# Least Squares and Generalized Least Squares in Models with Orthogonal Block Structure

Miguel Fonseca, João Tiago Mexia
Department of Mathematics, Faculty of Science and Technology
New University of Lisbon
Monte da Caparica 2829-516 Caparica Portugal
Email: fmig@fct.unl.pt

Roman Zmyślony
Faculty of Mathematics, Computer Science and Econometrics
University of Zielona Góra - Podgórna 50
65-246 Zielona Góra Poland
Email: rzmyslony@wmie.uz.zgora.pl

**Abstract**

Besides the basic model, Kronecker products of rotated models are used to isolate the variance components as parameters of a linear model. A characterization of BLUE given by Zmyślony is applied to the different models. Generalized least squares are used to complete the estimation.

## 1 Estimation of Variance in OBS Models

There are many approaches and many references to the problem of variance components estimation. One of them is the Bayesian approach, for example in [4] and [12]. Another one is the nonnegative approach. In [2] there were obtained two classes of invariant quadratic estimators for the first variance componentwhich are easy to compute and exhibit optimal properties under the mean squared error.

In our paper we will deal with unbiased estimation for fixed parameters and variance components in two derived models.

**Definition 1.** *A linear model has orthogonal block structure – OBS – if the covariance matrix of the observations vector* $\mathbf{y}$ *has the form*

$$\boldsymbol{\Sigma}_{\mathbf{y}} = \sum_{i=1}^{w} \gamma_i \mathbf{P}_i,$$

*were* $\mathbf{P}_1, \ldots, \mathbf{P}_w$ *are orthogonal projection matrices such that* $\sum_{i=1}^{w} \mathbf{P}_i = \mathbf{I}_n$ *and* $\mathbf{P}_i\mathbf{P}_j = \mathbf{0}$, $i \neq j$.

Variance components are $\gamma_1, \ldots, \gamma_w$ will be assumed non-negative, and the mean vector will have the form

$$\boldsymbol{\mu} = \boldsymbol{\Delta}'\boldsymbol{\nu}, \tag{1}$$

were the lines of $\mathbf{\Delta}$ are linearly independent. Let $\Omega = \mathrm{R}(\mathbf{\Delta}')$ and $\mathbf{T}$ the orthogonal projection matrix on $\Omega$.

**Definition 2.** *If a model as OBS and* $\mathbf{T}$ *commutes with* $\mathbf{P}_i$, $i = 1, \ldots, w$, *the model is said to have commutative orthogonal block structure – COBS.*

The goal is to estimate variance components and estimable functions $\mathbf{a}'\boldsymbol{\nu}$.

**Definition 3.** *A function* $g(\boldsymbol{\nu})$ *is estimable if there exists* $\mathbf{a}$ *such that*

$$\mathbb{E}[\mathbf{a}'\boldsymbol{\nu}]\,.$$

Besides obtaining conditions for obtaining estimators with optimal properties, BLUE and even UMVUE, computation techniques will be investigated. It will be shown that it is possible to obtain, from models with COBS, linear models in which the expectation has the form $\mathbf{X}\boldsymbol{\gamma}$, with $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_w)'$, which will permit a unified analysis with least squares – LS – and generalized least squares – GLS – for $\boldsymbol{\nu}$ and $\boldsymbol{\gamma}$. This is distinct of the usual analysis of models with OBS in which, according to [7] and [8], in which orthogonal projections of $\mathbf{y}$ on $\nabla_i = \mathrm{R}(\mathbf{P}_i)$, $i = 1, \ldots, w$. Although cases in which $w = 2$ and $w = 3$ are presented in [1], the methodology used seems difficult to apply when $w \geq 4$.

## 2 General Results

The analysis to be presented both for the initial an derived model is fitted in the approach proposed by [3]. Admit now that the model has the mean vector

$$\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta} \tag{2}$$

and that its covariance matrix is a linear combination of known matrices $\mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_w$, with $\mathrm{R}(\mathbf{V}_j) \subseteq \mathrm{R}(\mathbf{V}_0)$, $j = 1, \ldots, w$. Notice that, according to [6], for any linear model there exists a covariance matrix with maximal column space. It is admitted that $\mathbf{V}_0$ is such a matrix and that it is regular. These conditions will be verified in the models presented.

**Theorem 1** (Zmyślony 1980)**.** *Putting* $\mathbf{M} = \mathbf{I} - \mathbf{T}$, *where* $\mathbf{T}$ *is the orthogonal projection matrix on* $\mathrm{R}(\mathbf{X})$,

$$\mathbf{U} = \mathbf{X}'\mathbf{V}_0^{-1}\left(\sum_{i=1}^{w} \mathbf{V}_i\mathbf{M}\mathbf{V}_i\right)\mathbf{V}_0^{-1}\mathbf{X}$$

*and*

$$\mathbf{Z} = \mathbf{I} - \mathbf{U}^{+}\mathbf{U},$$

*a BLUE for* $\mathbf{c}'\boldsymbol{\beta}$ *exists if and only if*

$$\mathbf{c} \in \mathrm{R}\big(\mathbf{X}'\mathbf{V}_0^{-1}\mathbf{X}\mathbf{Z}\big)\,.$$

*Proof.* See [11]. □

The form of such estimator is also achieved.

**Theorem 2.** *when the conditions on theorem 1 are fulfilled, the BLUE for* $\mathbf{c}'\boldsymbol{\beta}$ *has the form*

$$\mathbf{c}'\widetilde{\boldsymbol{\beta}} = \mathbf{c}'(\mathbf{X}'\mathbf{V}_0\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_0^{-1}\mathbf{y}.$$

When $\mathbf{U} = \mathbf{0}$, $\mathbf{Z} = \mathbf{I}$ and

$$\mathrm{R}\big(\mathbf{X}'\mathbf{V}_0^{-1}\mathbf{X}\big) = \mathrm{R}(\mathbf{X}'), \tag{3}$$

BLUE will exist for all estimable functions.

# 3 Initial Model

The results of the previous section will now be applied to a model with OBS. As referred in previously, the observations vector $\mathbf{y}$ has the mean vector $\boldsymbol{\mu} = \boldsymbol{\Delta}'$ and covariance matrix $\boldsymbol{\Sigma}(\mathbf{y}) = \sum_{i=1}^{w} \gamma_i \mathbf{P}_i$. It is assumed that $\sum_{i=1}^{w} \mathbf{P}_i = \mathbf{I}$ and so, in particular,

$$\mathbf{V}_0 = \mathbf{I}. \tag{4}$$

Considering also $\mathbf{V}_i = \mathbf{P}_i$, $i = 1, \ldots, w$, and $\mathbf{X} = \boldsymbol{\Delta}'$,

$$\mathbf{U} = \boldsymbol{\Delta}\left(\sum_{i=1}^{w} \mathbf{P}_i\mathbf{M}\mathbf{P}_i\right)\boldsymbol{\Delta}' \tag{5}$$

and a BLUE $\mathbf{c}'\widetilde{\boldsymbol{\beta}}$ exists if and only if

$$\mathbf{c} \in \mathrm{R}(\mathbf{X}'\mathbf{X}\mathbf{Z}), \tag{6}$$

having thus

$$\widetilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{7}$$

Consider now $\mathbf{T}_1$ as the orthogonal projection matrix on $\mathrm{R}(\mathbf{X}'\mathbf{X}\mathbf{Z})$ and $\mathbf{M}_1 = \mathbf{I} - \mathbf{T}_1$. Theorem 1 can be reexpressed as

$$\mathbf{M}_1\mathbf{c} = \mathbf{0}. \tag{8}$$

This model will be used further on in GLS estimation.

# 4 First Derived Model

**Lemma 1.** *OBS is closed for orthogonal transformations.*

*Proof.* It is easily seen that, given a model $\mathbf{y}$ with OBS and an orthogonal matrix $\mathbf{K}$, $\mathbf{y}^* = \mathbf{K}\mathbf{y}$ has

$$\boldsymbol{\Delta}^*\boldsymbol{\nu} = \mathbf{K}\boldsymbol{\Delta}\boldsymbol{\nu}$$

and

$$\boldsymbol{\Sigma}(\mathbf{y}^*) = \sum_{i=1}^{w} \gamma_i \mathbf{K}\mathbf{P}_i\mathbf{K}'.$$

Furthermore, $\mathbf{P}_i^* = \mathbf{K}\mathbf{P}_i\mathbf{K}'$, $i = 1, \ldots, w$, are still orthogonal projection matrices, mutually orthogonal, and

$$\sum_{i=1}^{w} \mathbf{P}_i^* = \mathbf{I}.$$

□

Let $\mathbf{K}$ be an orthogonal matrix whose first $n-k$ line vectors form an orthonormal basis for $\Omega^\perp = \mathrm{R}(\boldsymbol{\Delta})^\perp$. Putting

$$\eta_j = \boldsymbol{\alpha}_j'\boldsymbol{\mu}; j = 1, \ldots, n \tag{9}$$

one gets $\eta_j = 0$, $j = 1, \ldots, n-k$. With $m = n-k$, putting

$$l = l(j,h) = \frac{h(h-1)}{2} + j; \quad j = 1, \ldots, h; h = 1, \ldots, m, \tag{10}$$

and

$$y_l^\circ = \mathbf{y}'\mathbf{A}_{j,h}\mathbf{y}; j = 1, \ldots, h; h = 1, \ldots, m; l = 1, \ldots, \frac{m(m+1)}{2}, \tag{11}$$

where

$$\mathbf{A}_{j,h} = \frac{1}{2}(\boldsymbol{\alpha}_j\boldsymbol{\alpha}_h' + \boldsymbol{\alpha}_h\boldsymbol{\alpha}_j'); j = 1, \ldots, h; h = 1, \ldots, m. \tag{12}$$

Then, see [9] pg. 352,

$$\begin{aligned}
\boldsymbol{\mu}_l^\circ = \mathbb{E}[y_l^\circ] = \mathrm{tr}(\mathbf{A}_{j,h}\boldsymbol{\Sigma}(\mathbf{y})) = \sum_{i=1}^{w} \gamma_i \mathrm{tr}(\mathbf{A}_{j,h}\mathbf{P}_i) \\
= \sum_{i=1}^{w} \frac{\gamma_i}{2}\mathrm{tr}\big((\boldsymbol{\alpha}_j\boldsymbol{\alpha}_h' + \boldsymbol{\alpha}_h\boldsymbol{\alpha}_j')\mathbf{P}_i\big) = \sum_{i=1}^{w} \frac{\gamma_i}{2}\mathrm{tr}\big(\boldsymbol{\alpha}_j'\mathbf{P}_i\boldsymbol{\alpha}_h + \boldsymbol{\alpha}_h'\mathbf{P}_i\boldsymbol{\alpha}_j\big) \\
= \sum_{i=1}^{w} \frac{\gamma_i}{2}(c_{i,j,h} + c_{i,h,j}) = \sum_{i=1}^{w} \gamma_i c_{i,h,j},
\end{aligned} \tag{13}$$

where

$$c_{i,j,h} = \boldsymbol{\alpha}_j'\mathbf{P}_i\boldsymbol{\alpha}_h = x_{l,i}^\circ. \tag{14}$$

Thus, with

$$\mathbf{X}^\circ = [x_{l,i}^\circ] \tag{15}$$

one gets the mean vector

$$\boldsymbol{\mu}^\circ = \mathbf{X}^\circ\boldsymbol{\gamma}. \tag{16}$$

In this way, $\boldsymbol{\gamma}$ is now the parameter in the mean vector.

In order to obtain the covariance of this model, it will be assumed that $\mathbf{y}$ has normal distribution. Then, with $y_l^\circ = \mathbf{y}'\mathbf{A}_{j,h}\mathbf{y}$ and $y_{l'}^\circ = \mathbf{y}'\mathbf{A}_{j',h'}\mathbf{y}$, see [9] pg. 395,

$$
\begin{aligned}
\mathrm{Cov}[y_l^\circ; y_{l'}^\circ] &= 2\mathrm{tr}(\mathbf{A}_{j,h}\boldsymbol{\Sigma}(\mathbf{y})\mathbf{A}_{j',h'}\boldsymbol{\Sigma}(\mathbf{y})) \\
&= 2\sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}\mathrm{tr}(\mathbf{A}_{j,h}\mathbf{P}_i\mathbf{A}_{j',h'}\mathbf{P}_{i'}) \\
&= \frac{1}{2}\sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}\mathrm{tr}(\boldsymbol{\alpha}_j\boldsymbol{\alpha}_h'\mathbf{P}_i\boldsymbol{\alpha}_{j'}\boldsymbol{\alpha}_{h'}'\mathbf{P}_{i'} + \boldsymbol{\alpha}_j\boldsymbol{\alpha}_h'\mathbf{P}_i\boldsymbol{\alpha}_{h'}\boldsymbol{\alpha}_{j'}'\mathbf{P}_{i'} \\
&\quad + \boldsymbol{\alpha}_h\boldsymbol{\alpha}_j'\mathbf{P}_i\boldsymbol{\alpha}_{j'}\boldsymbol{\alpha}_{h'}'\mathbf{P}_{i'} + \boldsymbol{\alpha}_h\boldsymbol{\alpha}_j'\mathbf{P}_i\boldsymbol{\alpha}_{h'}\boldsymbol{\alpha}_{j'}'\mathbf{P}_{i'}) \\
&= \frac{1}{2}\sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}\mathrm{tr}(\boldsymbol{\alpha}_h'\mathbf{P}_i\boldsymbol{\alpha}_{j'}\boldsymbol{\alpha}_{h'}'\mathbf{P}_{i'}\boldsymbol{\alpha}_j + \boldsymbol{\alpha}_h'\mathbf{P}_i\boldsymbol{\alpha}_{h'}\boldsymbol{\alpha}_{j'}'\mathbf{P}_{i'}\boldsymbol{\alpha}_j \\
&\quad + \boldsymbol{\alpha}_j'\mathbf{P}_i\boldsymbol{\alpha}_{j'}\boldsymbol{\alpha}_{h'}'\mathbf{P}_{i'}\boldsymbol{\alpha}_h + \boldsymbol{\alpha}_j'\mathbf{P}_i\boldsymbol{\alpha}_{h'}\boldsymbol{\alpha}_{j'}'\mathbf{P}_{i'}\boldsymbol{\alpha}_h) \\
&= \frac{1}{2}\sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}(c_{i,h,j'}c_{i',h',j} + c_{i,h,h'}c_{i',j',j} \\
&\quad + c_{i,j,j'}c_{i',h',h} + c_{i,j,h'}c_{i',j',h}).
\end{aligned} \tag{17}
$$

**Theorem 3.** *There exists a covariance matrix of* $\mathbf{y}^\circ$, $\boldsymbol{\Sigma}(\mathbf{y}^\circ)$, *that is regular.*

*Proof.* Because

$$\sum_{i=1}^{w} c_{i,j,h} = \boldsymbol{\alpha}_j'\mathbf{P}_i\boldsymbol{\alpha}_h = \boldsymbol{\alpha}_j'\boldsymbol{\alpha}_h = \delta_{j,h}; j,h = 1,\ldots,n, \tag{18}$$

follows

$$\sum_{i=1}^{w}\sum_{i'=1}^{w} c_{i,j,h'}c_{i',j',h} = \delta_{j,h'}\delta_{j',h}; j,h,j',h' = 1,\ldots,n. \tag{19}$$

Making $\boldsymbol{\gamma} = \mathbf{1}$, one gets

$$\mathrm{Cov}[y_l^\circ; y_{l'}^\circ] = \frac{1}{2}(\delta_{h,j'}\delta_{h',j} + \delta_{h,h'}\delta_{j',j} + \delta_{j,j'}\delta_{h',h} + \delta_{j,h'}\delta_{j',h}), \tag{20}$$

thus

$$\begin{cases} \{j,h\} \neq \{j',h'\}, & \mathrm{Cov}[y_l^\circ; y_{l'}^\circ] = 0 \\ \{j,h\} = \{j',h'\}, & \begin{cases} j \neq h, & \mathrm{Cov}[y_l^\circ; y_{l'}^\circ] = 1 \\ j = h, & \mathrm{Cov}[y_l^\circ; y_{l'}^\circ] = 2 \end{cases} \end{cases}. \tag{21}$$

Because $l = l'$ if and only if $\{j,h\} = \{j',h'\}$, with $\boldsymbol{\gamma} = \mathbf{1}$, $\boldsymbol{\Sigma}(\mathbf{y}^\circ)$ is a diagonal matrix with principal elements equal to 1 when $j \neq h$ and 2 when $j = h$. □

On the other hand, from (17), it is possible to obtain a set of generators for $\mathscr{V}$, the space of covariance matrices of $\mathbf{y}^\circ$. Take

$$d = (i-1)w + i'; i,i' = 1,\ldots,w, \tag{22}$$

as well as

$$b_{d,l,l'} = \frac{1}{2}(c_{i,h,j'}c_{i',h',l} + c_{i,h,h'}c_{i',j',j} + c_{i,j,j'}c_{i',h',h} + c_{i,j,h'}c_{i',j',h}), \tag{23}$$

where

$$\begin{aligned}
d &= 1, \dots, w^2;\\
i, j &= 1, \dots, w;\\
l, l' &= 1, \dots, \frac{m(m+1)}{2};\\
j &= 1, \dots, h;\\
j' &= 1, \dots, h';\\
h, h' &= 1, \dots, m.
\end{aligned}$$

One obtains, then, the set of generators $\{\mathbf{B}_0, \mathbf{B}_1, \dots, \mathbf{B}_{w^2}\}$ for $\mathscr{V}$, allowing the use of the general results.

# 5 Second Derived Model

Consider the second derived model as the crossed product between fixed and random parts of the model, namely the model whose observations are

$$y_l^{\circ\circ} = \mathbf{y}'\mathbf{A}_{j,h}\mathbf{y}; j = 1, \dots, m; h = m+1, \dots, n. \tag{24}$$

Index $l$ must be redefined in the following way:

$$\begin{cases} l = l(j,h) = \frac{h(h-1)}{2} + j; & j = 1, \dots, h; h = 1, \dots, m \\ l = l(j,h) = \frac{m(m-1)}{2} + (h-m)m + j; & j = 1, \dots, m; h = m+1, \dots, n \end{cases}. \tag{25}$$

**Lemma 2.** *For the second derived model,*

$$\mathbb{E}[y_l^{\circ\circ}] = \sum_{i=1}^{w} c_{i,j,h}\gamma_i$$

*and*

$$\mathrm{Cov}[y_l^{\circ\circ}; y_{l'}^{\circ\circ}] = \frac{1}{2}\sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}b_{d,l,l'} + \sum_{i=1}^{w}\gamma_i\eta_h c_{i,j,j'}\eta_{h'}.$$

*Proof.* For the expectation of $\mathbb{E}[y_l^{\circ\circ}]$, see [9], pg. 352, one gets

$$\begin{aligned}
\mathbb{E}[y_l^{\circ\circ}] &= \mathrm{tr}(\mathbf{A}_{j,h}\boldsymbol{\Sigma}(\mathbf{y})) + \boldsymbol{\mu}'\mathbf{A}_{j,h}\boldsymbol{\mu}\\
&= \sum_{i=1}^{w} c_{i,j,h}\gamma_i + \boldsymbol{\mu}'(\boldsymbol{\alpha}_j\boldsymbol{\alpha}_h' + \boldsymbol{\alpha}_h\boldsymbol{\alpha}_j')\boldsymbol{\mu}\\
&= \sum_{i=1}^{w} c_{i,j,h}\gamma_i + \eta_j\eta_h\\
&= \sum_{i=1}^{w} c_{i,j,h}\gamma_i,
\end{aligned}$$

with $\eta_j = 0$ when $j = 1\dots, m$. As for the covariance, note that because $j, j' \le m$ and [9], pg. 358,

$$\begin{aligned}
\mathrm{Cov}[y_l^{\circ\circ}; y_{l'}^{\circ\circ}] &= 2\mathrm{tr}(\mathbf{A}_{j,h}\boldsymbol{\Sigma}(\mathbf{y})\mathbf{A}_{j',h'}\boldsymbol{\Sigma}(\mathbf{y})) + 4\boldsymbol{\mu}'\mathbf{A}_{j,h}\boldsymbol{\Sigma}(\mathbf{y})\mathbf{A}_{j',h'}\boldsymbol{\mu} \\
&= \frac{1}{2}\sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}b_{d,l,l'} \\
&\quad + \sum_{i=1}^{w}\gamma_i\boldsymbol{\mu}'(\boldsymbol{\alpha}_j\boldsymbol{\alpha}_h + \boldsymbol{\alpha}_h\boldsymbol{\alpha}_j)\mathbf{P}_i(\boldsymbol{\alpha}_{j'}\boldsymbol{\alpha}_{h'} + \boldsymbol{\alpha}_{h'}\boldsymbol{\alpha}_{j'})\boldsymbol{\mu} \\
&= \frac{1}{2}\sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}b_{d,l,l'} + \sum_{i=1}^{w}\gamma_i\eta_h c_{i,j,j'}\eta_{h'}
\end{aligned}$$

$\square$

Matrices $\mathbf{B}_1, \dots, \mathbf{B}_{w^2}$ are sub-matrices of

$$\mathbf{G}_{d,i} = [b_{d,l,l'}]; d = 1, \dots, w^2, \tag{26}$$

with $l, l' = 1, \dots, \frac{m}{2}(2n - m + 1)$. Consider also

$$\mathbf{G}_{i,h,h'} = \boldsymbol{\delta}_h(n)\boldsymbol{\delta}'_{h'}(n) \otimes \mathbf{C}_i; h, h' = m + 1, \dots, n, \tag{27}$$

where $\boldsymbol{\delta}_h(n)$ is a vector of size $n$ whose only non-null component is the $h$-th one, which is equal to 1 and $\mathbf{C}_i = [c_{i,j,j'}]$, $j, j' = 1, \dots, m$, $i = 1, \dots, w$. Therefore,

$$\boldsymbol{\Sigma}(\mathbf{y}^{\circ\circ}) = \sum_{i=1}^{w}\sum_{i'=1}^{w}\gamma_i\gamma_{i'}\mathbf{B}^+ + \sum_{i=1}^{w}\sum_{h=m+1}^{n}\sum_{h'=m+1}^{n}\gamma_i\eta_h\eta_{h'}\mathbf{G}_{i,h,h'}. \tag{28}$$

In this case, $\boldsymbol{\beta}$ is substituted by $\boldsymbol{\eta} = (\eta_{m+1}, \dots, \eta_n)'$.

In order to obtain a regular, and therefore maximal, matrix of covariance for $\mathbf{y}^{\circ\circ}$, take $\boldsymbol{\gamma} = \mathbf{1}$ and $\boldsymbol{\eta} = \mathbf{0}$, thus obtaining

$$\mathbf{G}^\circ = \sum_{i=1}^{w}\sum_{i'=1}^{w}\mathbf{B}^+ = \left[\begin{array}{c|c} \mathbf{B}^\circ & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I} \end{array}\right]. \tag{29}$$

This makes the results shown applicable.

# 6 Generalized Least Squares

In the previous sections, conditions were described in order to obtain BLUE both for estimable functions and variance components, as well as linear combinations of them, but these conditions are not satisfied most of the times. An interesting technique is generalized least squares, seen for instance in [5].

In table 1 the three models' parameters and notations are depicted.

Interestingly, the only parameter in the first derived model is $\boldsymbol{\gamma}$.

As a starting point, consider the estimate

$$\widetilde{\gamma}_{1,0} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}_1. \tag{30}$$

Table 1: Model Parameter and Notation

| Model | $\mathbb{E}$ | $\boldsymbol{\Sigma}$ | Design Matrix | Observations Vector |
|---|---|---|---|---|
| Base | $\mathbb{E}[\mathbf{y} \mid \boldsymbol{\nu}]$ | $\boldsymbol{\Sigma}(\boldsymbol{\gamma})$ | $\mathbf{X}$ | $\mathbf{y}$ |
| $1^{\text{st}}$ derived | $\mathbb{E}[\mathbf{y}_1 \mid \boldsymbol{\gamma}]$ | $\boldsymbol{\Sigma}_1(\boldsymbol{\gamma})$ | $\mathbf{X}_1$ | $\mathbf{y}_1$ |
| $2^{\text{nd}}$ derived | $\mathbb{E}[\mathbf{y}_2 \mid \boldsymbol{\gamma}]$ | $\boldsymbol{\Sigma}_2(\boldsymbol{\gamma}, \boldsymbol{\nu})$ | $\mathbf{X}_2$ | $\mathbf{y}_2$ |

From this estimate one also gets an estimate for the covariance matrix of $\mathbf{y}_2$, which will be denoted has $\boldsymbol{\Sigma}_2(\widetilde{\boldsymbol{\gamma}}_{1,i})$, $i = 0, \ldots$, rendering possible an improved estimation

$$\widetilde{\boldsymbol{\gamma}}_{1,1} = \left(\mathbf{X}_1'\boldsymbol{\Sigma}_1(\widetilde{\boldsymbol{\gamma}}_{1,0})^{-1}\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\boldsymbol{\Sigma}_1(\widetilde{\boldsymbol{\gamma}}_{1,0})^{-1}\mathbf{y}_1, \tag{31}$$

leading to the iterative process

$$\widetilde{\boldsymbol{\gamma}}_{1,i+1} = \left(\mathbf{X}_1'\boldsymbol{\Sigma}_1(\widetilde{\boldsymbol{\gamma}}_{1,i})^{-1}\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\boldsymbol{\Sigma}_1(\widetilde{\boldsymbol{\gamma}}_{1,i})^{-1}\mathbf{y}_1. \tag{32}$$

A stopping criterion for this process is when $\|\widetilde{\boldsymbol{\gamma}}_{1,i+1} - \widetilde{\boldsymbol{\gamma}}_{1,i}\|^2 < \epsilon$. This method is similar to MINQUE estimation (see [10]). Let $\widetilde{\boldsymbol{\gamma}}$ be the resulting estimate of this method.

The estimate $\widetilde{\boldsymbol{\gamma}}$ can also be used in the estimation of $\boldsymbol{\nu}$, taking

$$\widetilde{\boldsymbol{\nu}} = \left(\mathbf{X}'\boldsymbol{\Sigma}(\widetilde{\boldsymbol{\gamma}})^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Sigma}(\widetilde{\boldsymbol{\gamma}})^{-1}\mathbf{y} \tag{33}$$

Similarly, an alternative estimate for $\boldsymbol{\gamma}$ can be obtained:

$$\widetilde{\widetilde{\boldsymbol{\gamma}}} = \left(\mathbf{X}_2'\boldsymbol{\Sigma}_2(\widetilde{\boldsymbol{\nu}}, \widetilde{\boldsymbol{\gamma}})^{-1}\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\boldsymbol{\Sigma}_2(\widetilde{\boldsymbol{\nu}}, \widetilde{\boldsymbol{\gamma}})^{-1}\mathbf{y}_2. \tag{34}$$

*Remark.* In this case, $\widetilde{\boldsymbol{\gamma}}$ and $\widetilde{\widetilde{\boldsymbol{\gamma}}}$ should not differ significantly.

# 7 Estimable Vectors

Let $\widetilde{\boldsymbol{\gamma}}$ be the estimator obtained in section 6 for $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}(\widetilde{\boldsymbol{\gamma}})$ the estimated covariance matrix for $\mathbf{y}$. Hence, a generalized least squares estimator for $\boldsymbol{\nu}$ is

$$\left(\mathbf{X}'\boldsymbol{\Sigma}(\widetilde{\boldsymbol{\gamma}})^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\boldsymbol{\Sigma}(\widetilde{\boldsymbol{\gamma}})^{-1}\mathbf{y}. \tag{35}$$

This leads easily to the natural assumption that, if $\mathbf{y}$ is normally distributed, then

$$\widetilde{\boldsymbol{\nu}} \dot{\sim} \mathcal{N}\left(\boldsymbol{\nu}, \widetilde{\mathbf{K}}\right), \tag{36}$$

with

$$\widetilde{\mathbf{K}} = \left(\mathbf{X}'\boldsymbol{\Sigma}(\widetilde{\boldsymbol{\gamma}})^{-1}\mathbf{X}\right)^{-1}. \tag{37}$$

Thus comes

**Lemma 3.** *Let* $\mathbf{A}$ *be a matrix with rank* $s$. *Then* $\widetilde{\boldsymbol{\psi}} = \mathbf{A}\widetilde{\boldsymbol{\nu}}$ *is an unbiased estimator of* $\boldsymbol{\psi} = \mathbf{A}\boldsymbol{\nu}$, *with*

$$\widetilde{\boldsymbol{\psi}} \dot{\sim} \mathcal{N}\left(\boldsymbol{\nu}, \widetilde{\mathbf{K}}\right),$$

*as well as*

$$(\boldsymbol{\psi} - \widetilde{\boldsymbol{\psi}})'\left(\mathbf{A}\widetilde{\mathbf{K}}\mathbf{A}'\right)^{-1}(\boldsymbol{\psi} - \widetilde{\boldsymbol{\psi}}) \dot{\sim} \chi^2_{(s)},$$

*were* $\dot{\sim}$ *stands for "approximately distributed".*

*Proof.* Taking $\boldsymbol{\Sigma}(\widetilde{\gamma})$ as an approximation of $\boldsymbol{\Sigma}(\mathbf{y})$ the conclusions are easily achieved. $\square$

Taking $x_{s,q}$ as the quantil of the chi square distribution with $s$ degrees of freedom for probability $q$,

$$(\boldsymbol{\psi} - \widetilde{\boldsymbol{\psi}})'\left(\mathbf{A}\widetilde{\mathbf{K}}\mathbf{A}'\right)^{-1}(\boldsymbol{\psi} - \widetilde{\boldsymbol{\psi}}) \leq x_{s,1-\alpha} \tag{38}$$

defines a $(1-\alpha) \times 100\%$ confidence ellipsoid for $\psi$. Through duality, these ellipsoids also make possible level $\alpha$ hypothesis tests for

$$\mathrm{H}_0 : \boldsymbol{\psi} = \boldsymbol{\psi}_0. \tag{39}$$

Furthermore, one can enunciate

**Theorem 4** (Scheffé's Theorem)**.**

$$\mathbb{P}\left[\bigcap_{\mathbf{d}\in\mathbb{R}^s} |\mathbf{d}'\boldsymbol{\psi} - \mathbf{d}'\widetilde{\boldsymbol{\psi}}| = \sqrt{x_{s,1-\alpha}\mathbf{d}'\left(\mathbf{A}\widetilde{\mathbf{K}}\mathbf{A}'\right)\mathbf{d}}\right] \approx 1 - \alpha.$$

*Proof.* A point is contained in an ellipsoid only and only if (see [**?**], pg. ??) it is between every two parallel plans tangent to the ellipsoid, each pair of tangent parallel plans generated by a vector $\mathbf{d} \neq \mathbf{0}$, *i. e.*,

$$|\mathbf{d}'\boldsymbol{\psi} - \mathbf{d}'\widetilde{\boldsymbol{\psi}}| \leq \sqrt{x_{s,1-\alpha}\mathbf{d}'\left(\mathbf{A}\widetilde{\mathbf{K}}\mathbf{A}'\right)\mathbf{d}},$$

which, along with lemma 3, establishes the thesis. $\square$

# References


[1] T. Calinski and S. Kageyama. *Block designs: a randomization approach. Vol. 1: Analysis.* Springer, New York, 2000.

[2] S. Gnot, J. Kleffe, and R. Zmyślony. Nonnegativity of admissible invariant quadratic estimates in mixed linear models with two variance components. *J. Stat. Plann. Inference*, 12:249–258, 1985.

[3] S. Gnot, W. Klonecki, and R. Zmyślony. Linear spaces and the theory of best linear unbiased estimation. *Bull. Acad. Pol. Sci., Sér. Sci. Math. Astron. Phys.*, (26):69—72, 1978.

[4] Stanisław Gnot, Jolanta Srzednicka, and Roman Zmyślony. Bayes invariant quadratic estimation in general linear regression models. *J. Stat. Plann. Inference*, 30(2):223–236, 1992.

[5] T. Kariya and H. Kurata. *Generalized least squares.* John Wiley & Sons, Chichester, 2004.

[6] L. R. LaMotte. A canonical form for the general linear model. *Ann. Stat.*, (5):787—789, 1977.

[7] J. A. Nelder. The analysis of randomized experiments with orthogonal block structure. I: Block structure and the null analysis of variance. *Proc. R. Soc. Lond., Ser. A*, (283):147—162, 1965.

[8] J. A. Nelder. The analysis of randomized experiments with orthogonal block structure. II: Treatment structure and the general analysis of variance. *Proc. R. Soc. Lond., Ser. A*, (283):163—178, 1965.

[9] J. R. Schott. *Matrix analysis for statistics.* Wiley, New York, 1997.

[10] S. R. Searle, G. Casella, and C. E. MacCulloch. *Variance components.* Wiley, New York, 1992.

[11] R. Zmyslony. A characterization of best linear unbiased estimators in the general linear model. *Mathematical statistics and probability theory, Proc. 6th int. Conf., Wisla/Pol. 1978, Lect. Notes Stat.*, (2):365—373, 1980.

[12] Roman Zmyślony and Hilmar Drygas. Jordan algebras and Bayesian quadratic estimation of variance components. *Linear Algebra Appl.*, 168:259–275, 1992.