this paper we consider a special case of the semidirect product of loops defined by Birkenmeier and Xiao in [2]. Starting with groups N and Q we obtain a loop L on $N \rtimes Q = \{(a, p)\colon a \in N,\ p \in Q\}$. The multiplication $*$ of L is defined as $(a, p) * (b, q) = (a^{\Phi(q)} \circ b^{\Psi(p)}, p \bullet q)$, where $\circ$ and $\bullet$ are the multiplications of N and Q. The mapping $\Phi(p)$ respectively $\Psi(p)$ from N into N is determined by a mapping Φ respectively Ψ from Q into the set of mappings from N into N. According to [2] we know that $(L,*)$ is a loop with neutral element $(1,1)$ if $\Phi(p)$ and $\Psi(p)$ are bijective, $1^{\Phi(p)} = 1^{\Psi(p)} = 1$ holds for all $p \in Q$ and $\Phi(1) = \Psi(1) = \mathrm{id}_N$. The constructed loops are associative if and only if the mappings Ψ, Φ, $\Phi(p)$ and $\Psi(p)$ are homomorphisms and $\Phi(p)$ and $\Psi(q)$ commute for all $p,q \in Q$.

Although the semidirect product treated by us here is a special case of the semidirect products defined in [1], [2] and [9] the construction presented here yields in general loops with no further identities. For example the 15 non-associative loops $L = C_3 \rtimes C_3$ of order 9, which are the smallest possible examples, are not even power associative and only three of them are commutative.

2 Bol-loops of order $3 \cdot 2^n$

We now construct loops of the form $L = C_3 \rtimes (C_2)^n$. These loops are all power-associative and under certain conditions Bol-loops.

Remark 1 The only two mappings of C_3 into C_3 which are one-to-one and keep the neutral element 1 fixed, are the identity and the inversion. Both mappings are automorphisms of C_3 and commute with each other.

Lemma 1 *All loops $L = C_3 \rtimes (C_2)^n$ are power-associative.*

Proof The restriction of Φ and Ψ to a subloop which is generated by a single element is a homomorphism. Therefore L is power-associative by the preceeding Remark. □

Proposition 1 *A semidirect product $L = C_3 \rtimes (C_2)^n$ is a left respectively right Bol-loop if and only if Φ respectively Ψ is a homomorphism.*

Proof Because of Remark 1 the left Bol-identity yields:

$$\left(a^{\Phi(qpr)}\left(b^{\Phi(pr)}\left(a^{\Phi(r)}c^{\Psi(p)}\right)^{\Psi(q)}\right)^{\Psi(p)}, pqpr\right)$$
$$= \left(\left(a^{\Phi(qp)}\left(b^{\Phi(p)}a^{\Psi(q)}\right)^{\Psi(p)}\right)^{\Phi(r)}c^{\Psi(pqp)}, pqpr\right) \tag{1}$$

If L is a left Bol-loop equation (1) implies for $a = c = 1$ that Φ is a homomorphism.

If Φ is a homomorphism we obtain for the first component of (1):

$$a^{\Phi(qpr)}\left(b^{\Phi(pr)}\right)^{\Psi(p)} = \left(a^{\Phi(qp)}\right)^{\Phi(r)}\left(\left(b^{\Phi(p)}\right)^{\Phi(r)}\right)^{\Psi(p)} \tag{2}$$

which is valid for all $a,b \in C_3$ and all $p,q,r \in (C_2)^n$. Therefore L is a left Bol-loop.

The proof for right Bol-loops is analogous. □

To classify the constructed loops up to isomorphism we now determine the order of the non-trivial elements in the loops.

Lemma 2 *Let* $L = C_3 \rtimes (C_2)^n$ *be a loop,* $a \in C_3 \setminus \{1\}$ *and* $p \in (C_2)^n \setminus \{1\}$. *Then the order of* (a, p) *is* 2 *if and only if* $\Phi(p) \neq \Psi(p)$ *and* 6 *if and only if* $\Phi(p) = \Psi(p)$.

Proof If $\Phi(p) \neq \Psi(p)$ then $(a, p)(a, p) = (1, 1)$ holds because of $\Phi(p), \Psi(p) \in \{\mathrm{id}, \mathrm{inv}\}$. If $\Phi(p) = \Psi(p)$ then the first component of $(a, p)^n$ is a power of a or a^{-1}. The second component alternates between 1 and p. Therefore the order of (a, p) is the least common multiple of 2 and 3.

Conversely if $(a, p)(a, p) = (1, 1)$ then $\Phi(p) \neq \Psi(p)$ because of $\Phi(p), \Psi(p) \in \{\mathrm{id}, \mathrm{inv}\}$. Assume the order of (a, p) to be 6 and $\Phi(p) \neq \Psi(p)$. This is a contradiction to the first part of the proof. □

Proposition 2 ($C_3 \rtimes (C_2)^2$**)** *Two proper loops of the form* $C_3 \rtimes (C_2)^2$ *are isomorphic if and only if both loops have the same number of elements with order* 6. *A loop* $L = C_3 \rtimes (C_2)^2$ *is a Bol-loop if and only if it has exactly zero or two elements of order* 6.

Proof Lemma 2 implies that a loop is a Bol-loop if and only if it has exactly zero or two elements of order 6.

Let L_1 and L_2 be loops with the same number of elements with order 6. Then it can be shown that

$$\iota\colon \begin{cases} (a, p) \mapsto (a, p) & \text{if } \Phi_1(p) = \Phi_2(p) \\ (a, p) \mapsto (a^{-1}, p) & \text{if } \Phi_1(p) \neq \Phi_2(p) \end{cases}$$

is an isomorphism between L_1 and L_2. The elements (a, p) with order 6 are assumed to have the same second component $p \in (C_2)^2$ because loops can be transferred in this form by obvious (anti-)isomorphisms. By Lemma 2 this implies that $\Phi_1(p) = \Psi_1(p)$ is equivalent to $\Phi_2(p) = \Psi_2(p)$.

Only the first component has to be analysed to check if ι is an isomorphism. The validity of the equation

$$\iota_{pq}(a^{\Phi_1(q)} b^{\Psi_1(p)}) = (\iota_p(a))^{\Phi_2(q)} (\iota_q(b))^{\Psi_2(p)} \tag{3}$$

is shown by case analysis.

Since $\Phi_1(p)$ in L_1 can be different from $\Phi_2(p)$ in L_2 there are four cases. The mapping Ψ is not considered in the following because it is determined by the order of the elements and the choice of Φ.

First the cases where $\Phi_1(p)$ and $\Phi_2(p)$ are unequal for all p or equal for exactly two elements $p, q \in V_4$: These loops can be trivially antiisomorphic by symmetry of Φ and Ψ. Otherwise they are not (anti-)isomorphic because out of every other pair of loops, which satisfies the preconditions, one and only one loop is a Bol-loop. Therefore in this cases it is not necessary to prove the validity of equation (3).

If there is exactly one element $r \in V_4$ for which $\Phi_1(r) = \Phi_2(r)$, then there are three possibilities, namely $\Phi_1(pq) = \Phi_2(pq)$, $\Phi_1(p) = \Phi_2(p)$ or $\Phi_1(q) = \Phi_2(q)$. In all three cases equation (3) holds for all combinations of $\Phi_1(q)$, $\Psi_1(p)$, $\Phi_2(q)$, $\Psi_2(p) \in \{\mathrm{id}, \mathrm{inv}\}$.

In the last case, which is $\Phi_1(p) = \Phi_2(p)$, $\Phi_1(q) = \Phi_2(q)$ and $\Phi_1(pq) = \Phi_2(pq)$, the validity of equation (3) is obvious. □

Corollary 1 *There are 32 Bol-loops of the form $C_3 \rtimes (C_2)^2$ which are distributed in two classes of isomorphism.*

Theorem 1 **($C_3 \rtimes (C_2)^n$)** *Two proper Bol-loops of the form $C_3 \rtimes (C_2)^n$ are isomorphic if and only if they have the same number of elements with order 6.*

Proof If L_1 and L_2 are proper Bol-loops of the form $C_3 \rtimes (C_2)^n$ then Φ or Ψ is a homomorphism by Proposition 1. Without loss of gererality we assume both loops to be left Bol-loops. If the loops have the same number of elements with order 6 then the mapping ι as in the proof of Proposition 2 can be shown to be an isomorphism from L_1 onto L_2: Any two elements $\bar{a} = (a, p)$ and $\bar{b} = (b, q)$ of $C_3 \rtimes (C_2)^n$ generate a subloop of $C_3 \rtimes (C_2)^n$ isomorphic to $C_3 \rtimes (C_2)^2$. Therefore ι is an isomorphism by the proof of Proposition 2. □

Corollary 2 *For $n \geq 3$ the proper Bol-loops of the form $C_3 \rtimes (C_2)^n$ are distributed in $2^n - 1$ classes of isomorphism.*

References

[1] Birkenmeier, G., Davis, B., Reeves, K., Xiao, S.: *Is a Semidirect Product of Groups Necessarily a Group?* Proc. Am. Math. Soc. **118** (1993), 689–692.

[2] Birkenmeier, G., Xiao, S.: *Loops which are Semidirect Products of Groups.* Commun. Algebra **23** (1995), 81–95.

[3] Burn, R. P.: *Finite Bol loops.* Math. Proc. Camb. Philos. Soc. **84** (1978), 377–385.

[4] Burn, R. P.: *Finite Bol loops II.* Math. Proc. Camb. Philos. Soc. **88** (1981), 445–455.

[5] Chein, O., Pflugfelder, O.: *The smallest Moufang loop.* Arch. Math. **22** (1971), 573–576.

[6] Chein, O.: *Moufang Loops of Small Order I.* Trans. Am. Math. Soc. **188** (1974), 31–51.

[7] Chein, O.: *Moufang Loops of Small Order.* Mem. Am. Math. Soc. **197** (1978), 1–131.

[8] Chein, O., Goodaire, E. G.: *A new construction of Bol-Loops of order 8k.* J. Algebra **287** (2005), 103–122.

[9] Figula, Á, Strambach, K.: *Loops which are semidirect product of groups.* Acta Math. Hung., to appear 2007.

[10] Goodaire, E. G., May, S.: Bol Loops of Order less than 32. *Department of Mathematics and Statistics, Memorial University of Newfoundland, Canada*, 1995.

[11] Pflugfelder, H. O.: Quasigroups and Loops: An Introduction. *Heldermann Verlag, Berlin*, 1990.

[12] Robinson, D. A.: *Bol Loops.* Trans. Am. Math. Soc. **123** (1966), 431–354.

ACTA
UNIVERSITATIS PALACKIANAE
OLOMUCENSIS

FACULTAS RERUM NATURALIUM
MATHEMATICA 46 (2007)

Published by the Palacký University Olomouc, Křížkovského 8, 771 47 Olomouc

First edition

ISBN 978-80-244-1819-3
ISSN 0231-9721