STK3100 Exercises, Week 12

Vinnie Ko, Jonas Moss

November 15, 2018

Exercise 9.7

The conditonal variance is

$$
\begin{aligned}
\mathrm{Var}(\boldsymbol{y}_i \mid \boldsymbol{X}_i) &= \mathrm{Var}\left(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{u}_i + \boldsymbol{\epsilon}_i \mid \boldsymbol{X}_i\right) \\
&= \boldsymbol{Z}_i \mathrm{Var}\left(\boldsymbol{u}_i\right) \boldsymbol{Z}_i^{\mathrm{T}} + \mathrm{Var}\left(\boldsymbol{\epsilon}_i\right) \\
&= \boldsymbol{Z}_i \boldsymbol{\Sigma}_u \boldsymbol{Z}_i^{\mathrm{T}} + \sigma_\epsilon^2 \boldsymbol{I}
\end{aligned}
$$

Exercise 9.8

Assume that $\mathrm{Cov}(\epsilon_{i,j}, \epsilon_{i,k}) = \sigma_\epsilon^2 \rho^{|j-k|}$, then

$$
\begin{aligned}
\mathrm{Cor}(y_{i,j}, y_{i,k}) &= \frac{\mathrm{Cov}(y_{i,j}, y_{i,k})}{\sqrt{\mathrm{Var}(y_{i,j})\mathrm{Var}(y_{i,k})}} \\
&= \frac{\mathrm{Cov}(u_{i,j} + \epsilon_{i,j}, u_{i,k} + \epsilon_{i,k})}{\sigma_u^2 + \sigma_\epsilon^2} \\
&= \frac{\mathrm{Cov}(u_{i,j}, u_{i,k}) + \mathrm{Cov}(\epsilon_{i,j}, \epsilon_{i,k})}{\sigma_u^2 + \sigma_\epsilon^2} \\
&= \frac{\mathrm{Var}(u_{i,j}) + \mathrm{Cov}(\epsilon_{i,j}, \epsilon_{i,k})}{\sigma_u^2 + \sigma_\epsilon^2} \\
&= \frac{\sigma_u^2 + \sigma_\epsilon^2 \rho^{|j-k|}}{\sigma_u^2 + \sigma_\epsilon^2}.
\end{aligned}
$$

Exercise 9.10

i)
When $\boldsymbol{V}$ is known, we have from (9.10) of the book

$$
\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}.
$$

So,

$$\begin{aligned}\mathrm{E}\left[\tilde{\boldsymbol{\beta}}\right] &= \mathrm{E}\left[\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y}\right] \\ &= \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\mathrm{E}\left[\boldsymbol{y}\right] \\ &= \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\boldsymbol{\beta} \\ &= \boldsymbol{\beta}\end{aligned}$$

which means that $\tilde{\boldsymbol{\beta}}$ is an unbiased estimator of $\boldsymbol{\beta}$.

ii)

$$\begin{aligned}\mathrm{Var}\left(\tilde{\boldsymbol{\beta}}\right) &= \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\mathrm{Var}\left(\boldsymbol{y}\right)\left(\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\right)^{\mathrm{T}} \\ &= \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{V}\boldsymbol{V}^{-1}\boldsymbol{X}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1} \\ &= \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\end{aligned}$$

iii)
When there are no random effects, $\boldsymbol{u} = \boldsymbol{0}$ and $\boldsymbol{V} = \boldsymbol{R}_{\boldsymbol{\epsilon}} = \sigma_\epsilon^2\boldsymbol{I}$. So,

$$\tilde{\boldsymbol{\beta}} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y} = \left(\boldsymbol{X}^{\mathrm{T}}\left(\sigma_\epsilon^2\right)^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\left(\sigma_\epsilon^2\right)^{-1}\boldsymbol{y} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

and

$$\mathrm{Var}\left(\tilde{\boldsymbol{\beta}}\right) = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1} = \left(\boldsymbol{X}^{\mathrm{T}}\left(\sigma_\epsilon^2\right)^{-1}\boldsymbol{X}\right)^{-1} = \sigma_\epsilon^2\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}.$$

When there is no random effects, those quantities become that of normal linear model.

Exercise 9.11

If $\boldsymbol{X}_i = \boldsymbol{X}_1$ and $\boldsymbol{V}_i = \boldsymbol{V}_1$ for $i = 1, \ldots, n$, we have

$$\begin{aligned}\tilde{\boldsymbol{\beta}} &= \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y} \\ &= \left(\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{V}_i^{-1}\boldsymbol{X}_i\right)^{-1}\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{V}_i^{-1}\boldsymbol{y}_i \\ &= \left(n\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{V}_1^{-1}\boldsymbol{X}_1\right)^{-1}\left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{V}_1^{-1}\sum_{i=1}^{n}\boldsymbol{y}_i\right) \\ &= \left(n\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{V}_1^{-1}\boldsymbol{X}_1\right)^{-1}\left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{V}_1^{-1}n\overline{\boldsymbol{y}}\right) \\ &= \left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{V}_1^{-1}\boldsymbol{X}_1\right)^{-1}\left(\boldsymbol{X}_1^{\mathrm{T}}\boldsymbol{V}_1^{-1}\overline{\boldsymbol{y}}\right)\end{aligned}$$

Thus the generalized least squares solution (9.10) can be expressed in terms of $\overline{\boldsymbol{y}} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\boldsymbol{y}_i$.

Exercise 9.12

Not sure about this one. I interpret exercise as asking us to show that

$$\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{V}^{-1}\boldsymbol{y} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

when $V = \sigma I + \sigma_u 1_{n\times n}$. But this is not true.

Additional Exercise 25

a)

Since

$$L = I - X\left(X^T X\right)^{-1} X^T,$$

$LX = X - X = 0$. Since $A = \begin{bmatrix} I_{N-p} & 0_{(N-p)\times p} \end{bmatrix} L$, $AX = 0$ as claimed.

b)

We know that $Y = X\beta + \epsilon$ with $\epsilon \sim N(0, V_\theta)$. Since $AX = 0$, we're left with $A\epsilon$, which is normal with covariance $AV_\theta A^T$.

c)

The density of the multivariate normal $N(0, V_\theta)$ is

$$\left|2\pi AV_\theta A^T\right|^{-1/2} \exp\left[-1/2 y^T \left(AV_\theta A^T\right)^{-1} y\right]$$

Hence the log-likelihood is

$$-1/2 \log\left|AV_\theta A^T\right| - 1/2 y^T \left(AV_\theta A^T\right)^{-1} y - \frac{n-p}{2}\log(2\pi)$$

d)

The log-likelihood is (ignoring the constant)

$$\begin{aligned} -1/2 \log\left|BAV_\theta A^T B\right| - 1/2 (By)^T \left(BAV_\theta A^T B\right)^{-1}(By) &= \\ -1/2 \log\left|AV_\theta A^T\right| - |B| - 1/2 y^T \left(AV_\theta A^T\right)^{-1} y & \end{aligned}$$

It only differs by the constant $-|B|$, so the maximum likelihood estimator is the same.

e)

Set $V_\theta = \sigma^2 V_0$ and put $C = AV_0A^T$, then differentiate with respect to σ:

$$\begin{aligned} \frac{d}{d\sigma}\left[-1/2 \log|C| - (n-p)\log\sigma - 1/2\sigma^{-2} y^T C^{-1} y\right] &= \\ \frac{n-p}{\sigma} - \sigma^{-3} y^T C^{-1} y &= 0 \end{aligned}$$

Thus

$$\sigma^2 = \frac{y^T C^{-1} y}{n-p}$$

as claimed.

f)

In general

$$E\left(y^T A y\right) = \mathrm{tr} A V + \mu^T A \mu$$

when $y \sim N(\mu, V)$. Thus

$$\begin{aligned} E\left(y^T C^{-1} y\right) &= \mathrm{tr} C^{-1} \sigma^2 A^T V_0 A \\ &= \sigma^2 \mathrm{tr}\left[\left(A^T V_0 A\right)^{-1} A^T V_0 A\right] \\ &= \sigma^2 (n-p) \end{aligned}$$