Explicit estimators in unbalanced mixed linear models with applications to SAE

Dietrich von Rosen[1,2], Tatjana von Rosen[3],

[1] Department of Energy and Technology, Swedish University of Agricultural Sciences

[2] Mathematical Statistics, Linköping University, Sweden

[3] Department of Statistics, Stockholm University, Sweden

Outline

- Two-way model with unequal number of observations
- Test of variance components
- Mixed linear model with two variance components
- Estimation of fixed effects

Area/Time model

Area	Time 1	2	3
1	5	6	5
2	4	2	7
3	7	5	3

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk},$$

$$i = 1, 2, \ldots, r$$
$$j = 1, 2, \ldots, s$$
$$k = 1, 2, \ldots, n_{ij}$$

Two-way model with unequal number of observations

$$y^s_{ijk} = \mu_{ijk} + \epsilon^s_{ijk}, \qquad \epsilon^s_{ijk} \sim N(0, \sigma^2_s) \qquad i = 1, 2, \ldots, r;\ j = 1, 2, \ldots, s;\ k = 1, 2, \ldots, n_{ij}$$

$$\mu_{ijk} = \theta_{ijk} + u_{ijk}, \qquad u_{ijk} \sim N(0, \sigma^2_u) \qquad i = 1, 2, \ldots, r;\ j = 1, 2, \ldots, s;\ k = 1, 2, \ldots, n_{ij}$$

$$\theta_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + (x_{ijk} - \overline{x}_{ij\bullet})\delta_{ij}$$

Two-way model with unequal number of observations

$$\theta_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + (x_{ijk} - \overline{x}_{ij\bullet})\delta_{ij}$$

Area effects are fixed:

$$\mu, \quad \alpha_i, \quad (x_{ijk} - \overline{x}_{ij\bullet})\delta_{ij}$$

Time effects are random:

$$\boldsymbol{\beta} = (\beta_j) \sim N_s(\mathbf{0}, \sigma_\beta^2 \boldsymbol{I}_s), \quad \boldsymbol{\gamma} = (\gamma_{ij}) \sim N_{rs}(\mathbf{0}, \sigma_\gamma^2 \boldsymbol{I}_{rs})$$

Two-way model with unequal number of observations

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + (x_{ijk} - \overline{x}_{ij\bullet})\delta_{ij} + \epsilon_{ijk}, \qquad i = 1, 2, \ldots, r$$

$$\epsilon_{ijk} = u_{ijk} + \epsilon^s_{ijk} \qquad j = 1, 2, \ldots, s$$

$$k = 1, 2, \ldots, n_{ij}$$

$$\boldsymbol{\beta} = (\beta_j) \sim N_s(\mathbf{0}, \sigma^2_\beta \boldsymbol{I}_s)$$

$$\boldsymbol{\gamma} = (\gamma_{ij}) \sim N_s(\mathbf{0}, \sigma^2_\gamma \boldsymbol{I}_{rs})$$

$$\boldsymbol{\epsilon} = (\epsilon_{ijk}) \sim N_N(\mathbf{0}, \sigma^2 \boldsymbol{I}_N), \qquad N = \sum_{ij} n_{ij}$$

Two-way model with unequal number of observations

Define $\boldsymbol{H}$ as an averaging operator on $\boldsymbol{y} = (y_{ijk})$ and put

$$\boldsymbol{z} = (\bar{y}_{ij\bullet}) = \boldsymbol{H}'\boldsymbol{y}, \quad \boldsymbol{z} : rs \times 1.$$

There exists a generator of $\mathcal{C}(\boldsymbol{H})^{\perp}$, say $\boldsymbol{H}^o$. Let $\boldsymbol{R}_x = (x_{ijk} - \overline{x}_{ij\bullet})$ and $\boldsymbol{\delta} = (\delta_{ij})$. Then,

$$\boldsymbol{H}^{o'}\boldsymbol{y} \sim N_{N-rs}(\boldsymbol{H}^{o'}\boldsymbol{R}_x\boldsymbol{\delta}, \sigma^2\boldsymbol{H}^{o'}\boldsymbol{H}^o) \quad \perp \ \boldsymbol{z} = \boldsymbol{H}'\boldsymbol{y}.$$

Moreover,

$$\begin{aligned} \boldsymbol{z} &= \mu\mathbf{1}_s \otimes \mathbf{1}_r + (\boldsymbol{I}_r \otimes \mathbf{1}_s)\boldsymbol{\alpha} + (\mathbf{1}_r \otimes \boldsymbol{I}_s)\boldsymbol{\beta} + (\boldsymbol{I}_r \otimes \boldsymbol{I}_s)\boldsymbol{\gamma} + \bar{\boldsymbol{\epsilon}} \\ &= \mu\mathbf{1}_{rs} + (\boldsymbol{I}_r \otimes \mathbf{1}_s)\boldsymbol{\alpha} + (\mathbf{1}_r \otimes \boldsymbol{I}_s)\boldsymbol{\beta} + \boldsymbol{\gamma} + \bar{\boldsymbol{\epsilon}}, \\ & \bar{\boldsymbol{\epsilon}} = \boldsymbol{H}'\boldsymbol{\epsilon} \sim N_{rs}(\mathbf{0}, \sigma^2\boldsymbol{H}'\boldsymbol{H}). \end{aligned}$$

Decomposition of linear spaces

$$\boldsymbol{z} = \mu \mathbf{1}_{rs} + (\boldsymbol{I}_r \otimes \mathbf{1}_s)\boldsymbol{\alpha} + (\mathbf{1}_r \otimes \boldsymbol{I}_s)\boldsymbol{\beta} + \boldsymbol{\gamma} + \bar{\boldsymbol{\epsilon}}.$$

First step: $\mathcal{C}(\boldsymbol{H}) \boxplus \mathcal{C}(\boldsymbol{H})^{\perp}$.

Moreover,

$$\mathcal{C}(\boldsymbol{H}) = \mathcal{C}(\boldsymbol{H}\boldsymbol{B}_1) \oplus \mathcal{C}(\boldsymbol{H}\boldsymbol{B}_2) \oplus \mathcal{C}(\boldsymbol{H}\boldsymbol{B}_3),$$

where

$$\begin{aligned} \boldsymbol{B}_1 &= (\boldsymbol{I}_r \otimes \mathbf{1}_s)s^{-1/2} \\ \boldsymbol{B}_2 &= (\mathbf{1}_r \otimes \mathbf{1}_s^o)(\mathbf{1}_s^{o\prime}\mathbf{1}_s^o)^{-1/2}r^{-1/2} \\ \boldsymbol{B}_3 &= (\mathbf{1}_r^o \otimes \mathbf{1}_s^o)((\mathbf{1}_r^{o\prime}\mathbf{1}_r^o)^{-1/2} \otimes (\mathbf{1}_s^{o\prime}\mathbf{1}_s^o)^{-1/2}) \end{aligned}$$

and $\boldsymbol{B}_1'\boldsymbol{B}_1 = \boldsymbol{I}_r$, $\boldsymbol{B}_2'\boldsymbol{B}_2 = \boldsymbol{I}_{s-1}$, $\boldsymbol{B}_3'\boldsymbol{B}_3 = \boldsymbol{I}_{(r-1)(s-1)}$.

Manipulations

It follows that

$$D\left[\begin{pmatrix} \boldsymbol{B}_2'\boldsymbol{z} \\ \boldsymbol{B}_3'\boldsymbol{z} \end{pmatrix}\right] = \text{diag}(r\sigma_\beta^2 \boldsymbol{I}_{s-1} + \sigma_\gamma^2 \boldsymbol{I}_{s-1}, \sigma_\gamma^2 \boldsymbol{I}_{(r-1)(s-1)}) + \sigma^2 \boldsymbol{L},$$

where

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{B}_2' \\ \boldsymbol{B}_3' \end{pmatrix} \boldsymbol{H}\boldsymbol{H}' \begin{pmatrix} \boldsymbol{B}_2 & \boldsymbol{B}_3 \end{pmatrix}$$

Manipulations

Main idea

Solve the following equation in $\boldsymbol{u} = (\boldsymbol{u}_1', \boldsymbol{u}_2')'$

$$D[\begin{pmatrix} \boldsymbol{B}_2'\boldsymbol{z} \\ \boldsymbol{B}_3'\boldsymbol{z} \end{pmatrix} + \boldsymbol{u}] = \text{diag}(r\sigma_\beta^2 \boldsymbol{I}_{s-1} + \sigma_\gamma^2 \boldsymbol{I}_{s-1}, \sigma_\gamma^2 \boldsymbol{I}_{(r-1)(s-1)}) + c\sigma^2 \boldsymbol{I}_{r(s-1)},$$

where c is a constant and $\boldsymbol{u}$ is independent of $\boldsymbol{z}$ and normally distributed.

Why is this important?

Manipulations

Solution

Since $(\boldsymbol{H}^{o'}\boldsymbol{H}^{o})^{-1/2}\boldsymbol{H}^{o'}\boldsymbol{y} \sim N_{N-rs}((\boldsymbol{H}^{o'}\boldsymbol{H}^{o})^{-1/2}\boldsymbol{H}^{o'}\boldsymbol{R}_x\boldsymbol{\delta}, \sigma^2\boldsymbol{I}_{N-rs})$ we may regress and define the residual $\boldsymbol{R}$ so that

$$\boldsymbol{R} \sim N_{N-rs-t}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_{N-rs-t}), \qquad t = r(\boldsymbol{H}^{o'}\boldsymbol{R}_x).$$

Let $\boldsymbol{v}$ be $r(s-1)$ variables drawn from $\boldsymbol{R}$ and let $\lambda_{\max}$ be the largest eigenvalue of $\boldsymbol{L}$. Then,

$$\boldsymbol{u} = (\lambda_{\max}\boldsymbol{I} - \boldsymbol{L})^{1/2}\boldsymbol{v}$$

solves

$$D[\begin{pmatrix} \boldsymbol{B}_2'\boldsymbol{z} \\ \boldsymbol{B}_3'\boldsymbol{z} \end{pmatrix} + \boldsymbol{u}] = \text{diag}(r\sigma_\beta^2\boldsymbol{I}_{s-1} + \sigma_\gamma^2\boldsymbol{I}_{s-1}, \sigma_\gamma^2\boldsymbol{I}_{(r-1)(s-1)}) + c\sigma^2\boldsymbol{I}_{r(s-1)}.$$

Exact tests

Let

$$\begin{aligned}
\boldsymbol{w}_1 &= \boldsymbol{B}_2'\boldsymbol{z} + \boldsymbol{u}_1 \sim N_{s-1}(\boldsymbol{0}, r\sigma_\beta^2 \boldsymbol{I}_{s-1} + \sigma_\gamma^2 \boldsymbol{I}_{s-1} + \lambda_{\max}\sigma^2 \boldsymbol{I}_{s-1}) \\
\boldsymbol{w}_2 &= \boldsymbol{B}_3'\boldsymbol{z} + \boldsymbol{u}_2 \sim N_{(r-1)(s-1)}(\boldsymbol{0}, \sigma_\gamma^2 \boldsymbol{I}_{(r-1)(s-1)} + \lambda_{\max}\sigma^2 \boldsymbol{I}_{(r-1)(s-1)})
\end{aligned}$$

Exact tests

Then,

$$\frac{\boldsymbol{w}_1'\boldsymbol{w}_1}{r\sigma_\beta^2 + \sigma_\gamma^2 + \lambda_{\max}\sigma^2} \sim \chi^2(s-1)$$

$$\frac{\boldsymbol{w}_2'\boldsymbol{w}_2}{\sigma_\gamma^2 + \lambda_{\max}\sigma^2} \sim \chi^2((r-1)(s-1))$$

and to test $H_0 : \sigma_\beta^2 = 0$ one can use

$$\frac{\boldsymbol{w}_1'\boldsymbol{w}_1/(s-1)}{\boldsymbol{w}_2'\boldsymbol{w}_2/(r-1)(s-1)} \sim F(s-1,(r-1)(s-1)).$$

Exact tests

To test $H_0 : \sigma_\gamma^2 = 0$ one can use

$$\frac{\boldsymbol{w}_2'\boldsymbol{w}_2/(r-1)(s-1)}{\lambda_{\max}\boldsymbol{v}_2'\boldsymbol{v}_2/(N-2rs+r)} \sim F((r-1)(s-1), N-2rs+r),$$

where $\boldsymbol{v}_2$ are $N-2rs+r$ variables from $\boldsymbol{R}$.

Some references

- Wald (1941), Spjøtvoll (1968), Thomsen (1975) presented exact tests for ratios of VC in one-way and two-way cross-classification models without interactions.
- Khuri (1986, 1987, 1990), Khuri & Little (1987), Gallo & Khuri (1990) constructed exact tests concerning the VC of the random effects and estimable linear functions of the fixed effects in an unbalanced mixed two-way cross-classification with interaction model.
- Öfversten (1993, 1995) presented methods for obtaining exact F-tests of VC for models with one random factor, models with nested classifications, and models with interaction between two random factors.
- Christensen (1996) commented on Öfversten's results and focused on linear spaces.
- Others: Seifert (1992), Mathew & Sinha (1992).

Estimation of fixed effects

Assume the model:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\gamma} + \epsilon,$$

where $\boldsymbol{\beta}$ is to be estimated and $\boldsymbol{y}$: $n \times 1$, the random vector, $\boldsymbol{X}$: $n \times k$, $\boldsymbol{Z}$: $n \times l$ are known matrices both supposed to be of full column rank, $\mathcal{C}(\boldsymbol{X}) \subseteq \mathcal{C}(\boldsymbol{Z})$, $\boldsymbol{\gamma} \sim N_l(\boldsymbol{0}, \sigma_\gamma^2 \boldsymbol{I}_l)$, and $\boldsymbol{\epsilon} \sim N_n(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n)$, where the variances are unknown.

Let $\boldsymbol{C}$: $n \times t$ be of full rank such that $\mathcal{C}(\boldsymbol{C}) = \mathcal{C}(\boldsymbol{X})^\perp \cap \mathcal{C}(\boldsymbol{Z})$.

We have $\mathcal{C}(\boldsymbol{Z}) = \mathcal{C}(\boldsymbol{X}) \boxplus \mathcal{C}(\boldsymbol{C})$, and

$$\boldsymbol{Z} = (\boldsymbol{X} : \boldsymbol{C})\boldsymbol{Q} = \boldsymbol{X}\boldsymbol{Q}_1 + \boldsymbol{C}\boldsymbol{Q}_2,$$

for some $\boldsymbol{Q}' = (\boldsymbol{Q}_1' : \boldsymbol{Q}_2')$: $l \times k$: $l \times t$.

Estimation

A canonical form of the model can be obtained as follows: Let

$$\boldsymbol{B}_1 = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1},$$

$$\boldsymbol{B}_2 = \boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1},$$

$$\boldsymbol{B}_3 \quad \text{be any matrix such that} \quad \mathcal{C}(\boldsymbol{B}_3) = \mathcal{C}(\boldsymbol{Z})^{\perp}.$$

Observe

$$R^n = \mathcal{C}(\boldsymbol{B}_1) \boxplus \mathcal{C}(\boldsymbol{B}_2) \boxplus \mathcal{C}(\boldsymbol{B}_3).$$

Then,

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\gamma} + \epsilon$$

is equivalent to

Estimation: $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \epsilon$

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{XQ}_1\gamma + \boldsymbol{CQ}_2\gamma + \epsilon$$

and we get that a 1-1 transformation of the model equals ($\boldsymbol{B}_1 = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$, $\boldsymbol{B}_2 = \boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1}$, $\mathcal{C}(\boldsymbol{B}_3) = \mathcal{C}(\boldsymbol{Z})^{\perp}$)

$$\boldsymbol{u}_1 := \boldsymbol{B}_1'\boldsymbol{y} = \boldsymbol{\beta} + \boldsymbol{Q}_1\gamma + \boldsymbol{B}_1'\epsilon, \quad (1)$$

$$\boldsymbol{B}_2'\boldsymbol{y} = \boldsymbol{Q}_2\gamma + \boldsymbol{B}_2'\epsilon, \quad (2)$$

$$\boldsymbol{B}_3'\boldsymbol{y} = \boldsymbol{B}_3'\epsilon. \quad (3)$$

(2) is equivalent to

$$\boldsymbol{v}_2 := (\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{B}_2'\boldsymbol{y} = (\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{Q}_2\gamma + (\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{B}_2'\epsilon \quad (4)$$

.

Estimation

Observe that

$$D[\boldsymbol{v}_2] = \sigma_\gamma^2 \boldsymbol{I}_t + (\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{B}_2'\boldsymbol{B}_2(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}.$$

In the next we perform the important step to add observations from (3) to (4) so that a diagonal dispersion is obtained, i.e.

$$\begin{aligned} \boldsymbol{u}_2 &:= (\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{B}_2'\boldsymbol{y} \\ &+ (\lambda \boldsymbol{I}_t - (\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{B}_2'\boldsymbol{B}_2(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2})^{1/2} sel(\boldsymbol{B}_3'\boldsymbol{y}, t), \end{aligned}$$

where the selection operator $sel(\boldsymbol{B}_3'\boldsymbol{y}, t)$ selects t observations from $\boldsymbol{B}_3'\boldsymbol{y} \sim N_{n-l}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_{n-l})$ and λ is so large that the square root exists.

Estimation

Now,

$$\begin{aligned} E[\boldsymbol{u}_2] &= \mathbf{0}, \\ D[\boldsymbol{u}_2] &= \sigma_\gamma^2 \boldsymbol{I}_t + \lambda\sigma^2 \boldsymbol{I}_t \end{aligned}$$

and $\boldsymbol{u}_2$ is normally distributed.

It is clear that if we would like to utilize $\boldsymbol{u}_2$ the parameter λ should be chosen as small as possible because the dispersion of $\boldsymbol{u}_2$ is proportional to λ.

Moreover, for $\boldsymbol{u}_1 = \boldsymbol{B}_1'\boldsymbol{y} = \boldsymbol{\beta} + \boldsymbol{Q}_1\boldsymbol{\gamma} + \boldsymbol{B}_1'\epsilon$,

$$C[\boldsymbol{u}_1, \boldsymbol{u}_2] = C[\boldsymbol{Q}_1\boldsymbol{\gamma}, (\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{Q}_2\boldsymbol{\gamma}] = \sigma_\gamma^2\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}.$$

Estimation

Below the main action takes place, namely condition $\boldsymbol{u}_1$ with $\boldsymbol{u}_2$ and we obtain

$$
\begin{aligned}
E[\boldsymbol{u}_1|\boldsymbol{u}_2] &= E[\boldsymbol{u}_1] + C[\boldsymbol{u}_1,\boldsymbol{u}_2]D[\boldsymbol{u}_2]^{-1}\boldsymbol{u}_2 \\
&= \boldsymbol{\beta} + \frac{\sigma_\gamma^2}{\sigma_\gamma^2+\lambda\sigma^2}\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2.
\end{aligned}
$$

Thus, with known variances

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}} &= \boldsymbol{u}_1 - \frac{\sigma_\gamma^2}{\sigma_\gamma^2+\lambda\sigma^2}\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2 \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - \frac{\sigma_\gamma^2}{\sigma_\gamma^2+\lambda\sigma^2}\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2
\end{aligned}
$$

Estimation

Finally we estimate the ratio of the variances. Note that

$$\frac{\sigma_\gamma^2}{\sigma_\gamma^2 + \lambda\sigma^2} = 1 - \frac{\lambda\sigma^2}{\sigma_\gamma^2 + \lambda\sigma^2},$$

and ($\boldsymbol{u}_2 \sim N_t(\mathbf{0}, (\sigma_\gamma^2 + \lambda\sigma^2)\boldsymbol{I}_t)$, $\boldsymbol{u}_3 \sim N_{n-t-l}(\mathbf{0}, \sigma^2\boldsymbol{I}_{n-t-l})$)

$$E[(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}] = \frac{1}{t-2}(\sigma_\gamma^2 + \lambda\sigma^2)^{-1}, \quad t > 2$$
$$E[\boldsymbol{u}_3'\boldsymbol{u}_3] = (n-t-l)\sigma^2.$$

Because $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ are independent we immediately find an estimator of $\frac{\sigma^2}{\sigma_\gamma^2 + \lambda\sigma^2}$, and thus $\boldsymbol{\beta}$ has been estimated.

Estimation

The expression equals

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{u}_1 - \boldsymbol{f} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{f},$$

where

$$\boldsymbol{f} = (1 - c\boldsymbol{u}_3'\boldsymbol{u}_3(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1})\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2$$

and

$$c = \lambda\frac{t-2}{n-t-l}.$$

Properties

Before going to calculate the mean and dispersion of the estimator we need to know

$$E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2'],$$
$$E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2'].$$

Since $\boldsymbol{\Gamma}\boldsymbol{u}_2$ has the same distribution as $\boldsymbol{u}_2$ for any orthogonal matrix $\boldsymbol{\Gamma}$ we have that $\boldsymbol{\Gamma}E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2']\boldsymbol{\Gamma}'$ equals $E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2']$. Thus

$$E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2'] = c_0\boldsymbol{I}_t$$

and taking the trace on both sides implies that $c_0 = 1/t$.

Properties

Hence,

$$E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2'] = \boldsymbol{I}_t/t.$$

Similarly

$$E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2'] = c_0\boldsymbol{I},$$

$$c_o = E[(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}]/t = \frac{1}{t(t-2)}(\sigma_\gamma^2 + \lambda\sigma^2)^{-1}.$$

Properties $\boldsymbol{\beta} = \boldsymbol{u}_1 - \boldsymbol{f} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{f}$

Now the mean and dispersion for $\widehat{\boldsymbol{\beta}}$ is derived. Since $E[\boldsymbol{u}_2] = \mathbf{0}$ and $E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}] = \mathbf{0}$ it follows that $\widehat{\boldsymbol{\beta}}$ is unbiased, i.e.

$$E[\widehat{\boldsymbol{\beta}}] = \boldsymbol{\beta}.$$

Moreover, we have

$$D[\widehat{\boldsymbol{\beta}}] = D[\boldsymbol{u}_1] + D[\boldsymbol{f}] - C[\boldsymbol{u}_1, \boldsymbol{f}] - C[\boldsymbol{f}, \boldsymbol{u}_1]$$

and the terms in the right hand side will be considered separately.

Properties

$$D[\boldsymbol{f}]$$
$$= E[(1 - c\boldsymbol{u}_3'\boldsymbol{u}_3(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1})^2\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2\boldsymbol{u}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{Q}_2\boldsymbol{Q}_1'].$$

Hence we have to perform the following calculation:

$$E[\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2\boldsymbol{u}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{Q}_2\boldsymbol{Q}_1']$$
$$= (\sigma_\gamma^2 + \lambda\sigma^2)\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1',$$

$$E[c\boldsymbol{u}_3'\boldsymbol{u}_3(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2\boldsymbol{u}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{Q}_2\boldsymbol{Q}_1']$$
$$= c(n-t-l)\sigma^2 E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2']\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'$$
$$= c\frac{(n-t-l)\sigma^2}{t}\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1',$$

Properties

$$E[c^2(\boldsymbol{u}_3'\boldsymbol{u}_3)^2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-2}\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2\boldsymbol{u}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{Q}_2\boldsymbol{Q}_1']$$
$$= c^2 2(n-t-l)(\sigma^2)^2 E[\boldsymbol{u}_2(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1}\boldsymbol{u}_2']\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'$$
$$= c^2\frac{2(n-t-l)(\sigma^2)^2}{t(t-2)(\sigma_\gamma^2+\lambda\sigma^2)}\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'.$$

$D[\widehat{\boldsymbol{\beta}}] = D[\boldsymbol{u}_1] + D[\boldsymbol{f}] - C[\boldsymbol{u}_1, \boldsymbol{f}] - C[\boldsymbol{f}, \boldsymbol{u}_1]$

Hence,

$$
\begin{aligned}
& D[\boldsymbol{f}] \\
& = ((\sigma_\gamma^2 + \lambda\sigma^2) - 2c\frac{(n-t-l)\sigma^2}{t} + c^2\frac{2(n-t-l)(\sigma^2)^2}{t(t-2)(\sigma_\gamma^2 + \lambda\sigma^2)}) \\
& \qquad \times \boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'. \qquad (5)
\end{aligned}
$$

Now $C[\boldsymbol{f}, \boldsymbol{u}_1]$ is calculated via conditioning $\boldsymbol{u}_1|\boldsymbol{u}_2$:

$$
C[\boldsymbol{f}, \boldsymbol{u}_1] = E_{u_2}E_{u_1|u_2}E_{u_3}[(1 - c\boldsymbol{u}_3'\boldsymbol{u}_3(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1})\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2\boldsymbol{u}_1']
$$

Properties

$$
\begin{aligned}
& C[\boldsymbol{f}, \boldsymbol{u}_1] \\
&= E_{u_2} E_{u_1|u_2} E_{u_3}[(1 - c\boldsymbol{u}_3'\boldsymbol{u}_3(\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1})\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2\boldsymbol{u}_1'] \\
&= E_{u_2}[(1 - cE[\boldsymbol{u}_3'\boldsymbol{u}_3](\boldsymbol{u}_2'\boldsymbol{u}_2)^{-1})\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{u}_2\boldsymbol{u}_2' \\
&\quad \times(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1/2}\boldsymbol{Q}_2\boldsymbol{Q}_1']\sigma_\gamma^2/(\sigma_\gamma^2 + \lambda\sigma^2)^{-1} \\
&= (\sigma_\gamma^2 - c\frac{\sigma_\gamma^2\sigma^2(n-t-l)}{t(\sigma_\gamma^2 + \lambda\sigma^2)})\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'. \qquad (6)
\end{aligned}
$$

Thus, from (5), (6) and $D[\boldsymbol{u}_1] = \sigma_\gamma\boldsymbol{Q}_1\boldsymbol{Q}_1' + \sigma^2\boldsymbol{B}_1'\boldsymbol{B}_1$, $D[\widehat{\boldsymbol{\beta}}]$ is obtained, i.e.

$$
D[\widehat{\boldsymbol{\beta}}] = D[\boldsymbol{u}_1] + D[\boldsymbol{f}] - 2C[\boldsymbol{u}_1, \boldsymbol{f}]
$$

It remains to express $\boldsymbol{Q}_1\boldsymbol{Q}_1'$ and $\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'$.

Properties

It remains to express $\boldsymbol{Q}_1\boldsymbol{Q}_1'$ and $\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'$ in the original matrices. The expression follow however immediately from

$$\begin{pmatrix} (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' \\ (\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}' \end{pmatrix} \boldsymbol{Z}\boldsymbol{Z}' \begin{pmatrix} \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} & \boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1} \end{pmatrix}$$
$$= \begin{pmatrix} \boldsymbol{Q}_1\boldsymbol{Q}_1' & \boldsymbol{Q}_1\boldsymbol{Q}_2' \\ \boldsymbol{Q}_2\boldsymbol{Q}_1' & \boldsymbol{Q}_2\boldsymbol{Q}_2' \end{pmatrix}.$$

$$\boldsymbol{Q}_1\boldsymbol{Q}_1' = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{X}'(\boldsymbol{X}'\boldsymbol{X})^{-1},$$
$$\boldsymbol{Q}_1\boldsymbol{Q}_2'(\boldsymbol{Q}_2\boldsymbol{Q}_2')^{-1}\boldsymbol{Q}_2\boldsymbol{Q}_1'$$
$$= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{C})^{-1}\boldsymbol{C}'\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{X}'(\boldsymbol{X}'\boldsymbol{X})^{-1}.$$