TESTING HYPOTHESES ABOUT STRUCTURE OF PARAMETERS IN MODELS WITH BLOCK COMPOUND SYMMETRIC COVARIANCE STRUCTURE

Roman Zmyślony[1], Arkadiusz Kozioł[2]

ABSTRACT

In this article we deal with testing the hypotheses of the so-called structured mean vector and the structure of a covariance matrix. For testing the above mentioned hypotheses Jordan algebra properties are used and tests based on best quadratic unbiased estimators (BQUE) are constructed. For convenience coordinate-free approach (see Kruskal (1968) and Drygas (1970)) is used as a tool for characterization of best unbiased estimators and testing hypotheses. To obtain the test for mean vector, linear function of mean vector with the standard inner product in null hypothesis is changed into equivalent hypothesis about some quadratic function of mean parameters (it is shown that both hypotheses are equivalent and testable). In both tests the idea of the positive and negative part of quadratic estimators is applied to get the test, statistics which have F distribution under the null hypothesis. Finally, power functions of the obtained tests are compared with other known tests like LRT or Roy test. For some set for parameters in the model the presented tests have greater power than the above mentioned tests. In the article we present new results of coordinate-free approach and an overview of existing results for estimation and testing hypotheses about BCS models.

Key words: coordinate-free approach, Jordan algebra, multivariate model, block compound symmetric covariance structure, best unbiased estimators, testing structure of mean vector, testing independence of block variables.

1. Coordinate-free approach and Jordan algebra

1.1. Expectation and covariance operator in finite dimensional space with inner product

Let $\mathcal{K}(\cdot, \cdot)$ be a finite dimensional space with an inner product $(\boldsymbol{a}, \boldsymbol{b})$.

Definition 1. *We say that the vector $\boldsymbol{\eta} \in \mathcal{K}$ is the expectation of a random vector $\boldsymbol{y} \in \mathcal{K}$ if there exists $\boldsymbol{\eta}$ such that for all $\boldsymbol{a} \in \mathcal{K}$ the expectation*

$$\mathrm{E}(\boldsymbol{a}, \boldsymbol{y}) = (\boldsymbol{a}, \boldsymbol{\eta}). \tag{1}$$

[1]Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Szafrana 4a, 65-516 Zielona Góra, Poland. E-mail: r.zmyslony@wmie.uz.zgora.pl. ORCID ID: https://orcid.org/0000-0001-8029-0578.

[2]Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Szafrana 4a, 65-516 Zielona Góra, Poland. E-mail: a.koziol@wmie.uz.zgora.pl. ORCID ID: https://orcid.org/0000-0002-5347-4109.

Lemma 1. *Expectation $\boldsymbol{\eta}$ is uniquely defined and does not depend on the choice of inner product.*

Proof. Suppose that $\boldsymbol{\eta}_1$ different from $\boldsymbol{\eta}_2$ are two expectation vectors, then for all $\boldsymbol{a} \in \mathscr{K}$ we have that $\mathrm{E}(\boldsymbol{a}, \boldsymbol{y}) = (\boldsymbol{\eta}_1, \boldsymbol{a}) = (\boldsymbol{\eta}_2, \boldsymbol{a})$. This is equivalent to $(\boldsymbol{\eta}_1 - \boldsymbol{\eta}_2, \boldsymbol{a}) = 0$ for all $\boldsymbol{a} \in \mathscr{K}$ and this is equivalent to $\boldsymbol{\eta}_1 = \boldsymbol{\eta}_2$.

To prove the second part of this lemma let $[\cdot, \cdot]$ be an arbitrary inner product. Then, from the characterization of all inner products it is implied there exists a self-adjoint positive definite operator $\boldsymbol{A} = \boldsymbol{A}^*$ such that for all $\boldsymbol{a}, \boldsymbol{b} \in \mathscr{K}$ we have $[\boldsymbol{a}, \boldsymbol{b}] = (\boldsymbol{a}, \boldsymbol{A}\boldsymbol{b}) = (\boldsymbol{A}\boldsymbol{a}, \boldsymbol{b})$. From the definition of expectation we have that for all $\boldsymbol{a} \in \mathscr{K}$

$$\mathrm{E}[\boldsymbol{a}, \boldsymbol{y}] = [\boldsymbol{a}, \boldsymbol{\eta}_{[\cdot, \cdot]}]. \tag{2}$$

On the other hand, for all $\boldsymbol{a} \in \mathscr{K}$

$$\mathrm{E}[\boldsymbol{a}, \boldsymbol{y}] = \mathrm{E}(\boldsymbol{a}, \boldsymbol{A}\boldsymbol{y}) = \mathrm{E}(\boldsymbol{A}\boldsymbol{a}, \boldsymbol{y}) = (\boldsymbol{A}\boldsymbol{a}, \boldsymbol{\eta}_{(\cdot, \cdot)}) = (\boldsymbol{a}, \boldsymbol{A}\boldsymbol{\eta}_{(\cdot, \cdot)}) = [\boldsymbol{a}, \boldsymbol{\eta}_{(\cdot, \cdot)}]. \tag{3}$$

From (2) and (3) we have that $\boldsymbol{\eta}_{[\cdot, \cdot]} = \boldsymbol{\eta}_{(\cdot, \cdot)}$. $\qquad \square$

Definition 2. *Operator $\boldsymbol{\Sigma}_{(\cdot, \cdot)}$ is a covariance operator if for all $\boldsymbol{a}, \boldsymbol{b} \in \mathscr{K}$*

$$\mathrm{cov}((\boldsymbol{a}, \boldsymbol{y}), (\boldsymbol{b}, \boldsymbol{y})) = (\boldsymbol{a}, \boldsymbol{\Sigma}_{(\cdot, \cdot)}\boldsymbol{b}). \tag{4}$$

The following lemma shows that the covariance operator depends on the choice of inner product.

Lemma 2. *Operator $\boldsymbol{\Sigma}_{(\cdot, \cdot)}$ is uniquely defined and depends on inner product $(\cdot, \cdot)$, i.e. under $[\cdot, \cdot] = (\cdot, \boldsymbol{A}\cdot)$ operator $\boldsymbol{\Sigma}_{[\cdot, \cdot]} = \boldsymbol{\Sigma}_{(\cdot, \cdot)}\boldsymbol{A}$.*

Proof. The proof of uniqueness of the covariance operator is similar to the proof of uniqueness of expectation. To prove the second part of the lemma note that from the definition we have

$$\mathrm{cov}([\boldsymbol{a}, \boldsymbol{y}], [\boldsymbol{b}, \boldsymbol{y}]) = [\boldsymbol{a}, \boldsymbol{\Sigma}_{[\cdot, \cdot]}\boldsymbol{b}]. \tag{5}$$

On the other hand,

$$\begin{aligned} \mathrm{cov}([\boldsymbol{a}, \boldsymbol{y}], [\boldsymbol{b}, \boldsymbol{y}]) &= \mathrm{cov}((\boldsymbol{A}\boldsymbol{a}, \boldsymbol{y}), (\boldsymbol{A}\boldsymbol{b}, \boldsymbol{y})) = (\boldsymbol{A}\boldsymbol{a}, \boldsymbol{\Sigma}_{(\cdot, \cdot)}\boldsymbol{A}\boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{A}\boldsymbol{\Sigma}_{(\cdot, \cdot)}\boldsymbol{A}\boldsymbol{b}) \quad (6) \\ &= [\boldsymbol{a}, \boldsymbol{\Sigma}_{(\cdot, \cdot)}\boldsymbol{A}\boldsymbol{b}] \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (7) \end{aligned}$$

From (5) and (7) it follows that $\boldsymbol{\Sigma}_{[\cdot, \cdot]} = \boldsymbol{\Sigma}_{(\cdot, \cdot)}\boldsymbol{A}$. $\qquad \square$

Remark 1. *Through the paper we deal with $\mathbb{R}^n$ and the standard inner product. In the space of $m \times n$ matrices, which is denoted by $\mathscr{M}^{m,n}$, the inner product is defined as $\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}')$. The space of $n \times n$ symmetric matrices will be denoted by $\mathscr{S}^n$. Because of symmetry the inner product in $\mathscr{S}^n$ is $\mathrm{tr}(\boldsymbol{A}\boldsymbol{B})$. Moreover, throughout the paper $\boldsymbol{A}'$ will stand for transpose of matrix $\boldsymbol{A}$.*

1.2. Special linear operator on space of $\mathcal{M}^{m,n}$

Definition 3. *Let A, B, C be matrices with such dimensions that multiplication ACB is possible. Then:*

$$(A \odot B)C = ACB.$$

In the following remark we will show the relation between the Kronecker product of two matrices A and B ($A \otimes B$), which have orders $k \times l$ and $p \times q$, respectively, and the special operator $\odot$. In this paper, the Kronecker product is defined as block matrix $A \otimes B = a_{ij}B$ for $i = 1, \ldots, k$ and $j = 1, \ldots, l$.

The operator vec is a linear transformation which converts a matrix into a column vector by stacking the columns of the matrix under another. The inverse operator to vec is vec_p^{-1} which converts a column vector into a matrix with p rows, such that $\text{vec}_p^{-1}(\text{vec}(X)) = X$ for all matrices X of order $p \times k$ and $\text{vec}(\text{vec}_p^{-1}(x)) = x$ for all vectors x with dimension $pk \times 1$.

Remark 2. *Let Y be a matrix order $q \times l$. The operator $\odot$ has a following properties:*

- $(A \otimes B)\text{vec}(Y) = \text{vec}\left((B \odot A')Y\right)$;

- $\text{vec}_p^{-1}\left((A \otimes B)\text{vec}(Y)\right) = (B \odot A')Y$;

- $(A \odot B)(C \odot D) = AC \odot DB$.

1.3. Jordan algebra and its properties

An associative algebra can be transformed into a Jordan algebra by the Jordan product $A \circ B = \frac{AB + BA}{2}$ (see Schafer (1966)). Through the paper we deal with Jordan algebras of matrices „formally real" in the sense that if $A^2 + B^2 + \ldots = 0$ then $A = B = \ldots = 0$ (see Jordan, Neumann and Wigner (1934)).

A full characterization of irreducible Jordan algebras of matrices is given by Jordan, Neumann and Wigner (1934) (for more details see also Massam (1994), Massam and Neher (1997), Letac and Massam (1998), Massam and Neher (1998), Faraut and Korányi (1994)):

- The algebra $\mathcal{S}^n$ of all $n \times n$ ($n \geq 1$) symmetric matrices with trace inner product and operation $A \circ B$;

- The algebra $\mathcal{L}^n$ (Lorentz spin algebra);

- The algebra $\mathcal{H}_n$ of all $n \times n$ complex Hermitian matrices with trace inner product and operation $A \circ B$;

- The algebra $\mathcal{Q}_n$ of all $n \times n$ quaternion Hermitian matrices with trace inner product operation $A \circ B$;

- The algebra $\mathcal{O}_3$ of all 3×3 octonion Hermitian matrices with trace inner product and operation $A \circ B$.

Remark 3. *Note that all Jordan algebras can be represented as Cartesian product of all above Jordan algebras. For statistics, the most important are Cartesian product of $\mathbb{R}$ (as a special case of the first one with $n = 1$, where multiplication is commutative $(ab = ba)$) and $\mathcal{S}^n$ (for $n \geq 2$). They were named as quadratic subspaces by Seely (1971). If matrices in Jordan algebra commute, i.e. $\boldsymbol{AB} = \boldsymbol{BA}$, then this algebra is isomorphic to Cartesian product of $\mathbb{R}$.*

Some of properties which we will use through the paper are given in the lemma below.

Lemma 3. *Let $\boldsymbol{\vartheta}$ be a quadratic subspace of $\mathcal{S}^n$. Then:*

1. *$\boldsymbol{A} \in \boldsymbol{\vartheta} \Rightarrow \boldsymbol{A}^k \in \boldsymbol{\vartheta}$;*

2. *$\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{\vartheta} \Rightarrow \boldsymbol{ABA} \in \boldsymbol{\vartheta}$;*

3. *$\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \boldsymbol{\vartheta} \Rightarrow \boldsymbol{ABC} + \boldsymbol{CBA} \in \boldsymbol{\vartheta}$;*

4. *$\boldsymbol{P}^2 = \boldsymbol{P}$, $\boldsymbol{P} = \boldsymbol{P}'$ and $\forall \boldsymbol{V} \in \boldsymbol{\vartheta} \; \boldsymbol{PV} = \boldsymbol{VP} \Rightarrow \boldsymbol{M}\boldsymbol{\vartheta}\boldsymbol{M}' = \boldsymbol{M}\boldsymbol{\vartheta}$ is a quadratic subspace, where matrix $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{P}$, while $\boldsymbol{I}$ stands for identity matrix;*

5. *If $\boldsymbol{Q}$ is an orthogonal matrix then $\boldsymbol{Q}\boldsymbol{\vartheta}\boldsymbol{Q}'$ is also a quadratic subspace.*

For the proof see Seely (1971) and also Zmyślony (1979).

2. Estimation and testing hypotheses in mixed models for univariate case

In this section we deal with estimation and testing hypotheses using coordinate-free approach and properties of Jordan algebra.

2.1. Estimation of parameters in mixed models

The well-known normal mixed model can be expressed as follows

$$\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{X\beta}, \boldsymbol{V}(\boldsymbol{\sigma})), \tag{8}$$

where $\boldsymbol{\sigma} = \left(\sigma_1^2, \ldots, \sigma_m^2\right)'$ and $\boldsymbol{V}(\boldsymbol{\sigma}) = \sum_{i=1}^m \sigma_i^2 \boldsymbol{V}_i$. We shall note that $\mathcal{K} = \mathbb{R}^n$ with the standard inner product, $\mathcal{X} = \{\boldsymbol{X\beta} : \boldsymbol{\beta} \in \mathbb{R}^p\}$ and $\boldsymbol{\vartheta} = \mathrm{sp}\left\{\sum_{i=1}^m \sigma_i^2 \boldsymbol{V}_i : \sigma_i \geq 0, \boldsymbol{V}_i \text{ are known}\right\}$.

Remark 4. *We assume that there exists $\boldsymbol{\sigma}_0$ such that $\boldsymbol{V}(\boldsymbol{\sigma}_0) = \boldsymbol{I}$.*

Let $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \boldsymbol{\sigma}\}$ and $g(\boldsymbol{\theta})$ be a real-valued function. We consider the following classes of linear and quadratic estimators, respectively
$\mathcal{A} = \{(\boldsymbol{a}, \boldsymbol{y}) : \boldsymbol{a} \in \mathbb{R}^n, \mathrm{E}(\boldsymbol{a}, \boldsymbol{y}) = g(\boldsymbol{\theta})\}$,
$\mathcal{B} = \left\{\langle \boldsymbol{B}, \boldsymbol{yy}'\rangle : \boldsymbol{B} \in \mathcal{S}^n, \mathrm{E}\langle \boldsymbol{B}, \boldsymbol{yy}'\rangle = g(\boldsymbol{\theta})\right\}$.

Remark 5. *In the class $\mathscr{B}$ with $\mathscr{K} = \mathscr{S}^n$ and the inner product $\langle A, B \rangle = \mathrm{tr}(AB)$ we get that the expectation of yy' is:*

$$\mathrm{E}(yy') = V(\sigma) + X\beta\beta'X = V(\sigma) + \mathrm{E}(y)\mathrm{E}(y)'$$

and covariance of yy' is:

$$\mathrm{cov}(yy') = 2\left[\mathrm{E}(yy') \otimes \mathrm{E}(yy') - \mathrm{E}(y)\mathrm{E}(y)' \otimes \mathrm{E}(y)\mathrm{E}(y)'\right].$$

We recall the definition of estimable function of parameters β and σ.

Definition 4. *A function $[c, \beta]$ is said to be estimable if there exists a linear unbiased estimator for this function i.e. $\mathrm{E}(a, y) = [c, \beta]$.*

In the following two theorems conditions for the existence of estimators with optimal properties in a mixed linear model are given.

Theorem 1. *For any estimable function $[c, \beta]$ there exists its best linear unbiased estimator if and only if for $i = 1, \ldots, m$ holds $PV_i = V_iP$, where $P = XX^+$, while X^+ is Moore-Penrose inverse of matrix X.*

Theorem 2. *For any estimable function $[c, \sigma]$ there exists its best quadratic unbiased estimator if and only if $M\vartheta M$ is quadratic subspace, where $M = I - XX^+$.*

Theorem 3. *For any quadratic estimable function there exists best quadratic unbiased estimators (BQUE) if and only if $\mathrm{sp}\{X\beta\beta'X, V_1, \ldots, V_m\}$ is quadratic subspace.*

For proofs of these theorems see Zmyślony (1978, 1980). From Seely (1972, 1977) and Zmyślony (1980), and since the estimators are functions of complete sufficient statistics, the following remarks follows.

Remark 6. *Best linear unbiased estimators and best quadratic unbiased estimators are best unbiased estimators.*

Suppose that

$$\underset{n \times 1}{y} \sim \mathcal{N}(\mu 1, V(\sigma)),$$

where $\mu \in \mathbb{R}$ and $V(\sigma) = \sigma_1^2 V_1 + \sigma_2^2 V_2 + \sigma_3^2 V_3$, while

$$V_1 = \begin{bmatrix} \underset{n_1 \times n_1}{11'} & 0 \\ 0 & \underset{n_2 \times n_2}{0} \end{bmatrix}, V_2 = \begin{bmatrix} \underset{n_1 \times n_1}{0} & 0 \\ 0 & \underset{n_2 \times n_2}{11'} \end{bmatrix}, V_3 = \begin{bmatrix} \underset{n_1 \times n_1}{I} & 0 \\ 0 & \underset{n_2 \times n_2}{I} \end{bmatrix}.$$

Since the expectation of yy' is

$$\mathrm{E}(yy') = \mu^2 11' + V(\sigma) = \mu^2 11' + \sigma_1^2 V_1 + \sigma_2^2 V_2 + \sigma_3^2 V_3,$$

three following conditions for this model are satisfied:

1. $\vartheta = \mathrm{sp}\left\{\mathbf{1}\mathbf{1}', \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\right\}$ is a quadratic subspace,

2. $\frac{1}{n}\mathbf{1}\mathbf{1}'$ does not commute with $\mathbf{V}_1$ and $\mathbf{V}_2$,

3. according to first characterization of Jordan algebra it means that ϑ can be represented as Cartesian product of 2×2 symmetric matrices and $\sigma_3^2 \mathbf{I}$.

Remark 7. *According to Theorem 1 note that $\mathbf{P} = \frac{1}{n}\mathbf{1}\mathbf{1}'$ does not commute with $\mathbf{V}(\boldsymbol{\sigma})$ and thus the BLUE of μ does not exist. However, from Theorem 3, there exists the BQUE for μ^2.*

2.2. Tests for variance components based on unbiased estimators

For the normal model of the form given in (8) we consider the following hypotheses

$$H_0 : \sigma_i^2 = 0 \;\; vs \;\; H_1 : \sigma_i^2 > 0.$$

Let $\mathbf{y}'\mathbf{A}\mathbf{y}$ be an unbiased estimator of σ_i^2. Moreover, let $\mathbf{A}_+$, $\mathbf{A}_-$ stand for positive and negative part of matrix $\mathbf{A}$, respectively.

Remark 8. *For $i < k$ the estimator $\mathbf{y}'\mathbf{A}\mathbf{y}$ is "not defined", that is $\mathbf{A} = \mathbf{A}_+ - \mathbf{A}_-$, where $\mathbf{A}_+, \mathbf{A}_- \geq 0$, i.e. $\mathbf{A}_+, \mathbf{A}_-$ are nonnegative definite matrices different than $\mathbf{0}$. Note that*

- *if H_0 is true, then $\mathrm{E}\left(\mathbf{y}'\mathbf{A}_+\mathbf{y}\right) = \mathrm{E}\left(\mathbf{y}'\mathbf{A}_-\mathbf{y}\right)$,*

- *if H_1 is true, then $\mathrm{E}\left(\mathbf{y}'\mathbf{A}_+\mathbf{y}\right) > \mathrm{E}\left(\mathbf{y}'\mathbf{A}_-\mathbf{y}\right)$.*

Corollary 1. *The test should reject hypothesis*

$$H_0 : \sigma_i^2 = 0$$

if statistic

$$F = \frac{\mathbf{y}'\mathbf{A}_+\mathbf{y}}{\mathbf{y}'\mathbf{A}_-\mathbf{y}}$$

is sufficiently large.

Let us consider three conditions for commutative Jordan algebra, i.e. for all elements $\mathbf{A}$ and $\mathbf{B}$ of such algebra $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$:

1. $\mathrm{sp}\left\{\mathbf{M}\mathbf{V}_1\mathbf{M}, \ldots, \mathbf{M}\mathbf{V}_k\mathbf{M}\right\}$ is a commutative Jordan algebra,

2. $\mathrm{sp}\left\{\left\{\mathbf{M}\mathbf{V}_1\mathbf{M}, \ldots, \mathbf{M}\mathbf{V}_k\mathbf{M}\right\} \setminus \left\{\mathbf{M}\mathbf{V}_i\mathbf{M}\right\}\right\}$ is a commutative Jordan algebra,

3. $F = \frac{\mathbf{y}'\mathbf{A}_+\mathbf{y}}{\mathbf{y}'\mathbf{A}_-\mathbf{y}}$ has F-Snedecor distribution under $H_0 : \sigma_i^2 = 0$.

Theorem 4. *The first and second from the above conditions imply the third condition.*

For proof see Michalski and Zmyślony (1996).

Theorem 5. *The first and third from the above conditions imply the second condition.*

Theorem 6. *Let us assume that a subspace*

$$\mathrm{sp}\left\{ \boldsymbol{MV}_1\boldsymbol{M}, \ldots, \boldsymbol{MV}_k\boldsymbol{M} \right\}$$

is a commutative Jordan algebra, while

$$\mathrm{sp}\left\{ \left\{ \boldsymbol{MV}_1\boldsymbol{M}, \ldots, \boldsymbol{MV}_k\boldsymbol{M} \right\} \setminus \left\{ \boldsymbol{MV}_i\boldsymbol{M} \right\} \right\}$$

is not a commutative Jordan algebra. Then, statistic

$$F = \frac{\boldsymbol{y}'\boldsymbol{A}_+\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{A}_-\boldsymbol{y}}.$$

has a generalized F-Snedecor distribution under $H_0 : \sigma_i^2 = 0$*, where* $\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}$ *is BQUE of parameter* σ_i^2 *(see Fonseca et al. (2002)).*

3. Block compound symmetric covariance structure in doubly multivariate data

3.1. Covariance structure

The $(mu \times mu)$–dimensional BCS covariance structure for m-variate observations over u factor levels is defined as:

$$
\begin{aligned}
\boldsymbol{\Gamma} &= \begin{bmatrix} \boldsymbol{\Gamma}_0 & \boldsymbol{\Gamma}_1 & \cdots & \boldsymbol{\Gamma}_1 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \boldsymbol{\Gamma}_1 & \boldsymbol{\Gamma}_1 & \cdots & \boldsymbol{\Gamma}_0 \end{bmatrix} \\
&= \left(\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1 \right) \odot \boldsymbol{I}_u + \boldsymbol{\Gamma}_1 \odot \boldsymbol{J}_u \\
&= \boldsymbol{\Gamma}_0 \odot \boldsymbol{I}_u + \boldsymbol{\Gamma}_1 \odot \left(\boldsymbol{J}_u - \boldsymbol{I}_u \right)
\end{aligned}
$$

with $\boldsymbol{J}_u = \boldsymbol{1}_u\boldsymbol{1}_u'$. The above BCS structure can be also written as a sum of two orthogonal matrices (i.e. the product of orthogonal matrices is equal to matrix $\boldsymbol{0}$):

$$\boldsymbol{\Gamma} = (\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1) \odot \left(\boldsymbol{I}_u - \frac{1}{u}\boldsymbol{J}_u \right) + \left(\boldsymbol{\Gamma}_0 + (u-1)\boldsymbol{\Gamma}_1 \right) \odot \frac{1}{u}\boldsymbol{J}_u.$$

The following assumptions for matrices $\boldsymbol{\Gamma}_0$ and $\boldsymbol{\Gamma}_1$ in BCS structure

1. $\boldsymbol{\Gamma}_0$ is a positive definite symmetric $m \times m$ matrix,

2. $\boldsymbol{\Gamma}_1$ is a symmetric $m \times m$ matrix,

3. $\boldsymbol{\Gamma}_0 + (u-1)\boldsymbol{\Gamma}_1$ is a positive definite matrix,

4. $\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1$ is a positive definite matrix.

imply that the $um \times um$ matrix $\boldsymbol{\Gamma}$ is positive definite (for the proof see Lemma 2.1 in Roy and Leiva (2011)). This result follows also from the property of rank for strong orthogonality of matrices.

3.2. Normal model with BCS covariance structure

The normal BCS model can be written in the following way:

$$\underset{um \times n}{\boldsymbol{Y}} = \left[\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_n\right] \sim \mathcal{N}\left((\boldsymbol{I}_{um} \odot \boldsymbol{1}'_n)\boldsymbol{\mu}, \boldsymbol{\Gamma} \odot \boldsymbol{I}_n\right) \tag{9}$$

with $\boldsymbol{\Gamma}$ defined in 3.1. In this model we assume that the mean vector changes over sites or over time points so $\boldsymbol{\mu}$ has um components. Matrix $\boldsymbol{Y}$ contains n independent normally distributed random column vectors, which are identically distributed with the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Gamma}$.

Let us consider orthogonal transformation $\boldsymbol{I}_{um} \odot \boldsymbol{Q}$ on $\boldsymbol{Y}_{um \times n}$, where $\boldsymbol{Q}$ is an orthogonal matrix of order n.

Proposition 1. *If* $\operatorname{cov}(\boldsymbol{Y}) = \boldsymbol{\Sigma}_{\boldsymbol{Y}} = \boldsymbol{\Sigma} \odot \boldsymbol{I}$ *with any covariance matrix* $\boldsymbol{\Sigma}$ *then covariance is invariant with respect to transformation* $\boldsymbol{I} \odot \boldsymbol{Q}$ *on* $\boldsymbol{Y}$.

In the next proposition we show that orthogonal transformation saves commutativity of projectors with covariance matrices as well as the property of quadratic subspace.

Proposition 2. *Let* $\boldsymbol{\vartheta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}$ *be the space generated by covariance matrices* $\boldsymbol{\Sigma}$ *and let* $\boldsymbol{P}_{\mathrm{E}(\boldsymbol{Y})}$ *denote orthogonal projector onto the subspace of mean matrix of a random matrix* $\boldsymbol{Y}$. *Moreover, let* $\boldsymbol{U} = \boldsymbol{Q}(\boldsymbol{Y})$, *where* $\boldsymbol{Q}$ *is an arbitrary orthogonal operator. Then:*

(i) *If* $\boldsymbol{P}_{\mathrm{E}(\boldsymbol{Y})}\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \boldsymbol{\Sigma}_{\boldsymbol{Y}}\boldsymbol{P}_{\mathrm{E}(\boldsymbol{Y})}$ *then* $\boldsymbol{P}_{\mathrm{E}(\boldsymbol{U})}\boldsymbol{\Sigma}_{\boldsymbol{U}} = \boldsymbol{\Sigma}_{\boldsymbol{U}}\boldsymbol{P}_{\mathrm{E}(\boldsymbol{U})}$. (10)

(ii) *If* $\boldsymbol{\vartheta}_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}$ *is a quadratic subspace then* $\boldsymbol{\vartheta}_{\boldsymbol{\Sigma}_{\boldsymbol{U}}}$ *is a quadratic subspace.* (11)

For the special case of $\boldsymbol{Q} = \boldsymbol{Q}_1 \odot \boldsymbol{Q}_2$ we get the following:

Lemma 4. *Since the space* $\boldsymbol{\vartheta}_{\operatorname{cov}(\boldsymbol{Y})}$ *generated by covariance matrices* $\boldsymbol{\Gamma} \odot \boldsymbol{I}$ *is a quadratic subspace and orthogonal projector* $\boldsymbol{P}_{\mathrm{E}(\boldsymbol{Y})} = \boldsymbol{I}_{um} \odot \frac{1}{n}\boldsymbol{J}_n$ *commutes with covariance matrices, we have:*

$\boldsymbol{P}_{\mathrm{E}(\boldsymbol{U})}$ *commutes with* $\operatorname{cov}(\boldsymbol{U})$ *and* $\boldsymbol{\vartheta}_{\operatorname{cov}(\boldsymbol{U})}$ *is a quadratic subspace.*

For the proof that for the model *(9)* $\boldsymbol{\vartheta}_{\operatorname{cov}(\boldsymbol{Y})}$ is a quadratic subspace and assumption that commutativity of $\boldsymbol{P}_{\mathrm{E}(\boldsymbol{Y})}$ holds see Roy et al. (2016).

3.3. Testing hypotheses about structure of expectation

In this section we consider testing hypotheses about the parameters of the mean vector. These results can be also found in Zmyślony et al. (2018). For this reason

we use the two following orthogonal transformations:

1. $U = [u_1, \ldots, u_n] = (I_{um} \odot Q_2)Y$, where $Q_2 = \left[\frac{1}{\sqrt{n}} 1_n \vdots K_{1_n} \right]$ is Helmert matrix, such that $K'_{1_n} K_{1_n} = I_{n-1}$ and $K'_{1_n} 1_n = 0$,

2. $W_i = (I \odot Q_1)U_i$, where $U_i = \text{vec}^{-1}(u_i)$ is a matrix of size $m \times u$ and $Q_1 = \left[\frac{1}{\sqrt{u}} 1_u \vdots K_{1_u} \right]$

which are useful for constructing the test statistic.

Now, we formulate the null hypothesis for structure of mean

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_u.$$

This hypothesis can be written equivalently as

$$H_0 : \mu_2^{(c)} = \mu_3^{(c)} = \ldots = \mu_u^{(c)} = 0,$$

where $\mu_j^{(c)} = \sqrt{nu} \sum_{l=1}^{u} k_{l,j-1} \mu_l$, while $k_{l,j-1}$ is $l, j-1$-th element of K_{1_u}.

Following the idea of Michalski and Zmyślony (1999) this hypothesis is equivalent to

$$H_0 : \sum_{j=2}^{u} \mu_j^{(c)} \mu_j^{(c)'} = 0.$$

One can prove that quadratic estimator of $\sum_{j=2}^{u} \mu_j^{(c)} \mu_j^{(c)'}$ is a function of complete sufficient statistics (see Roy et al. (2016)) and has the following form:

$$\widehat{\sum_{j=2}^{u} \mu_j^{(c)} \mu_j^{(c)'}} = \sum_{j=2}^{u} \widehat{\mu}_j^{(c)} \widehat{\mu}_j^{(c)'} - (u-1)\widehat{\Gamma}_0 - \widehat{\Gamma}_1, \tag{12}$$

where $\widehat{\Gamma}_0$ and $\widehat{\Gamma}_1$ are best unbiased estimators (BUE) for Γ_0 and Γ_1, respectively. For details see Roy et al. (2016).

Note that

$$\sum_{j=2}^{u} \widehat{\mu}_j^{(c)} \widehat{\mu}_j^{(c)'} \stackrel{\text{df}}{=} (u-1)\widehat{\Delta}_2$$

is the positive part and

$$(u-1)\widehat{\Gamma}_0 - \widehat{\Gamma}_1 \stackrel{\text{df}}{=} (u-1)\widehat{\Delta}_1$$

is the negative part of estimator in (12).

Under the null hypothesis the positive part has Wishart distribution and the negative part multiplied by $(n-1)$ is Wishart distributed with the same covariance matrix

$\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1$

$$(n-1)(u-1)\widehat{\mathbf{\Delta}}_1 \sim \mathcal{W}_m\left(\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1, (n-1)(u-1)\right),$$
$$(u-1)\widehat{\mathbf{\Delta}}_2 \sim \mathcal{W}_m\left(\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1, u-1\right),$$

where $\widehat{\mathbf{\Delta}}_1$ and $\widehat{\mathbf{\Delta}}_2$ are independent.

Lemma 5. *If $\boldsymbol{W}_1 \sim \mathcal{W}_m(\mathbf{\Sigma}, n_1)$ and $\boldsymbol{W}_2 \sim \mathcal{W}_m(\mathbf{\Sigma}, n_2)$ are independent, then for every fixed vector $\boldsymbol{x} \neq 0 \in \mathbb{R}^m$:*

$$F = \frac{n_2 \boldsymbol{x}' \boldsymbol{W}_1 \boldsymbol{x}}{n_1 \boldsymbol{x}' \boldsymbol{W}_2 \boldsymbol{x}} \sim F_{n_1, n_2}.$$

Now, we give the theorem from Zmyślony et al. (2018).

Theorem 7. *Under the null hypothesis, the statistic*

$$F = \frac{\boldsymbol{x}' \sum_{j=2}^{u} \widehat{\boldsymbol{\mu}}_j^{(c)} \widehat{\boldsymbol{\mu}}_j^{(c)'} \boldsymbol{x}}{(u-1)\boldsymbol{x}'\left(\widehat{\mathbf{\Gamma}}_0 - \widehat{\mathbf{\Gamma}}_1\right)\boldsymbol{x}} = \frac{\boldsymbol{x}'\widehat{\mathbf{\Delta}}_2 \boldsymbol{x}}{\boldsymbol{x}'\widehat{\mathbf{\Delta}}_1 \boldsymbol{x}} \tag{13}$$

has F distribution with $(u-1)$ and $(n-1)(u-1)$ degrees of freedom for any fixed $\boldsymbol{x}$.

3.4. Testing hypotheses about $\mathbf{\Gamma}_1$

In this section we consider the following hypotheses about parameters in matrix $\mathbf{\Gamma}_1$ under assumption that all elements of $\mathbf{\Gamma}_1$ are nonnegative or nonpositive:

$$H_0 : \mathbf{\Gamma}_1 = \mathbf{0} \ \ vs. \ \ H_1 : \mathbf{\Gamma}_1 \neq \mathbf{0}.$$

The presented results can be also found in Fonseca et al. (2018). From Roy et al. (2015) we get that matrices:

$$(n-1)(u-1)\widehat{\mathbf{\Delta}}_1 = (n-1)(u-1)(\widehat{\mathbf{\Gamma}}_0 - \widehat{\mathbf{\Gamma}}_1) \sim \mathcal{W}_m(\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1, (n-1)(u-1)),$$

$$(n-1)\widehat{\mathbf{\Delta}}_2 = (n-1)(\widehat{\mathbf{\Gamma}}_0 + (u-1)\widehat{\mathbf{\Gamma}}_1) \sim \mathcal{W}_m(\mathbf{\Gamma}_0 + (u-1)\mathbf{\Gamma}_1, (n-1))$$

are independent. It is easy to show that:

$$\widehat{\mathbf{\Gamma}}_1 = \frac{\widehat{\mathbf{\Delta}}_2 - \widehat{\mathbf{\Delta}}_1}{u}.$$

Under the framework given in Michalski and Zmyślony (1996) a positive part of $\widehat{\mathbf{\Gamma}}_1$ is given by:

$$\widehat{\mathbf{\Gamma}}_{1+} = \frac{\widehat{\mathbf{\Delta}}_2}{u}$$

and a negative part is given by:

$$\widehat{\mathbf{\Gamma}}_{1-} = \frac{\widehat{\mathbf{\Delta}}_1}{u}.$$

Note that estimator of $\mathbf{\Gamma}_1$ is given by:

$$\widehat{\mathbf{\Gamma}}_1 = \widehat{\mathbf{\Gamma}}_{1+} - \widehat{\mathbf{\Gamma}}_{1-} = \frac{\widehat{\mathbf{\Delta}}_2 - \widehat{\mathbf{\Delta}}_1}{u}.$$

For the proof of the following theorem see Fonseca et al. (2018).

Theorem 8. *Under the hypothesis $H_0 : \mathbf{\Gamma}_1 = 0$ the test statistic:*

$$F = \frac{\mathbf{1}' \widehat{\mathbf{\Gamma}}_{1+} \mathbf{1}}{\mathbf{1}' \widehat{\mathbf{\Gamma}}_{1-} \mathbf{1}} \tag{14}$$

has F distribution with $(n-1)$ and $(n-1)(u-1)$ degrees of freedom.

3.5. Testing hypotheses about single parameters

$$H_0 : \sigma_{ij}^{(1)} = 0 \quad vs. \quad H_1 : \sigma_{ij}^{(1)} \neq 0$$

In order to conduct F test for testing hypotheses about single parameter, i.e. $H_0 : \sigma_{ii}^{(1)} = 0$ for given $i = 1, \ldots, m$, vectors $\mathbf{1}$ in (14) should be replaced by

$$\boldsymbol{e}_i = (0, \ldots, 0, \underbrace{1}_{i\text{th position}}, 0, \ldots, 0)'.$$

If $\sigma_{ii}^{(1)}$ and $\sigma_{jj}^{(1)}$ are equal to zeros then for parameters $\sigma_{ij}^{(1)}$, $i < j$, $i = 1, \ldots, m$, instead of vectors $\mathbf{1}$ in (14) one should insert

$$\boldsymbol{e}_i - \boldsymbol{e}_j = (0, \ldots, 0, \underbrace{1}_{i\text{th position}}, 0, \ldots, \underbrace{-1}_{j\text{th position}}, 0, \ldots, 0)'.$$

Remark 9. *Testing single contrast of parameters can be done in a similar way using vector e_i defined above instead of $\mathbf{1}_u$.*

4. Data application

In this section we use a data set from Johnson and Wichern (2007) for estimation parameters and testing hypotheses, presented in previous section, about the structure of expectation and covariance parameters in model (9). These data contain measures of mineral content of three bones for 25 women: radius, humerus and ulna. Each measurement was recorded on the dominant and non-dominant side.

Using the formula (4.13) and Theorem 1 from Roy et al. (2016) we get that BLUE for $\boldsymbol{\mu}$ is

$$\widehat{\boldsymbol{\mu}} = \begin{bmatrix} 0.84380 & 1.79268 & 0.70440 & 0.81832 & 1.73484 & 0.69384 \end{bmatrix},$$

where, in accordance with the order of variables, the first three values are the means for measurements of mineral content in dominant side of radius, humerus and ulna, respectively, while the last three values are the means for measurements of mineral content in non-dominant side for these bones.

From the same paper, using formulas (3.4) and (3.5) and Theorem 1 we get that BQUE for $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_1$ are

$$\widehat{\mathbf{\Gamma}}_0 = \begin{bmatrix} 0.01221 & 0.02172 & 0.00901 \\ 0.02172 & 0.07492 & 0.01682 \\ 0.00901 & 0.01682 & 0.01108 \end{bmatrix} \text{ and } \widehat{\mathbf{\Gamma}}_1 = \begin{bmatrix} 0.01038 & 0.01931 & 0.00824 \\ 0.01931 & 0.06678 & 0.01529 \\ 0.00824 & 0.01529 & 0.00807 \end{bmatrix},$$

respectively.

For testing hypotheses about the structure of expectation in test statistic (13) we take the vector $\mathbf{x} = \mathbf{1}_m$, so we consider the sum of elements of the positive and negative part of the estimator $\sum_{j=2}^{u} \widehat{\boldsymbol{\mu}_j^{(c)} \boldsymbol{\mu}_j^{(c)\prime}}$. Our test was compared with two well-known tests: likelihood ratio test (LRT) and Roy's test. Formulas of these tests statistics were given in Zmyślony et al. (2018) in Section 4. Calculated p-values for considered data example for all three tests are given in the table below.

Table 1: P-values in testing hypotheses about the structure of expectation and elements of $\mathbf{\Gamma}_1$ with the use of three different tests

Name of test	Test for $\boldsymbol{\mu}$	Test for $\mathbf{\Gamma}_1$
F test	0.0363	$1.06073 \cdot 10^{-9}$
LRT	0.1725	$1.807443 \cdot 10^{-13}$
Roy's test	0.1725	——

The same p-values for LRT and Roy's test in Table 1 follow from the fact that in case $u = 2$ both tests are equivalent. On the standard 5% level of significance we conclude on the p-value for F test that means are significantly different between two sides. For more details about the comparison of these three tests see Zmyślony et al. (2018).

Test F for testing hypotheses about elements of $\mathbf{\Gamma}_1$, whose statistic was given in (14), was compared with LRT, whose statistic was given in formula (3.3) in Fonseca et al. (2018). For both tests, on 5% level of significance, we can conclude that at least one element of $\mathbf{\Gamma}_1$ is different than 0.

5. Conclusion

This paper contains a review of results concerning estimation and testing hypotheses for univariate and multivariate linear models. The presented results are based on the properties of Jordan algebra. Moreover, the coordinate-free approach simplifies inference in linear models for both the univariate and multivariate case. It was presented how the methods of estimation and testing for a univariate model can be

extended to the multivariate case. Estimators of the parameters of the presented BCS covariance structure model and the data presenting measures of mineral content of bones can be found in Roy et al. (2016). The power of the proposed tests for expectation and covariance parameters, in the multivariate case, is compared with well-known tests such as LRT and Roy's test in Fonseca et al. (2018) and Zmyślony et al. (2018). As a result of the simulation study we can say that in some cases (for some alternatives) the tests proposed in this paper have greater power than LRT and Roy's test. The same data example, as for the estimation purpose, was used for testing hypotheses for covariance structure in Fonseca et al. (2018) and for testing hypotheses about the mean structure in Zmyślony et al. (2018).

REFERENCES

DRYGAS, H., (1970). The Coordinate-Free Approach to Gauss-Markov Estimation, Berlin, Heidelberg: Springer.

FARAUT, J., KORÁNYI, A., (1994). Analysis on symmetric cones, Oxford University Press, Oxford.

FONSECA, M., MEXIA, J. T., ZMYŚLONY, R., (2002). Exact distribution for the generalized F tests. Discussiones Mathematicae Probability and Statistics, 22, pp. 37–51.

FONSECA, M., KOZIOŁ, A., ZMYŚLONY, R., (2018). Testing hypotheses of covariance structure in multivariate data. Electronic Journal of Linear Algebra: International Conference on Matrix Analysis and its Applications, MAT TRIAD 2017, 33, pp. 53–62.

JOHNSON, R. A., WICHERN, D. W., (2007). Applied Multivariate Statistical Analysis, New Jersey, Pearson Prentice Hall, sixth ed., Englewood Cliffs.

JORDAN, P., NEUMANN, von, J., WIGNER, E., (1934). On an algebraic generalization of the quantum mechanical formalism. The Annals of Mathematics, 35 (1), pp. 29–64.

KOZIOŁ, A., ROY, A., FONSECA, M., ZMYŚLONY, R., LEIVA, R., (2018). Free-coordinate estimation for doubly multivariate data. Linear Algebra Appl., 547C, pp. 217–239.

KRUSKAL, W., (1968). When are Gauss-Markov and Least Squares Estimators Identical? A Coordinate-Free Approach. The Annals of Mathematical Statistics, 39 (1), pp. 70–75.

LETAC, G., MASSAM, H., (1998). Quadratic and inverse regressions for Wishart distributions, The Annals of Statistics, 26 (2), pp. 573–595.

MASSAM, H., (1994). An exact decomposition theorem and a unified view of some related distributions for a class of exponential transformations on symmetric cones, The Annals of Statistics, 22 (1), pp. 369–394.

MASSAM, H., NEHER, E., (1997). On transformations and determinants of Wishart variables on symmetric cones, Journal of Theoretical Probability, 10, pp. 867–902.

MASSAM, H., NEHER, E., (1998). Estimation and testing for lattice conditional independence models on Euclidean Jordan algebras, The Annals of Statistics, 26 (3), pp. 1051–1082.

MICHALSKI, A., ZMYŚLONY, R., (1996). Testing hypothesis for variance components in mixed linear models. Statistics, 27, pp. 297–310.

MICHALSKI, A., ZMYŚLONY, R., (1999). Testing hypotheses for linear functions of parameters in mixed linear models. Tatra Mountains Mathematical Publications, 17, pp. 103–110.

ROY, A., LEIVA, R., ŽEŽULA, I., KLEIN, D., (2015). Testing the equality of mean vectors for paired doubly multivariate observations in blocked compound symmetric covariance matrix setup. Journal of Multivariate Analysis, 137, pp. 50–60.

ROY, A., ZMYŚLONY, R., FONSECA, M., LEIVA, R., (2016). Optimal estimation for doubly multivariate data in blocked compound symmetric covariance structure, Journal of Multivariate Analysis, 144, pp. 81–90.

SCHAFER, R., (1966). An Introduction to Nonassociative Algebras, Academic Press New York and London.

SEELY, J. F., (1971). Quadratic subspaces and completeness. The Annals of Mathematical Statistics, 42 (2), pp. 710–721.

SEELY, J. F., (1972). Completeness for a family of multivariate normal distributions. The Annals of Mathematical Statistics, 43, pp. 1644–1647.

SEELY, J. F., (1977). Minimal sufficient statistics and completeness for multivariate normal families. Sankhya (Statistics). The Indian Journal of Statistics, 39(2), pp. 170–185.

ZMYŚLONY, R., (1978). A characterization of best linear unbiased estimators in the general linear model, Lecture Notes in Statistics, 2, pp. 365–373.

ZMYŚLONY, R. (1980). Completeness for a family of normal distributions, Mathematical Statistics, Banach Center Publications, 6, pp. 355–357.

ZMYŚLONY, R., ŽEŽULA, I., KOZIOŁ, A., (2018). Application of Jordan Algebra for testing hypotheses about structure of mean vector in model with block compound symmetric covariance structure. Electronic Journal of Linear Algebra: International Conference on Matrix Analysis and its Applications, MAT TRIAD 2017, 33, pp. 41–52.