Testing hypotheses about structure of parameters in models with block compound symmetric covariance structure

Roman Zmyślony [1] Ivan Žežula [2] Arkadiusz Kozioł[1]

[1]Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Szafrana 4a, 65-516 Zielona Góra, Poland

[2]Institute of Mathematics, Faculty of Science, P. J. Šafárik University, Jesenná 5, 040 01 Košice, Slovakia

II Kongres Statystyki Polskiej

Warszawa, July 10-12, 2018

Tests for variance components based on an unbiased estimator

Definition

Jordan algebra is closed under operation $A \circ B = \frac{AB+BA}{2}$.

A full characterization of irreducible Jordan algebra gave Jordan, Neumann i Wigner (1934):

- $\mathbb{R}$ with addition and multiplication;
- S^n - set of symmetric matrices with operation $A \circ B$;
- quaternions;
- special algebra.

Properties of Jordan Algebra

Let ϑ be Jordan Algebra then:

1. $\boldsymbol{A} \in \vartheta \Rightarrow \boldsymbol{A}^k \in \vartheta$;
2. $\boldsymbol{A}, \boldsymbol{B} \in \vartheta \Rightarrow \boldsymbol{ABA} \in \vartheta$;
3. $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \vartheta \Rightarrow \boldsymbol{ABC} + \boldsymbol{CBA} \in \vartheta$;
4. $\boldsymbol{P}^2 = \boldsymbol{P}$ and $\boldsymbol{PV} = \boldsymbol{VP} \Rightarrow \boldsymbol{M}\vartheta\boldsymbol{M}' = \boldsymbol{M}\vartheta$ is Jordan Algebra;
5. Let $\boldsymbol{Q}$ be orthogonal matrix $(\boldsymbol{QQ}' = \boldsymbol{I}) \Rightarrow \boldsymbol{Q}\vartheta\boldsymbol{Q}'$ is quadratic subspace.

Tests for variance components based on an unbiased estimator

$$\underset{n\times 1}{\boldsymbol{y}} \sim N\left(\mu\mathbf{1}, \boldsymbol{\Sigma}\right),$$

where $\mu \in \mathbb{R}$ and $\boldsymbol{\Sigma} = \sigma_1^2 \boldsymbol{V}_1 + \sigma_2^2 \boldsymbol{V}_2 + \sigma_3^2 \boldsymbol{V}_3$.

$$\boldsymbol{V}_1 = \begin{bmatrix} \underset{n_1\times n_1}{\mathbf{11}'} & \mathbf{0} \\ \mathbf{0} & \underset{n_2\times n_2}{\mathbf{0}} \end{bmatrix}, \boldsymbol{V}_2 = \begin{bmatrix} \underset{n_1\times n_1}{\mathbf{0}} & \mathbf{0} \\ \mathbf{0} & \underset{n_2\times n_2}{\mathbf{11}'} \end{bmatrix}, \boldsymbol{V}_3 = \begin{bmatrix} \underset{n_1\times n_1}{\boldsymbol{I}} & \mathbf{0} \\ \mathbf{0} & \underset{n_2\times n_2}{\boldsymbol{I}} \end{bmatrix}.$$

$$E\left(\boldsymbol{yy'}\right) = \mu^2\mathbf{11'} + \boldsymbol{\Sigma} = \mu^2\mathbf{11'} + \sigma_1^2 \boldsymbol{V}_1 + \sigma_2^2 \boldsymbol{V}_2 + \sigma_3^2 \boldsymbol{V}_3.$$

Tests for variance components based on an unbiased estimator

1. $\vartheta = sp\left\{\mathbf{1}\mathbf{1}', \boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{V}_3\right\}$ is a quadratic subspace.
2. $\frac{1}{n}\mathbf{1}\mathbf{1}'$ not commute with $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$.
3. Thus according theorem of characterization of Jordan Algebra it means that ϑ can be represented as Cartesian product of 2×2 symmetric matrices and $\sigma_3^2 \boldsymbol{I}$.

Remark

Note that $\boldsymbol{P} = \frac{1}{n}\mathbf{1}\mathbf{1}'$ does not commute with $\boldsymbol{\Sigma} \Rightarrow$ for μ does not exist BLUE.
Moreover there exists BQUE for μ^2.

Tests for variance components based on an unbiased estimator

$$H_0 : \sigma_i^2 = 0 \quad vs. \quad H_1 : \sigma_i^2 \neq 0$$

Let $y'Ay$ be unbiased estimator of σ_i^2. Moreover, let A_+, A_- stand for positive and negative part of matrix A, respectively.

Remark

For $i < k$ estimator $y'Ay$ is "not defined", that is $A = A_+ - A_-$, where $A_+, A_- \neq 0$. Note that

- *if H_0 is true, then $E\left(y'A_+y\right) = E\left(y'A_-y\right)$,*
- *if H_1 is true, then $E\left(y'A_+y\right) > E\left(y'A_-y\right)$.*

Tests for variance components based on an unbiased estimator

Test should reject hypothesis

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

if statistic

$$F = \frac{y' A_+ y}{y' A_- y}$$

is sufficiently large.

Tests for variance components based on an unbiased estimator

Normal model has the following form:

$$y \sim N\left(X\beta, \sum_{i=1}^{k} \sigma_i^2 V_i\right)$$

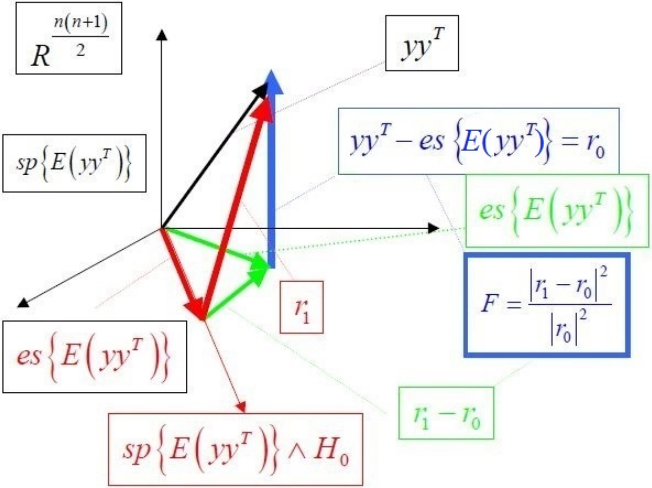
Let us consider three conditions:

1. $sp\left\{MV_1M, \ldots, MV_kM\right\}$ is Jordan algebra,
2. $sp\left\{\{MV_1M, \ldots, MV_kM\} \setminus \{MV_iM\}\right\}$ is commutative Jordan algebra,
3. $F = \frac{y'A_+y}{y'A_-y}$ has F-Snedecor distribution under $H_0 : \sigma_i^2 = 0$.

Theorem (1996): 1. $\wedge$ 2. $\Rightarrow$ 3.

Theorem (2002): 1. $\wedge$ 3. $\Rightarrow$ 2.

F-Test for hypothesis $H_0 : \sigma_i = 0$



Lens of residual is not a function of minimal sufficient statistics

Tests for variance components based on an unbiased estimator

Theorem

Let assume, that subspace

$$sp\left\{MV_1M, \ldots, MV_kM\right\}$$

is commutative Jordan algebra, while

$$sp\left\{\left\{MV_1M, \ldots, MV_kM\right\} \setminus \left\{MV_iM\right\}\right\}$$

is not commutative. Then statistic

$$F = \frac{y'A_+y}{y'A_-y}$$

has generalized F-Snedecor distribution under $H_0 : \sigma_i^2 = 0$, *where* $y'Ay$ *is BIQUE of parameter* σ_i^2.

Special self-adjoint linear operator

Definition

Let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ be matrices with such dimensions that multiplication $\boldsymbol{ACB}$ is possible. Then:

$$(\boldsymbol{A} \odot \boldsymbol{B})\boldsymbol{C} = \boldsymbol{ACB}.$$

Remark

Operator $\odot$ has a following properties:

- $(\boldsymbol{A} \otimes \boldsymbol{B})\mathrm{vec}(\boldsymbol{Y}) = \mathrm{vec}\left((\boldsymbol{B}' \odot \boldsymbol{A})\boldsymbol{Y}\right)$;
- $\mathrm{vec}^{-1}\left((\boldsymbol{A} \otimes \boldsymbol{B})\mathrm{vec}(\boldsymbol{Y})\right) = (\boldsymbol{B}' \otimes \boldsymbol{A})\boldsymbol{Y}$ *if $\boldsymbol{A}$ and $\boldsymbol{B}$ are square matrices and then vec^{-1} is well defined;*
- $(\boldsymbol{A} \odot \boldsymbol{B})(\boldsymbol{C} \odot \boldsymbol{D}) = \boldsymbol{AC} \odot \boldsymbol{DB}$.

Block compound symmetric covariance structure in doubly multivariate data - form

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

$$\begin{aligned}
\mathbf{\Gamma} &= \begin{bmatrix} \mathbf{\Gamma}_0 & \mathbf{\Gamma}_1 & \dots & \mathbf{\Gamma}_1 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{\Gamma}_1 & \mathbf{\Gamma}_1 & \dots & \mathbf{\Gamma}_0 \end{bmatrix} \\
&= (\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1) \odot \boldsymbol{I}_u + \mathbf{\Gamma}_1 \odot \boldsymbol{J}_u \\
&= \mathbf{\Gamma}_0 \odot \boldsymbol{I}_u + \mathbf{\Gamma}_1 \odot (\boldsymbol{J}_u - \boldsymbol{I}_u)\,.
\end{aligned}$$

The above BCS structure can be also written as a **sum of two mutually orthogonal matrices** (i.e. the product of such matrices is equal to matrix $\mathbf{0}$):

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

Block compound symmetric covariance structure in doubly multivariate data - assumptions

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. $-\frac{1}{u-1}\boldsymbol{\Gamma}_0 \prec \boldsymbol{\Gamma}_1$ which means that: $\boldsymbol{\Gamma}_0 + (u-1)\boldsymbol{\Gamma}_1$ is positive definite matrix,
4. $\boldsymbol{\Gamma}_1 \prec \boldsymbol{\Gamma}_0$ which means that: $\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1$ is positive definite matrix.

So that the $um \times um$ matrix $\boldsymbol{\Gamma}$ is positive definite $\big($for a proof, see Lemma 2.1 in Roy and Leiva (2011)$\big)$.

Model with unstructured mean vector

In this model we assume that covariance structure is BCS and **the mean vector changes over sites or over time points**. So $\boldsymbol{\mu}$ has *um* components.

The BCS model can be written in the following way:

$$\underset{um \times n}{\boldsymbol{Y}} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n] \sim N\big((\boldsymbol{I}_{um} \odot \mathbf{1}_n')\boldsymbol{\mu}, \boldsymbol{\Gamma}_{um} \odot \boldsymbol{I}_n\big). \quad (1)$$

It means that matrix $\boldsymbol{Y}$ contains n independent normally distributed random column vectors which are identically distributed with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Gamma}$.

Orthogonal transformation

Let us consider orthogonal transformation $\boldsymbol{I}_{um} \odot \boldsymbol{Q}_n$ on $\boldsymbol{Y}_{um \times n}$.

Proposition

If $Var(\boldsymbol{Y}) = \boldsymbol{\Sigma} \odot \boldsymbol{I}$ with any covariance matrix $\boldsymbol{\Sigma}$ then the model is invariant with respect to transformation $\boldsymbol{I} \odot \boldsymbol{Q}$.

Proposition

Let $\vartheta_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}}$ be the space generated by covariance matrices $\boldsymbol{\Sigma}$ and let $\boldsymbol{P}_{E(\boldsymbol{Y})}$ denote orthogonal projector on 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 we have

$$\text{If } \boldsymbol{P}_{E(\boldsymbol{Y})}\boldsymbol{\Sigma}_{\boldsymbol{Y}} = \boldsymbol{\Sigma}_{\boldsymbol{Y}}\boldsymbol{P}_{E(\boldsymbol{Y})} \text{ then } \boldsymbol{P}_{E(\boldsymbol{U})}\boldsymbol{\Sigma}_{\boldsymbol{U}} = \boldsymbol{\Sigma}_{\boldsymbol{U}}\boldsymbol{P}_{E(\boldsymbol{U})}. \quad (2)$$

$$\text{If } \vartheta_{\boldsymbol{\Sigma}_{\boldsymbol{Y}}} \text{ is a quadratic subspace then } \vartheta_{\boldsymbol{\Sigma}_{\boldsymbol{U}}} \text{ is a quadratic subspace.} \quad (3)$$

Orthogonal transformation - special case

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

Lemma

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

$\boldsymbol{P}_{E(\boldsymbol{U})}$ commutes with $Var(\boldsymbol{U})$ and $\vartheta_{Var(\boldsymbol{U})}$ is a quadratic subspace.

Remark

The proof that for the model (1) $\vartheta_{Var(\boldsymbol{Y})}$ is a quadratic subspace and assumption that commutativity of $\boldsymbol{P}_{E(\boldsymbol{Y})}$ holds see [13].

Orthogonal transformation - step I

Lemma

Let $\boldsymbol{U} = (\boldsymbol{I}_{um} \odot \boldsymbol{Q}_2)\boldsymbol{Y}$, *where* $\boldsymbol{Q}_2 = \left[\frac{1}{\sqrt{n}}\mathbf{1}_n \vdots \boldsymbol{K}_{\mathbf{1}_n}\right]$ *is Helmert matrix, so that* $\boldsymbol{K}'_{\mathbf{1}_n}\boldsymbol{K}_{\mathbf{1}_n} = \boldsymbol{I}_{n-1}$ *and* $\boldsymbol{K}'_{\mathbf{1}_n}\mathbf{1}_n = \mathbf{0}$. *Then* $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n]$ *has independent column vectors, where*

$$\boldsymbol{u}_1 \sim N(\sqrt{n}\boldsymbol{\mu}, \boldsymbol{\Gamma}) \text{ and } \boldsymbol{u}_i \sim N(\mathbf{0}, \boldsymbol{\Gamma}) \text{ for } i = 2, \ldots n.$$

For each $\boldsymbol{u}_i$ we define matrix $\boldsymbol{U}_i$ of size $m \times u$ dividing vector $\boldsymbol{u}_i$ using vec^{-1} for column vector of dim $m \times 1$ i.e.

$$\boldsymbol{U}_i = \left[\boldsymbol{u}_1^{(i)}, \ldots, \boldsymbol{u}_u^{(i)}\right]$$

with distribution

$$\boldsymbol{U}_1 \sim N\left(\sqrt{n}\left[\boldsymbol{\mu}_1^{(1)}, \ldots, \boldsymbol{\mu}_u^{(1)}\right], \boldsymbol{\Gamma}\right),$$

$$\boldsymbol{U}_i \sim N\left(\mathbf{0}_{m\times u}, \boldsymbol{\Gamma}\right) \text{ for } i = 2, \ldots, n.$$

Orthogonal transformation - step II

Now we use the same orthogonal mapping for each matrix $\boldsymbol{U}_i$ which according to the previous proposition saves the property of quadratic subspace generated by covariance structure. Let $\boldsymbol{W}_i = (\boldsymbol{I} \odot \boldsymbol{Q}_1)\boldsymbol{U}_i$, where $\boldsymbol{Q}_1 = \left[\frac{1}{\sqrt{u}}\mathbf{1}_u \vdots \boldsymbol{K}_{\mathbf{1}_u}\right]$. Each matrix $\boldsymbol{W}_i$ can be expressed as

$$\boldsymbol{W}_i = \left[\boldsymbol{w}_1^{(i)}, \ldots, \boldsymbol{w}_u^{(i)}\right],$$

where $\boldsymbol{w}_j^{(i)}$ is $m \times 1$ vector. On can easily prove that

$$Var(\boldsymbol{W}_i) = (\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1) \odot \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \boldsymbol{I}_{u-1} \end{bmatrix} + (\boldsymbol{\Gamma}_0 + (u-1)\boldsymbol{\Gamma}_1) \odot \begin{bmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{0}_{u-1} \end{bmatrix}$$

Orthogonal transformation - step II

Corollary

Vectors $\boldsymbol{w}_j^{(i)}$ *are independent and*

$$\boldsymbol{w}_1^{(1)} \sim N\left(\sqrt{nu}\sum_{j=1}^{u}\boldsymbol{\mu}_j, \boldsymbol{\Gamma}_0 + (u-1)\boldsymbol{\Gamma}_1\right),$$

$$\boldsymbol{w}_1^{(i)} \sim N\left(\boldsymbol{0}, \boldsymbol{\Gamma}_0 + (u-1)\boldsymbol{\Gamma}_1\right) \ \text{for } i = 2, \ldots, n,$$

$$\boldsymbol{w}_j^{(1)} \sim N\left(\sqrt{nu}\sum_{l=1}^{u}\boldsymbol{k}_{l,j-1}\boldsymbol{\mu}_l, \boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1\right) \ \text{for } j = 2, \ldots, u,$$

where $\boldsymbol{k}_{lj}$ *is lj-th element of* $\boldsymbol{K}_{\boldsymbol{1}_u}$.

$$\boldsymbol{w}_j^{(i)} \sim N\left(\boldsymbol{0}, \boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1\right) \ \text{for } i = 2, \ldots, n, \ j = 2, \ldots, u.$$

Hypotheses for structure of mean

Remark

According to full characterization of Jordan Algebra, note that covariance structure is isomorphic to Cartesian product of Jordan Algebra of $n(u-1)$ *and* n *full* $m \times m$ *symmetric matrices* $\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1$ *and* $\mathbf{\Gamma}_0 + (u-1)\mathbf{\Gamma}_1$*, respectively, see [6].*

Now we formulate null hypothesis for structure of mean

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots = \boldsymbol{\mu}_u,$$

This hypothesis can be written equivalently as

$$H_0 : \boldsymbol{\mu}_2^{(c)} = \boldsymbol{\mu}_3^{(c)} = \ldots = \boldsymbol{\mu}_u^{(c)} = 0,$$

where $\boldsymbol{\mu}_j^{(c)} = \sqrt{nu} \sum_{l=1}^{u} \boldsymbol{k}_{l,j-1} \boldsymbol{\mu}_l$.
Following idea of [10] this hypothesis is equivalent

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

One can prove that quadratic estimator of $\sum_{j=2}^{u} \boldsymbol{\mu}_j^{(c)} \boldsymbol{\mu}_j^{(c)'}$ is a function of complete sufficient statistics (see [13]) and has the following form:

$$\widehat{\sum_{j=2}^{u} \boldsymbol{\mu}_j^{(c)} \boldsymbol{\mu}_j^{(c)'}} = \sum_{j=2}^{u} \widehat{\boldsymbol{\mu}}_j^{(c)} \widehat{\boldsymbol{\mu}}_j^{(c)'} - (u-1)\widehat{\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1}. \tag{4}$$

Note that

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

is positive part and

$$(u-1)\widehat{\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1} = \frac{u-1}{(n-1)(u-1)} \sum_{i=2}^{n} \sum_{j=2}^{u} w_j^{(i)} w_j^{(i)'} \stackrel{\mathrm{df}}{=} (u-1)\widehat{\boldsymbol{\Delta}}_1$$

is negative part of estimator in (4).

Under null hypothesis positive part has Wishart distribution and negative part multiplied by $(n-1)$ is Wishart distributed with the same covariance matrix $\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1$

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

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

Test statistic for F-test

Lemma

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

$$T = \frac{n_2 \boldsymbol{x}' \boldsymbol{W}_1 \boldsymbol{x}}{n_1 \boldsymbol{x}' \boldsymbol{W}_2 \boldsymbol{x}} \sim F_{n_1, n_2}.$$

Theorem

Under null hypothesis test statistic

$$T = \frac{\boldsymbol{x}' \sum_{j=2}^{u} \widehat{\boldsymbol{\mu}}_j^{(c)} \widehat{\boldsymbol{\mu}}_j^{(c)'} \boldsymbol{x}}{(u-1)\boldsymbol{x}' \widehat{\boldsymbol{\Gamma}_0 - \boldsymbol{\Gamma}_1} \boldsymbol{x}} = \frac{\boldsymbol{x}' \widehat{\boldsymbol{\Delta}}_2 \boldsymbol{x}}{\boldsymbol{x}' \widehat{\boldsymbol{\Delta}}_1 \boldsymbol{x}} \tag{5}$$

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

Since the distribution of (5) is the same for any $\boldsymbol{x}$, we can look for higher values of T in order to get higher power of the test. Let us denote

$$\boldsymbol{y} = \widehat{\boldsymbol{\Delta}}_1^{1/2}\boldsymbol{x}.$$

This is a regular transformation, since we assume $\boldsymbol{\Delta}_1 > \mathbf{0}$. If the number of degrees of freedom is greater than the dimensionality, i.e. $(n-1)(u-1) > m$, then also $\widehat{\boldsymbol{\Delta}}_1 > \mathbf{0}$ with probability 1. That is why

$$T_m \stackrel{\mathrm{df}}{=} \max_{\boldsymbol{x}} T = \max_{\boldsymbol{y}} \frac{\boldsymbol{y}'\widehat{\boldsymbol{\Delta}}_1^{-1/2}\widehat{\boldsymbol{\Delta}}_2\widehat{\boldsymbol{\Delta}}_1^{-1/2}\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{y}} = \lambda_{max}\left(\widehat{\boldsymbol{\Delta}}_2\widehat{\boldsymbol{\Delta}}_1^{-1}\right).$$

Test statistic for Roy's test

Using the Definition 3.7.2 and Equation 3.7.12 of [9], we can tell that the distribution of

$$R = \frac{\frac{1}{(n-1)} T_m}{1 + \frac{1}{(n-1)} T_m}$$

is Roy's largest root distribution with parameters m, $(n-1)(u-1)$, and $u-1$ if $n-1 > m$. Thus, the hypothesis can also be tested using critical values of Roy's distribution.

However, the maximizing vector $\boldsymbol{x}$ is the eigenvector $\boldsymbol{u}_1$ corresponding to the largest eigenvalue, which is no more fixed but depends on the data. As a consequence, Roy's test **does not necessarily have higher power than the F-test** and performs better than other ones only when the largest eigenvalue is substantially greater than the remaining ones.

Test statistic for likelihood ratio test

LRT for this situation was developed by Fleiss in [4]. The test statistic is of the form

$$L = \frac{\left|\widehat{\boldsymbol{\Delta}}_1\right|}{\left|\widehat{\boldsymbol{\Delta}}_1 + \frac{1}{n}\widehat{\boldsymbol{\Delta}}_2\right|},$$

where $\frac{1}{n}\widehat{\boldsymbol{\Delta}}_2 = \frac{1}{n(u-1)}\sum_{j=2}^{u} \widehat{\boldsymbol{\mu}}_j^{(c)} \widehat{\boldsymbol{\mu}}_j^{(c)\prime} = \frac{1}{u-1}\overline{X}\left(I - \frac{1}{u}J_u\right)\overline{X}'$,
$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $X_i = \mathrm{vec}^{-1}\boldsymbol{y}_i$.

This statistic has under H_0 Wilks lambda distribution with parameters m, $u-1$, and $(n-1)(u-1)$ if $n-1 > m$ (compare with Definition 3.7.1 in [9]). We obtain critical values for both tests by 1 000 000 simulations using Monte Carlo method.

Simulation study

In our test statistic we take vector $\boldsymbol{x} = \mathbf{1}_m$. Using argument of minimal sufficiency we need only to generate independently $\boldsymbol{w}_2^{(1)}, \ldots, \boldsymbol{w}_u^{(1)}$ according $N(\mathbf{0}_m, \boldsymbol{I}_m)$ and random matrix with Wishart distribution $W_m((n-1)(u-1), \boldsymbol{I}_m)$.

In each step of simulation we add randomly chosen vectors $\boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_u$ to the vectors $\boldsymbol{w}_j^{(1)}$ for $j = 2, \ldots, u$ multiplied by fixed λ to obtain power function of the test. Here λ is between 0 and some suitable value Λ such that power is close to 1. Naturally, for $\lambda = 0$ we have null hypothesis. When λ increases then power should also increase. We have compared powers of all three tests as a function of λ.

Figure: $n = 25, u = 2, m = 3$ and all elements of the contrast vector are positive

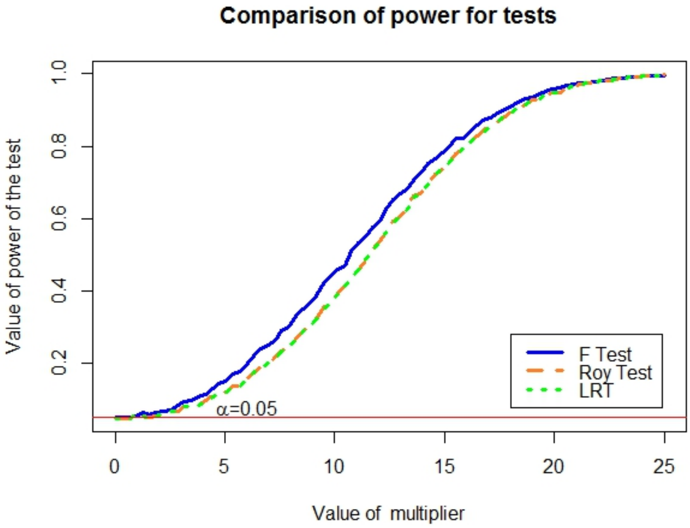


Figure: $n = 10, u = 3, m = 3$ and all elements of the contrast vector are positive

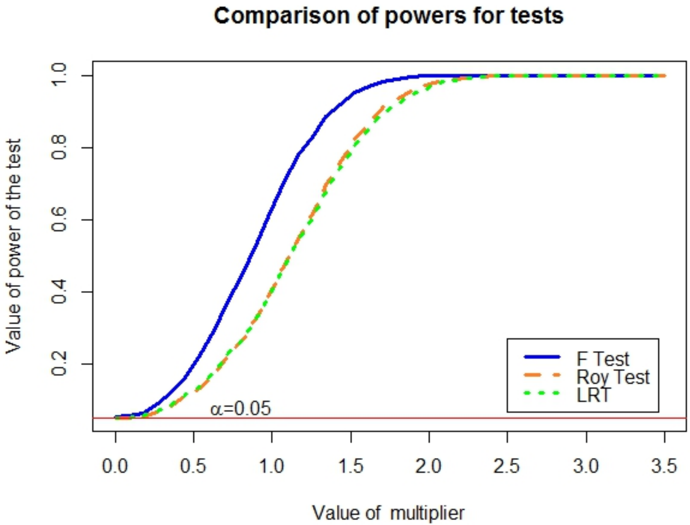


Figure: $n = 25, u = 3, m = 3$ and all elements of the contrast vector are negative

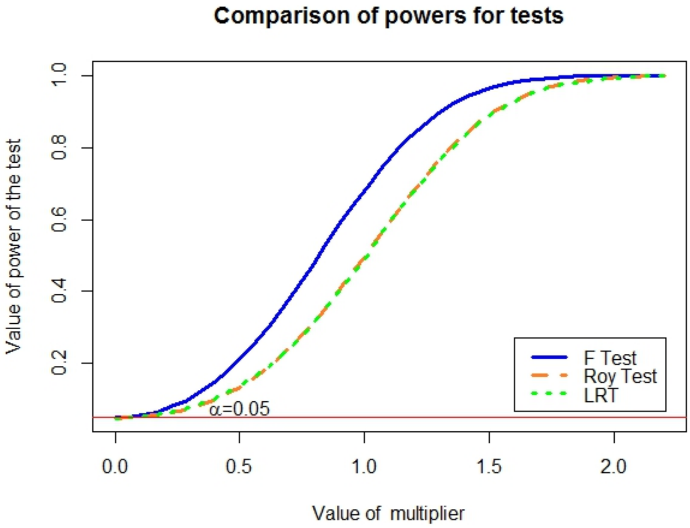
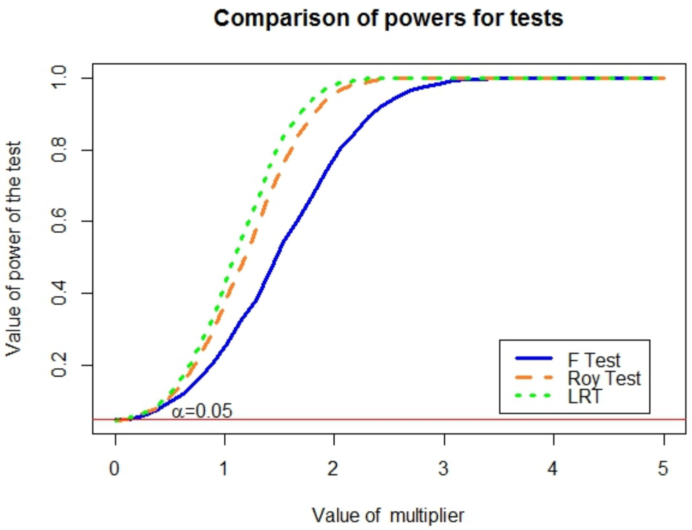


Figure: $n = 25, u = 3, m = 3$ and all elements of contrast vector have different signs



Hypotheses of covariance structure in double multivariate data

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

Applying the idea in [A. Michalski, R. Zmyślony, *Testing hypotheses for variance components in mixed linear models*,Statistics 27(1996), 297-310] for testing hypothesis under assumption that all elements of $\mathbf{\Gamma}_1$ are nonnegative or nonpositive.

Lemma

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

$$T = \frac{\frac{\boldsymbol{x}'\boldsymbol{W}_1\boldsymbol{x}}{n_1}}{\frac{\boldsymbol{x}'\boldsymbol{W}_2\boldsymbol{x}}{n_2}} \sim F_{n_1,n_2}.$$

Hypotheses of covariance structure in double multivariate data

Proof.

According to well-known theorem if $\boldsymbol{W} \sim W_m(\boldsymbol{\Sigma}, n)$ then for every $\boldsymbol{x} \neq 0 \in \mathbb{R}^m$:

$$\frac{\boldsymbol{x}'\boldsymbol{W}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{\Sigma}\boldsymbol{x}} \sim \chi^2_n.$$

Now if we calculate ratio of $\frac{\boldsymbol{x}'\boldsymbol{W}_1\boldsymbol{x}}{n_1}$ and $\frac{\boldsymbol{x}'\boldsymbol{W}_2\boldsymbol{x}}{n_2}$ we get:

$$\frac{\frac{\boldsymbol{x}'\boldsymbol{W}_1\boldsymbol{x}}{n_1}}{\frac{\boldsymbol{x}'\boldsymbol{W}_2\boldsymbol{x}}{n_2}} = \frac{\frac{\boldsymbol{x}'\boldsymbol{W}_1\boldsymbol{x}}{n_1\boldsymbol{x}'\boldsymbol{\Sigma}\boldsymbol{x}}}{\frac{\boldsymbol{x}'\boldsymbol{W}_2\boldsymbol{x}}{n_2\boldsymbol{x}'\boldsymbol{\Sigma}\boldsymbol{x}}} \sim \frac{\frac{\chi^2_{n_1}}{n_1}}{\frac{\chi^2_{n_2}}{n_2}} \sim F_{n_1,n_2}.$$

$\square$

Hypotheses of covariance structure in double multivariate data

From [A. Roy, R. Leiva, I. Žežula, D. Klein, *Testing the equality of mean vectors for paired doubly multivariate observations in blocked compound symmetric covariance matrix setup*, Journal of Multivariate Analysis, 137, 50-60] we get that that matrices:

$$(n-1)(u-1)\widetilde{\mathbf{\Delta}}_1 = (n-1)(u-1)(\widetilde{\mathbf{\Gamma}}_0 - \widetilde{\mathbf{\Gamma}}_1) \sim W_m(\mathbf{\Gamma}_0 - \mathbf{\Gamma}_1, (n-1)(u-1)),$$

$$(n-1)\widetilde{\mathbf{\Delta}}_2 = (n-1)(\widetilde{\mathbf{\Gamma}}_0 + (u-1)\widetilde{\mathbf{\Gamma}}_1) \sim W_m(\mathbf{\Gamma}_0 + (u-1)\mathbf{\Gamma}_1, (n-1))$$

are independent.

Negative and positive part of estimator $\widetilde{\mathbf{\Gamma}}_1$

It is easy to show that:

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

Under the framework of Michalski and Zmyślony positive part of $\widetilde{\mathbf{\Gamma}}_1$ is given by:

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

and negative part is given by:

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

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

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

The test statistic:

$$T = \frac{\mathbf{1}'\widetilde{\mathbf{\Gamma}}_{1+}\mathbf{1}}{\mathbf{1}'\widetilde{\mathbf{\Gamma}}_{1-}\mathbf{1}}$$

is distributed as an F random variable with $(n-1)$ and $(n-1)(u-1)$ degrees of freedom under the hypothesis $H_0 : \mathbf{\Gamma}_1 = 0$.

Likelihood ratio test

In the case of the likelihood ratio test (LRT) to verify hypotheses:

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

there is no need to assume that all elements in matrix $\mathbf{\Gamma}_1$ have the same sign.

Test statistic for this test has the following form:

$$L = \frac{\left| \frac{(n-1)(u-1)\widetilde{\mathbf{\Delta}}_1 + (n-1)\widetilde{\mathbf{\Delta}}_2}{nu} \right|^{-\frac{nu}{2}}}{\left| \frac{n-1}{n} \widetilde{\mathbf{\Delta}}_1 \right|^{-\frac{n(u-1)}{2}} \left| \frac{n-1}{n} \widetilde{\mathbf{\Delta}}_2 \right|^{-\frac{n}{2}}}.$$

Under null hypothesis $H_0 : \mathbf{\Gamma}_1 = 0$ statistic $-2\ln(L)$ has approximately a chi-squared distribution with $\frac{m(m+1)}{2}$ degrees of freedom.

Comparison of power of tests: F and LRT (all elements in matrix $\mathbf{\Gamma}_1$ are positive)

In order to compare power of both test, assumed that $u = 2$, $n = 25$ and matrices $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_1$ are:

$$\mathbf{\Gamma}_0 = \begin{bmatrix} 0.01221 & 0.02172 & 0.00901 \\ 0.02172 & 0.07492 & 0.01682 \\ 0.00901 & 0.01682 & 0.01108 \end{bmatrix},$$

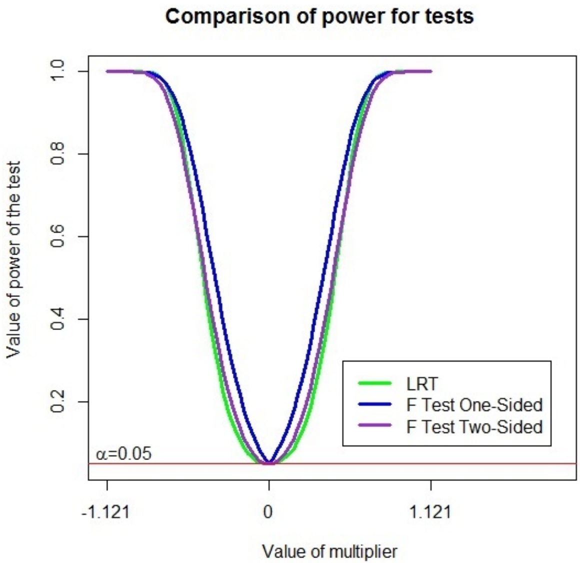
$$\mathbf{\Gamma}_1 = \begin{bmatrix} 0.01038 & 0.01931 & 0.00824 \\ 0.01931 & 0.06678 & 0.01529 \\ 0.00824 & 0.01529 & 0.00807 \end{bmatrix}.$$

Comparison of power of tests: F and LRT (all elements in matrix $\mathbf{\Gamma}_1$ are positive)

For these matrices $\mathbf{\Gamma}_0$, $\mathbf{\Gamma}_1$ and value of u, determined interval for values of multiplier λ, so that the following two conditions are satisfied:

1. $\mathbf{\Gamma}_0 + (u-1)\lambda\mathbf{\Gamma}_1$ is positive definite matrix,
2. $\mathbf{\Gamma}_0 - \lambda\mathbf{\Gamma}_1$ is positive definite matrix.

These conditions ensure positive definite of matrix $\mathbf{\Gamma}$.

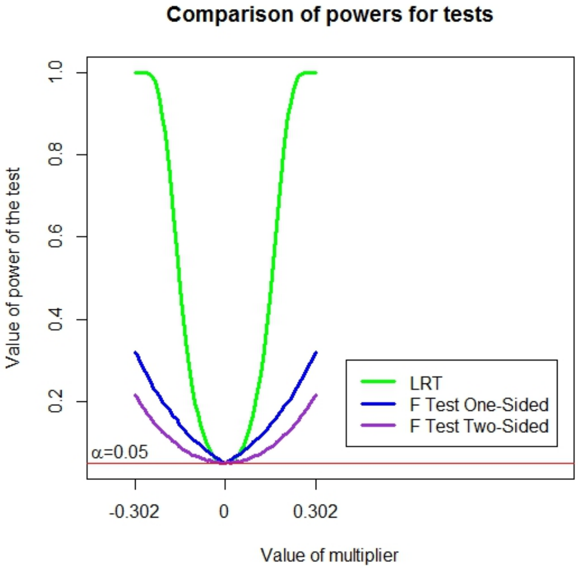

Comparison of power for tests
Value of power of the test
1.0
0.8
0.6
0.4
0.2
α=0.05
LRT
F Test One-Sided
F Test Two-Sided
-1.121
0
1.121
Value of multiplier

Comparison of power of tests: F and LRT (elements in matrix $\mathbf{\Gamma}_1$ have different signs)

It is worth to analyze case with **either negative and positive elements** in matrix $\mathbf{\Gamma}_1$. Intuition suggests that power of F test in this case should be lower than power of LRT. Two examples in this case will be considered.

In this example the last element on main diagonal in matrix $\mathbf{\Gamma}_1$ is multiplied by -1 thus matrix $\mathbf{\Gamma}_1$ is (matrix $\mathbf{\Gamma}_0$, parameters u and n stay unchanged):

$$\mathbf{\Gamma}_1 = \begin{bmatrix} 0.01038 & 0.01931 & 0.00824 \\ 0.01931 & 0.06678 & 0.01529 \\ 0.00824 & 0.01529 & -0.00807 \end{bmatrix}.$$

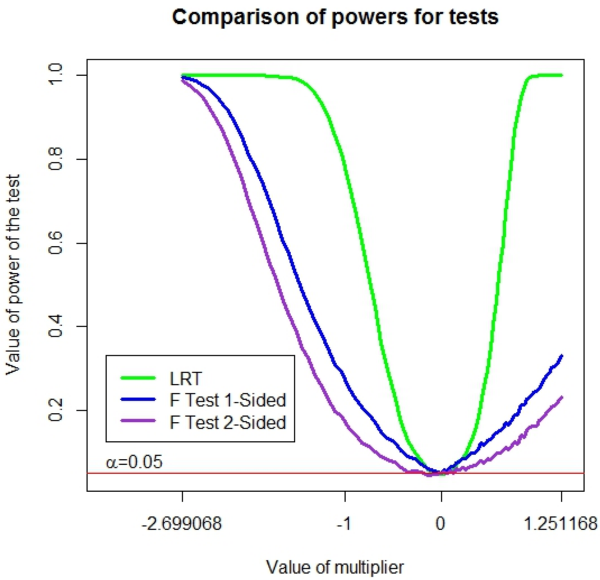

Comparison of powers for tests
Value of power of the test
1.0
0.8
0.6
0.4
0.2
α=0.05
-0.302
0
0.302
Value of multiplier
LRT
F Test One-Sided
F Test Two-Sided

Comparison of power of tests: F and LRT (elements in matrix $\mathbf{\Gamma}_1$ have different signs)

Now consider second example. This time we assumed that $u = 5$, $n = 25$ and matrices $\mathbf{\Gamma}_0$ and $\mathbf{\Gamma}_1$ are:

$$\mathbf{\Gamma}_0 = \begin{bmatrix} 16.25767 & -2.44727 & 1.2296 \\ -2.44727 & 20.40595 & -4.1875 \\ 1.2296 & -4.1875 & 18.56094 \end{bmatrix},$$

$$\mathbf{\Gamma}_1 = \begin{bmatrix} -0.278602 & 1.87846 & 1.26189 \\ 1.87846 & -3.19609 & 1.11567 \\ 1.26189 & 1.11567 & -2.15724 \end{bmatrix}.$$


Comparison of powers for tests
Value of power of the test
1.0
0.8
0.6
0.4
0.2
LRT
F Test 1-Sided
F Test 2-Sided
α=0.05
-2.699068
-1
0
1.251168
Value of multiplier

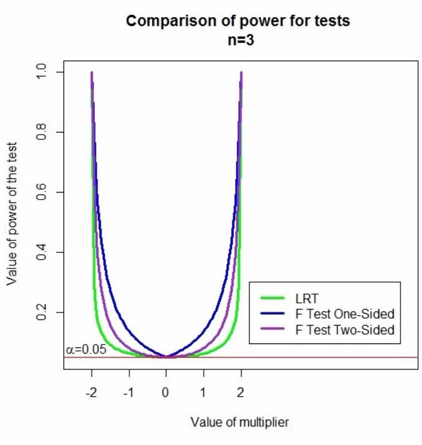
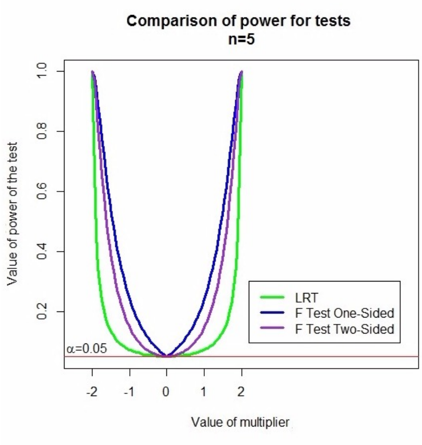
Comparison of power of tests: F and LRT (matrices Γ_0 and Γ_1 are scalars)

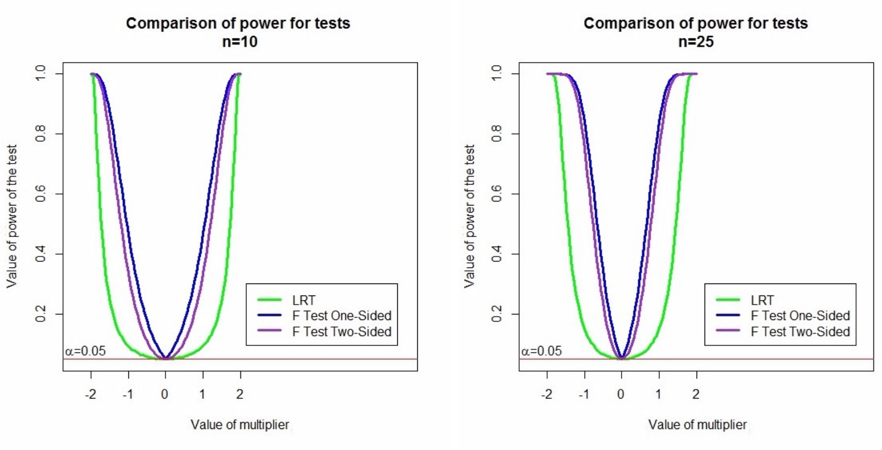
Let us consider another case. Very special one, because Γ_0 and Γ_1 are scalars, thus $m = 1$. Let $\Gamma_0 = 2$ and $\Gamma_1 = 1$. Additionally is assumed that $u = 2$, and parameter n will be one of values from set $\{3, 5, 10, 25\}$.

Matrix $\mathbf{\Gamma}$ has the following form:

$$\mathbf{\Gamma} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

From conditions of positive definiteness of matrix $\mathbf{\Gamma}$ it is easy to show that values of multiplier λ should be from interval $[-2, 2]$.


Comparison of power for tests
n=3
Value of power of the test
1.0
0.8
0.6
0.4
0.2
-2
-1
0
1
2
Value of multiplier
α=0.05
LRT
F Test One-Sided
F Test Two-Sided

Comparison of power for tests
n=5
Value of power of the test
1.0
0.8
0.6
0.4
0.2
-2
-1
0
1
2
Value of multiplier
α=0.05
LRT
F Test One-Sided
F Test Two-Sided


Comparison of power for tests
n=10
Value of power of the test
Value of multiplier
LRT
F Test One-Sided
F Test Two-Sided
α=0.05
Comparison of power for tests
n=25
Value of power of the test
Value of multiplier
LRT
F Test One-Sided
F Test Two-Sided
α=0.05

Test F for single parameters in matrix $\mathbf{\Gamma}_1$

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

In order to conduct F test for above hypotheses, if considered parameter is $\sigma_{ii}^{(1)}, \quad i = 1, \ldots, m$, vectors $\mathbf{1}$ in formula for value of F test statistic should be replaced by

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

For parameters $\sigma_{ij}^{(1)}, \quad i < j, \; i = 1, \ldots, m$, instead of vectors $\mathbf{1}$ in formula for value of F test statistic one should insert

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

Test F for single parameters in matrix $\mathbf{\Gamma_1}$

Matrix with p-values for two sided F test for single parameters in first presented (all elements in $\mathbf{\Gamma_1}$ are positive) example is:

$$\begin{bmatrix} 3.86 * 10^{-8} & 2.57 * 10^{-9} & 0.0599 \\ & 1.02 * 10^{-9} & 5.43 * 10^{-8} \\ & & 2.46 * 10^{-5} \end{bmatrix}.$$

For significance level 0.05, using Bonferroni correction, p-values should be compared with:

$$\alpha_0 = \frac{0.05}{6} \approx 0.0083$$

[1] Arnold, S. F. 1979. Linear models with exchangeably distributed errors. Journal of the American Statistical Association. 74, 194-199.

[2] Arnold, S. F. 1973. Application of the thoery of products of problems to certain patterned covariance matrices. Ann. Statist. 1(4), 682-699.

[3] Drygas, H., 1970. The Coordinate-Free Approach to Gauss-Markov Estimation, Berlin, Heidelberg: Springer.

[4] Fleiss, J. L., 1966. Assessing the Accuracy of Multivariate Observations. Journal of the American Statistical Association 61 (314), Part 1, pp. 403-412.

[5] R.A. Johnson, D.W. Wichern, Applied Multivariate Statistical Analysis, sixth ed., Pearson Prentice Hall, Englewood Cliffs, NJ, 2007.

[6] Jordan, P., Neumann, von, J. and Wigner, E., 1934. On an algebraic generalization of the quantum mechanical formalism. The Annals of Mathematics, 35(1), 29-64.

[7] Kozioł, A., Roy, A., Fonseca, M., Zmyślony, R., Leiva, R., 2018. Free-coordinate estimation for doubly multivariate data. Linear Algebra Appl., https://doi.org/10.1016/j.laa.2018.02.019.

[8] 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.

[9] Mardia, K. V., Kent, J. T., Bibby, J. M., 1979. Multivariate Analysis. New York: Academic Press Inc.

[10] Michalski, A., Zmylony, R., 1999. Testing hypotheses for linear functions of parameters in mixed linear models. Tatra Mountains Mathematical Publications 17, 103-110.

[11] Roy, A., Leiva, R., Žežula, I., and 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, 50-60.

[12] Roy, A. and Leiva, R. 2011. Estimating and testing a structured covariance matrix for three-level multivariate data. Communications in Statistics - Theory and Methods, 40(11), 1945-1963.

[13] Roy, A., Zmyślony, R., Fonseca, M. and Leiva, R. 2016. Optimal estimation for doubly multivariate data in blocked compound symmetric covariance structure, Journal of Multivariate Analysis, Vol. 144, s. 81–90.

[14] Seely, J. F., 1971. Quadratic subspaces and completeness. The Annals of Mathematical Statistics, 42(2), 710-721.

[15] Seely, J. F., 1972. Completeness for a family of multivariate normal distributions. The Annals of Mathematical Statistics, 43, 1644-1647.

[16] Seely, J. F., 1977. Minimal sufficient statistics and completeness for multivariate normal families. Sankhya (Statistics). The Indian Journal of Statistics. Series A, 39(2), 170-185.

[17] Szatrowski, T., 1976. Estimation and testing for block compound symmetry and other patterned covariance matrices with linear and non-linear structure. Technical report No. 107, Dept. of statistics, Stanford University.

[18] Szatrowski, T. H. 1982. Testing and estimation in the block compound symmetry problem. J. Educ. Stat. 7 (1), 318.

[19] Zmyślony, R. 1978. A characterization of best linear unbiased estimators in the general linear model, Lecture Notes in Statistics, 2, 365-373.

[20] Zmyślony, R. 1980. Completeness for a family of normal distributions, Mathematical Statistics, Banach Center Publications 6, 355-357.

200 GBP Prize

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be symmetric, real, nonnegative and not commutative matrices of $n \times n$ sizes. Let define inner product $\langle \boldsymbol{A}, \boldsymbol{B} \rangle = tr\,(\boldsymbol{AB})$. Moreover, let for arbitrary symmetric matrix $\boldsymbol{C}$, matrix $\boldsymbol{C}_+$ be positive part of real symmetric matrix, what means that summation in spectral decomposition only for positive values α_i. For matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and nonnegative x define function:

$$F(x) = \langle (\boldsymbol{A} - x\boldsymbol{B})_+, \boldsymbol{B} \rangle \,.$$

Prove that $F(x)$ is a convex function or not for any value of n.

For $n = 2$ was proved that $F(x)$ is convex function.

My email address is: **rzmyslony@op.pl**