Longitudinal Homework 3

Tim Vigers

03 October 2019

1. Cell counts

Starting with the subject-level model, define Z, G, and R matrices:

$$Z_i = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$G_i = \sigma_0^2$$

$$R_i = \begin{bmatrix} 1 & \phi & \phi^2 & \phi^3 \\ \phi & 1 & \phi & \phi^2 \\ \phi^2 & \phi & 1 & \phi \\ \phi^3 & \phi^2 & \phi & 1 \end{bmatrix}$$

V_i is the variance of Y_i, so:

$$Var(Y_i) = Z_i G_i Z_i^t + \sigma_\epsilon^2 R_i$$

$$= \begin{bmatrix} \sigma_0^2 & \sigma_0^2 & \sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 & \sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 & \sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 & \sigma_0^2 & \sigma_0^2 \end{bmatrix} + \begin{bmatrix} \sigma_\epsilon^2 & \phi\sigma_\epsilon^2 & \phi^2\sigma_\epsilon^2 & \phi^3\sigma_\epsilon^2 \\ \phi\sigma_\epsilon^2 & \sigma_\epsilon^2 & \phi\sigma_\epsilon^2 & \phi^2\sigma_\epsilon^2 \\ \phi^2\sigma_\epsilon^2 & \phi\sigma_\epsilon^2 & \sigma_\epsilon^2 & \phi\sigma_\epsilon^2 \\ \phi^3\sigma_\epsilon^2 & \phi^2\sigma_\epsilon^2 & \phi\sigma_\epsilon^2 & \sigma_\epsilon^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_0^2 + \sigma_\epsilon^2 & \sigma_0^2 + \phi\sigma_\epsilon^2 & \sigma_0^2 + \phi^2\sigma_\epsilon^2 & \sigma_0^2 + \phi^3\sigma_\epsilon^2 \\ \sigma_0^2 + \phi\sigma_\epsilon^2 & \sigma_0^2 + \sigma_\epsilon^2 & \sigma_0^2 + \phi\sigma_\epsilon^2 & \sigma_0^2 + \phi^2\sigma_\epsilon^2 \\ \sigma_0^2 + \phi^2\sigma_\epsilon^2 & \sigma_0^2 + \phi\sigma_\epsilon^2 & \sigma_0^2 + \sigma_\epsilon^2 & \sigma_0^2 + \phi\sigma_\epsilon^2 \\ \sigma_0^2 + \phi^3\sigma_\epsilon^2 & \sigma_0^2 + \phi^2\sigma_\epsilon^2 & \sigma_0^2 + \phi\sigma_\epsilon^2 & \sigma_0^2 + \sigma_\epsilon^2 \end{bmatrix}$$

Specifying a G and an R matrix gives a more flexible model that accounts for both within-subject correlation and the decaying correlation between time points. If you only used the AR(1) structure, then the variance will go to 0 as the time points get farther apart. When σ_0^2 is added, then this can't happen, which is why the model is more flexible.

2. Mt. Kilimanjaro

$$G_i = \begin{pmatrix} \sigma_I^2 & \sigma_{IS}^2 \\ \sigma_{IS}^2 & \sigma_S^2 \end{pmatrix} Z_i = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} Z_i^t = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} R_i = \begin{pmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_e^2 & 0 \\ 0 & 0 & \sigma_e^2 \end{pmatrix} V_i = Var(Y_i) = Z_i G_i Z_i^t + R_i = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \sigma_I^2 & \sigma_{IS}^2 \\ \sigma_{IS}^2 & \sigma_S^2 \end{pmatrix}$$

In order to show this, you compare the covariance for times 0 and 1, and for times 0 and 2. If there's more correlation between time 0 and 1 than there is between time 0 and 2, then cov(0,1) > cov(0,2). This turns out to be fairly easy to rearrange, and shows that there can be decay as time between measurements increases, as long as the below conditions are met:

$$\sigma^2_{IS} + \sigma^2_I > 2\sigma^2_{IS} + \sigma^2_I < 3\sigma^2_{IS} + \sigma^2_I + 2\sigma^2_S 0 > \sigma^2_{IS} < 2\sigma^2_{IS} + 2\sigma^2_S$$

There must be an inverse relationship between the random effects ($\sigma^2_{IS} < 0$) and $\sigma^2_{IS} + 2\sigma^2_S$ must be greater than 0.

3. Beta carotene data

Convert the data to long form and make a continuous time variable (using baseline 2 as time 0):

```
# Read in
carotene <- read.csv("/Users/timvigers/Documents/GitHub/School/Analysis of Longitudinal Data/Homework 3/
# Wide to long
long <- melt(carotene,id.vars = c("Id","Prepar"))
# Use baseline 2 as time 0
long <- long %>% filter(variable != "Base1lvl") %>% arrange(Id,variable)
# Continuous time
long$time <- long$variable
long$time <- plyr::mapvalues(long$time,
          from = c("Base2lvl","Wk6lvl","Wk8lvl","Wk10lvl","Wk12lvl"),
          to = c(0,6,8,10,12))
long$time <- as.numeric(as.character(long$time))
long$Prepar <- as.factor(long$Prepar)
kable(head(long,10))
```

Id	Prepar	variable	value	time
71	1	Base2lvl	116	0
71	1	Wk6lvl	174	6
71	1	Wk8lvl	178	8
71	1	Wk10lvl	218	10
71	1	Wk12lvl	190	12
72	3	Base2lvl	162	0
72	3	Wk6lvl	432	6
72	3	Wk8lvl	336	8
72	3	Wk10lvl	440	10
72	3	Wk12lvl	472	12

Fit a polynomial model for time and compare AIC and BIC to determine the sufficient degree:

```
# Models
# Time polynomials
lin_mod <- gls(value ~ time*Prepar,
            data = long, method = "ML",correlation=corSymm(form = ~1|Id),
            weights = varIdent(form = ~1|time))
quad_mod <- gls(value ~ time*Prepar +
                    I(time^2)*Prepar,
            data = long, method = "ML",correlation=corSymm(form = ~1|Id),
            weights = varIdent(form = ~1|time))
cub_mod <- gls(value ~ time*Prepar +
                    I(time^2)*Prepar +
```

```
                   I(time^3)*Prepar,
            data = long, method = "ML",correlation=corSymm(form = ~1|Id),
            weights = varIdent(form = ~1|time))
quart_mod <- gls(value ~ time*Prepar +
                    I(time^2)*Prepar +
                   I(time^3)*Prepar +
                     I(time^4)*Prepar,
            data = long, method = "ML",correlation=corSymm(form = ~1|Id),
            weights = varIdent(form = ~1|time))
kable(AIC(lin_mod,quad_mod,cub_mod,quart_mod))
```

	df	AIC
lin_mod	23	1251.998
quad_mod	27	1246.079
cub_mod	31	1243.108
quart_mod	35	1245.336

```
kable(BIC(lin_mod,quad_mod,cub_mod,quart_mod))
```

	df	BIC
lin_mod	23	1315.132
quad_mod	27	1320.192
cub_mod	31	1328.201
quart_mod	35	1341.408

The cubic model is slightly lower by AIC and definitely better by BIC, so we'll continue with this model.

a. Compare to class variable model

The cubic model can be written:

$$Y_{hi} = \mu + \alpha_1 + \alpha_2 + \alpha_3 + \tau_h + \gamma_{1h} + \gamma_{2h} + \gamma_{3h} + b_i + \epsilon_{hi} b_i \quad \text{iid } N(0, \sigma_b^2) \epsilon_{hi} \quad \text{iid } N(0, \sigma_\epsilon^2) \text{ and } \epsilon_i \quad \text{iid } N(0, R_i) \text{ where } R_i \text{ is unstructu}$$

Here h represents group and i represents subject. The way this model is written, $\alpha_1, \alpha_2,$ and α_3 represent the effect of time, time squared, and time cubed respectively, and $\gamma_{1h}, \gamma_{2h},$ and γ_{3h} represent the interaction effects of group and time. b_i is the random intercept for subject and ϵ_{hi} is the error term.

Now compare this to the linear model with group, time and group*time as categorical variables:

```
class_mod <- gls(value ~ factor(variable)*Prepar,
            data = long, method = "ML",correlation=corSymm(form = ~1|Id),
            weights = varIdent(form = ~1|time))
AIC(cub_mod,class_mod)
```

```
##           df      AIC
## cub_mod   31 1243.108
## class_mod 35 1245.336
```

```
BIC(cub_mod,class_mod)
```

```
##           df      BIC
## cub_mod   31 1328.201
## class_mod 35 1341.408
```

I think I would include the cubic model in the report, even though I'm not particularly comfortable interpreting polynomial models, and I think they can sometimes be difficult to interpret. However, in this case the class variable model includes a lot of parameters and the cubic model was slightly better by AIC and much better by BIC. That said, the class model is easier to interpret and not much worse by AIC, so I think there are good reasons to report either one.

b. Contrast

```r
# By group, time, and group*time
emm_group <- emmeans(cub_mod, specs = ~Prepar)
emm_group
```

```
##  Prepar emmean   SE   df lower.CL upper.CL
##  1         256 45.5 22.7    162.2      351
##  2         193 45.5 22.7     99.2      288
##  3         318 49.9 22.7    214.9      421
##  4         316 45.5 22.7    221.5      410
##
## Degrees-of-freedom method: boot-satterthwaite
## Confidence level used: 0.95
```

```r
group1 <- c(1,0,0,0)
group4 <- c(0,0,0,1)
contrast(emm_group, method = list("Group 1 vs. group 4" = group4 - group1))
```

```
##  contrast            estimate   SE   df t.ratio p.value
##  Group 1 vs. group 4     59.3 64.4 22.7 0.922   0.3664
##
## Degrees-of-freedom method: boot-satterthwaite
```

The estimate above compares the group 1 mean to the group 4 mean at the average time (7.2). The difference between the two is not statistically significant (p=0.37).

4. Children and schools measured over time

Write out the model:

$$Y_{hij} = \text{fixed effects} + b_h + b_{i(h)} + \epsilon_{hij} b_h \quad N(0, \sigma_s^2) b_{i(h)} \quad N(0, \sigma_{i(s)}^2) \epsilon \quad N(0, \sigma_\epsilon^2)$$

Next write out the Z, G and matrices for a school h:

$$Z_h = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} G_h = \begin{bmatrix} \sigma_h^2 & 0 & 0 & 0 \\ 0 & \sigma_{1(h)}^2 & 0 & 0 \\ 0 & 0 & \sigma_{2(h)}^2 & 0 \\ 0 & 0 & 0 & \sigma_{3(h)}^2 \end{bmatrix} R_h = \sigma_\epsilon^2 I_{8x8}$$

Then use $V_h = Z_h G_h Z_h^t + R_h$:

$$Z_hG_hZ_h^t = \begin{bmatrix} \sigma_h^2 & \sigma_{1(h)}^2 & 0 & 0 \\ \sigma_h^2 & \sigma_{1(h)}^2 & 0 & 0 \\ \sigma_h^2 & \sigma_{1(h)}^2 & 0 & 0 \\ \sigma_h^2 & 0 & \sigma_{2(h)}^2 & 0 \\ \sigma_h^2 & 0 & \sigma_{2(h)}^2 & 0 \\ \sigma_h^2 & 0 & \sigma_{2(h)}^2 & 0 \\ \sigma_h^2 & 0 & 0 & \sigma_{3(h)}^2 \\ \sigma_h^2 & 0 & 0 & \sigma_{3(h)}^2 \end{bmatrix} * \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$Z_hG_hZ_h^t = \begin{bmatrix} \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{3(h)}^2 & \sigma_h^2+\sigma_{3(h)}^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{3(h)}^2 & \sigma_h^2+\sigma_{3(h)}^2 \end{bmatrix}$$

$$Z_hG_hZ_h^t+R_h = \left[\begin{matrix} \sigma_h^2+\sigma_{1(h)}^2+\sigma_\epsilon^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2+\sigma_\epsilon^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2 & \sigma_h^2+\sigma_{1(h)}^2+\sigma_\epsilon^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{2(h)}^2+\sigma_\epsilon^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2+\sigma_\epsilon^2 & \sigma_h^2+\sigma_{2(h)}^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2 & \sigma_h^2+\sigma_{2(h)}^2+\sigma_\epsilon^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \\ \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 & \sigma_h^2 \end{matrix}\right.$$