of the block diagonal matrix on slide 21

End lecture 21
03-12-25

Thus, $\mathrm{Var}(\boldsymbol{y}) = \sigma_p^2 \boldsymbol{Z}\boldsymbol{Z}^\top + \sigma_e^2 \boldsymbol{I}$ is a block diagonal matrix.

The first block is

pot 1

$\mathrm{Var}(y_{111})$ $\mathrm{Cov}(y_{111}, y_{112})$

$$\mathrm{Var}\begin{bmatrix} y_{111} \\ y_{112} \\ y_{113} \end{bmatrix} = \begin{bmatrix} \sigma_p^2 + \sigma_e^2 & \sigma_p^2 & \sigma_p^2 \\ \sigma_p^2 & \sigma_p^2 + \sigma_e^2 & \sigma_p^2 \\ \sigma_p^2 & \sigma_p^2 & \sigma_p^2 + \sigma_e^2 \end{bmatrix}.$$

Structure of the first matrix on the diagonal of ZZ^T

$$(\sigma_p^2 + \sigma_e^2) = \sigma_e^2\left(\underline{\frac{\sigma_p^2}{\sigma_e^2} + 1}\right)$$

Note that

- $\text{Var}(y_{ijk}) = \sigma_p^2 + \sigma_e^2 \qquad \forall\, i, j, k.$
- $\text{Cov}(y_{ijk},\ y_{ijk^*}) = \sigma_p^2 \qquad \forall\, i, j,$ and $k \neq k^*.$
  plants within the same pot
- $\text{Cov}(y_{ijk},\ y_{i^*j^*k^*}) = 0 \qquad$ if $i \neq i^*$ or $j \neq j^*$.
  plants from different pots – independent
- Any two observations from the same pot have covariance σ_p^2.
- Any two observations from different pots are uncorrelated.

Alternative Derivation of Variances and Covariances

$$
\begin{aligned}
\mathrm{Var}(y_{ijk}) &= \mathrm{Var}(\overbrace{\mu + \alpha_i}^{\text{constant}} + \overbrace{p_{ij} + e_{ijk}}^{\text{random}}) = \mathrm{Var}(p_{ij} + e_{ijk}) \\
&\overset{\text{by assump.}}{=} \mathrm{Var}(p_{ij}) + \mathrm{Var}(e_{ijk}) + \underbrace{\mathrm{Cov}(p_{ij}, e_{ijk})}_{=0} + \underbrace{\mathrm{Cov}(e_{ijk}, p_{ij})}_{=0} \\
&= \sigma_p^2 + \sigma_e^2 + 0 + 0 = \boxed{\sigma_p^2 + \sigma_e^2.}
\end{aligned}
$$

For $k \neq k^*$,

$$
\begin{aligned}
\mathrm{Cov}(y_{ijk},\ y_{ijk^*}) &= \mathrm{Cov}(\cancel{\mu + \alpha_i} + p_{ij} + e_{ijk},\ \cancel{\mu + \alpha_i} + p_{ij} + e_{ijk^*}) \\
&= \mathrm{Cov}(p_{ij} + e_{ijk},\ p_{ij} + e_{ijk^*}) \\
&= \mathrm{Cov}(p_{ij},\ p_{ij}) + \mathrm{Cov}(p_{ij},\ e_{ijk^*}) \qquad = 0 \text{ by assumption} \\
&\quad + \mathrm{Cov}(e_{ijk}, p_{ij}) + \mathrm{Cov}(e_{ijk},\ e_{ijk^*}) \\
&= \mathrm{Var}(p_{ij}) + 0 + 0 + 0 = \sigma_p^2.
\end{aligned}
$$

- Note that $\mathrm{Var}(\boldsymbol{y})$ may be written as $\sigma_e^2 \boldsymbol{V}$ where $\boldsymbol{V}$ is a block diagonal matrix with blocks of the form

block element of V =

$$\begin{bmatrix} 1+\sigma_p^2/\sigma_e^2 & \sigma_p^2/\sigma_e^2 & \dots & \sigma_p^2/\sigma_e^2 \\ \sigma_p^2/\sigma_e^2 & 1+\sigma_p^2/\sigma_e^2 & \dots & \sigma_p^2/\sigma_e^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_p^2/\sigma_e^2 & \sigma_p^2/\sigma_e^2 & \dots & 1+\sigma_p^2/\sigma_e^2 \end{bmatrix}$$

- Thus, if σ_p^2/σ_e^2 were known, we would have the Aitken Model.

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \text{ where } \boldsymbol{\epsilon} = \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e} \sim \mathcal{N}(\boldsymbol{0}, \ \sigma^2\boldsymbol{V}), \ \sigma^2 \equiv \sigma_e^2.$$

- Thus, if σ_p^2/σ_e^2 were known, we would use GLS to estimate any estimable $\boldsymbol{C\beta}$ by $\boldsymbol{C}\widehat{\boldsymbol{\beta}}_{\boldsymbol{V}} = \boldsymbol{C}(\boldsymbol{X}^\top \boxed{\boldsymbol{V}^{-1}} \boldsymbol{X})^- \boldsymbol{X}^\top \boxed{\boldsymbol{V}^{-1}} \boldsymbol{y}$.

- However, we seldom know σ_p^2/σ_e^2 or, more generally, $\boldsymbol{\Sigma}$ or $\boldsymbol{V}$.

- For the general problem where $\mathrm{Var}(\boldsymbol{y}) = \boldsymbol{\Sigma}$ is an unknown positive definite matrix, we can rewrite $\boldsymbol{\Sigma}$ as $\sigma^2\boldsymbol{V}$, where σ^2 is an unknown positive variance and $\boldsymbol{V}$ is an unknown positive definite matrix.

- As in our simple example, each entry of $\boldsymbol{V}$ is usually assumed to be a known function of few unknown parameters.

- Thus, our strategy for estimating an estimable $C\beta$ involves estimating the unknown parameters in V to obtain

$$C\widehat{\beta}_{\widehat{V}} = C(X^\top \widehat{V}^{-1} X)^- X^\top \widehat{V}^{-1} y.$$

- In general,

$$C\widehat{\beta}_{\widehat{V}} = C(X^\top \widehat{V}^{-1} X)^- X^\top \widehat{V}^{-1} y$$

is an nonlinear estimator that is an approximation to

$$C\widehat{\beta}_{V} = C(X^\top V^{-1} X)^- X^\top V^{-1} y,$$

which would be the BLUE of $C\beta$ if V were known.

- In special cases, $\boldsymbol{C}\widehat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{V}}}$ may be a linear estimator.

- However, even for our simple example involving seedling height, $\boldsymbol{C}\widehat{\boldsymbol{\beta}}_{\widehat{\boldsymbol{V}}}$ is a nonlinear estimator of $\boldsymbol{C\beta}$ for

$$\begin{aligned}
\boldsymbol{C} = [1, 1, 0] \quad &\Longleftrightarrow \quad \boldsymbol{C\beta} = \mu + \alpha_1, \\
\boldsymbol{C} = [1, 0, 1] \quad &\Longleftrightarrow \quad \boldsymbol{C\beta} = \mu + \alpha_2, \text{ and} \\
\boldsymbol{C} = [0, 1, -1] \quad &\Longleftrightarrow \quad \boldsymbol{C\beta} = \alpha_1 - \alpha_2.
\end{aligned}$$

- Confidence intervals and tests for these estimable functions are not exact.

- In our simple example involving seedling height, there was only one random factor (pot).

- When there are m random factors, we can partition $\boldsymbol{Z}$ and $\boldsymbol{u}$ as

$$\boldsymbol{Z} = [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_m] \text{ and } \boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_m \end{bmatrix},$$

where $\boldsymbol{u}_j$ is the vector of random effects associated with factor j ($j = 1, \ldots, m$).

- We can write $\boldsymbol{Zu}$ as

$$[\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_m] \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_m \end{bmatrix} = \sum_{j=1}^{m} \boldsymbol{Z}_j \boldsymbol{u}_j.$$

- We often assume that all random effects (including random errors) are mutually independent and that the random effects associated with the jth random factor have variance σ_j^2 $(j = 1, \ldots, m)$. Under these assumptions,

$$\text{Var}(\boldsymbol{y}) = \boldsymbol{ZGZ}^\top + \boldsymbol{R} = \sum_{j=1}^{m} \underline{\sigma_j^2 \boldsymbol{Z}_j \boldsymbol{Z}_j^\top} + \sigma_e^2 \boldsymbol{I}.$$

Sum of variance components associated with the m random effects

Example 2

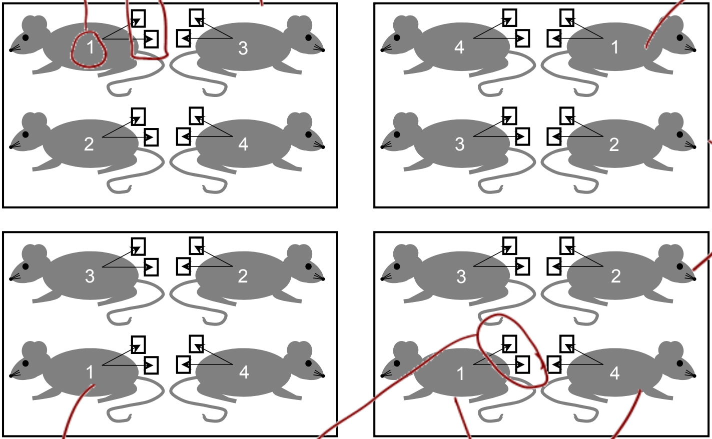
mice

- Consider an experiment involving 4 litters of 4 animals each.
- Suppose 4 treatments are randomly assigned to the 4 animals in each litter. litter effect
- Suppose we obtain $\boxed{\text{two replicate muscle}}$ samples from each animal and measure the response of interest for each muscle sample.

two measurements from the same mouse are more similar to each other
$\hat{=}$ mouse effect

2 replicate measurement from each mouse

litter 1



litter 2

3

indep.

4

litter effect + mouse effect

litter effect σ_e^2

Let y_{ijk} denote the kth measure of the response for the animal from litter j that received treatment i
$(\underbrace{i = 1, 2, 3, 4}_{\text{trt}}; \underbrace{j = 1, 2, 3, 4}_{\text{litter}}; \underbrace{k = 1, 2}_{\text{repl.}})$. Suppose

random litter + random mouse effect

$$y_{ijk} = \mu + \tau_i + \boxed{\ell_j + a_{ij}} + e_{ijk},$$

where $\boldsymbol{\beta} = [\mu, \tau_1, \tau_2, \tau_3, \tau_4]^\top \in \mathbb{R}^5$ is an unknown vector of fixed parameters,

$$\boldsymbol{u} = [\ell_1, \ell_2, \ell_3, \ell_4, a_{11}, a_{21}, a_{31}, a_{41}, a_{12}, \ldots, a_{34}, a_{44}]^\top$$

is a vector of random effects, and

$$\boldsymbol{e} = [e_{111}, e_{112}, e_{211}, e_{212}, e_{311}, e_{312}, e_{411}, e_{412}, \ldots, e_{441}, e_{442}]^\top$$

is a vector of random errors.

Obs. are ordered by repl. within treatment within litter

With

$$\underline{\boldsymbol{y}} = [\underline{y_{111}, y_{112}, y_{211}, y_{212}, y_{311}, y_{312}, y_{411}, y_{412}}, \ldots, y_{441}, y_{442}]^\top,$$

we can write the model as a linear mixed-effects model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e},$$

where

y_{ijk} k - repl. i = trt
j = litter

4 litters 16 mice

litter 1

litters 2-4

$$\boldsymbol{X} = \begin{array}{ccccc} \mu & \tau_1 & \tau_2 & \tau_3 & \tau_4 \end{array}
\begin{bmatrix}
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 \\
\multicolumn{5}{c}{\text{The matrix above repeated three more times.}}
\end{bmatrix}, \quad
\boldsymbol{Z} = \begin{bmatrix}
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 1
\end{bmatrix}$$

$$
\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \otimes \boldsymbol{B} \tag{1}
$$

$$
\tag{2}
$$

$$
= \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix} \tag{3}
$$

$$I_{4\times4} \otimes 1_{2\times1} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Big\}$$ two obs. per mouse

We can write less and be more precise using Kronecker product notation.

8 obs. for 4 mice within each litter

$$\underset{32\times5}{\underline{\underline{\boldsymbol{X}}}} = \underset{4\times1}{\mathbf{1}} \otimes \Big[\overset{\mu}{\underset{8\times1}{\mathbf{1}}}, \underset{4\times4}{\boldsymbol{I}} \otimes \underset{2\times1}{\mathbf{1}}\Big], \quad \boldsymbol{Z} = \Big[\underbrace{\underset{4\times4}{\boldsymbol{I}} \otimes \underset{8\times1}{\mathbf{1}}}_{\text{4 litters}}, \underset{16\times16}{\boldsymbol{I}} \otimes \underset{2\times1}{\mathbf{1}}\Big].$$

2 obs. for 16 mice

In this experiment, we have two random factors: litter and animal.

We can partition our random effects vector $\boldsymbol{u}$ into a vector of litter effects and a vector of animal effects:

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{\ell} \\ \boldsymbol{a} \end{bmatrix}, \quad \boldsymbol{\ell} = \begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix}, \quad \boldsymbol{a} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \\ a_{12} \\ \vdots \\ a_{44} \end{bmatrix}.$$

We make the usual assumption that

ℓ & a are independent

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{\ell} \\ \boldsymbol{a} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_\ell^2 \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \sigma_a^2 \boldsymbol{I} \end{bmatrix} \right),$$

where $\sigma_\ell^2, \sigma_a^2 \in \mathbb{R}^+$ are unknown parameters.

We can partition

$$\begin{aligned} \boldsymbol{Z} &= [\underset{4\times 4}{\boldsymbol{I}} \otimes \underset{8\times 1}{\boldsymbol{1}}, \underset{16\times 16}{\boldsymbol{I}} \otimes \underset{2\times 1}{\boldsymbol{1}}] \\ &= [\boldsymbol{Z}_\ell, \boldsymbol{Z}_a]. \end{aligned}$$

We have

$$\begin{aligned} \boldsymbol{Z}\boldsymbol{u} &= [\boldsymbol{Z}_\ell, \boldsymbol{Z}_a] \begin{bmatrix} \boldsymbol{\ell} \\ \boldsymbol{a} \end{bmatrix} \\ &= \boldsymbol{Z}_\ell \boldsymbol{\ell} + \boldsymbol{Z}_a \boldsymbol{a} \end{aligned}$$

and

end lecture 22

3-14-25

$$
\begin{aligned}
\mathrm{Var}(\boldsymbol{Z}\boldsymbol{u}) &= \underline{\boldsymbol{Z}\boldsymbol{G}\boldsymbol{Z}^\top} \\
&= [\boldsymbol{Z}_\ell, \boldsymbol{Z}_a] \begin{bmatrix} \sigma_\ell^2 \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_a^2 \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{Z}_\ell^\top \\ \boldsymbol{Z}_a^\top \end{bmatrix} \\
&= \boldsymbol{Z}_\ell(\sigma_\ell^2 \boldsymbol{I})\boldsymbol{Z}_\ell^\top + \boldsymbol{Z}_a(\sigma_a^2 \boldsymbol{I})\boldsymbol{Z}_a^\top \\
&= \sigma_\ell^2 \boldsymbol{Z}_\ell \boldsymbol{Z}_\ell^\top + \sigma_a^2 \boldsymbol{Z}_a \boldsymbol{Z}_a^\top \\
&= \sigma_\ell^2 \underset{4\times 4}{\boldsymbol{I}} \otimes \underset{8\times 8}{\boldsymbol{1}\boldsymbol{1}^\top} + \sigma_a^2 \underset{16\times 16}{\boldsymbol{I}} \otimes \underset{2\times 2}{\boldsymbol{1}\boldsymbol{1}^\top}.
\end{aligned}
$$