picture\_book

Jonathan Tay

November 17, 2017

## R Markdown

Firstly, let’s examine the relationship between age and cognition. This graph is a terse description of the relationship between the two variables:

# {r echo=F, warning=F} # library(ggplot2) # # ggplot(dfLong, aes(age, cognitiveindex, group=rundmcs)) + # geom\_point() + geom\_line() + # geom\_smooth(aes(group=1), method="lm", colour="red3") + # geom\_smooth(aes(group=1), method="lm", formula = y~poly(x, 2), colour="royalblue2") + # scale\_x\_continuous(limits=c(50, 91), expand=c(0,0)) + # theme\_classic() #

# 

# This contains all the information we can use in linear mixed effects regression (lmer). Let’s break it down, step by step, so that each term in the model is interpretable.

# 

# First, the effects of age can be broken down into two variables: baseline age and time between follow-ups. First, the linear term for baseline age is simply expressing the cross-sectional correlation between age and cognition:

# 

# The quadratic is positive

# 

# {r echo=FALSE, warning=F, message=F} # ggplot(dfWide, aes(Age\_2006, cognitiveindex06)) + # geom\_point() + geom\_smooth(method="lm", colour="red3") + # scale\_x\_continuous(limits=c(50, 85), expand=c(0,0)) + # theme\_classic() #

# 

# The fixed effect of time between follow-ups expresses the linear trend between cognition and the passage of time:

# 

# {r echo=FALSE, warning=F, message=F} # ggplot(dfLong, aes(age, cognitiveindex, group=rundmcs)) + # geom\_point() + geom\_line() + # geom\_smooth(aes(group=1), method="lm", colour="red3") + # scale\_x\_continuous(limits=c(50, 91), expand=c(0,0)) + # theme\_classic() #

# 

# And the fixed effect of time squared between follow-ups expresses the quadratic trend between cognition and the passage of time:

# 

# {r echo=F, warning=F, message=F} # ggplot(dfLong, aes(age, cognitiveindex, group=rundmcs)) + # geom\_point() + geom\_line() + # geom\_smooth(aes(group=1), method="lm", formula = y~poly(x, 2), colour="royalblue2") + # scale\_x\_continuous(limits=c(50, 91), expand=c(0,0)) + # theme\_classic() #

# 

# Then, two random effects are specified for intercept and slopes, grouped by

# subject. This expresses subject specific changes in cognition over time. The

# correlation between slope and intercept implies that we expect the rate of

# change for each subject to depend on their age (as evidenced by how cognition

# declines quadratically towards 80-90.

# 

# The random effect of each subject can be visualised as the total effect of each

# individual line segment:

# 

# {r echo=F, warning=F} # ggplot(dfLong, aes(age, cognitiveindex, group=rundmcs)) + # geom\_point() + geom\_line() + # scale\_x\_continuous(limits=c(50, 91), expand=c(0,0)) + # theme\_classic() #

# 

# To more easily see this relationship, we can examine individual line segments, plotted against each other with the same vertical axis (i.e., the baseline visit):

# 

# {r message=F, warning=F, echo=F} # library(lmerTest) # library(sjPlot) # fit <- lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject), data = dfLong) # sjp.lmer(fit, vars = "time", type = "rs.ri") #

# 

# It is quite difficult to see individual trends, so we’ll only plot a subset (n = 15):

# 

# {r message=F, warning=F, echo=F} # sjp.lmer(fit, vars = "time", type = "rs.ri", sample.n=15, sort.est = "sort.all") #

# 

# These lines are fairly flat, suggesting that change in cognition for each individual subject is low between time points. This is reflected in the low variance of the slope term for time in the initial LMER, which only includes terms for age and time:

# 

# {r message=F, warning=F} # fit <- lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject), data = dfLong) # summary(fit) #

# 

# The slope term is close to 0. This suggests that the timeframe for follow-ups is not sufficient to detect change on the individual level; indeed, the nonlinear relationship between age and cognitive change was only seen in the population range of 50-90. 10 years for a subject at 50 isn’t likely going to show as much cognitive decline compared to, say, 10 years for a subject at 80. This likely reflects inclusion bias (e.g., younger participants were more likely to be enrolled because they have better functioning, as evidenced by how average age at enrollment was ~65; these may be more pronounced if average age was 75).

# 

# Indeed, there are some participants with a non-normal (compared to other subjects) random slope term, as evidenced by a qq-plot for all subjects:

# 

# {r message=F, warning=F, echo=F} # sjp.lmer(fit, vars = "time", type = "re.qq") #

# 

# Now, on to adding terms for WMH and HV. Since WMH progression has a significant

# quadratic change with time (van Leijsen et al., 2017), we must first examine

# interactions between WMH and either linear or quadratic terms. This is done

# at the population level (i.e., modelled as a fixed effect). We also include a

# random slope term for WMH to capture subject variability (which covers any

# higher polynomial that might govern an individual’s developmental trajectory).

# 

# Let’s see which interaction (linear or quadratic) best describes the data:

# 

# {r warnings=F} # fit1 <- lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject) + # lnwmh + lnwmh\*time, data = dfLong) # fit2 <- lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject) + # lnwmh + lnwmh\*timesq, data = dfLong) # summary(fit1) # summary(fit2) # anova(fit1, fit2) #

# 

# Interestingly, the quadratic interaction seems to better describe the data than the linear interaction, as well as the main effect of WMH.

# 

# Now to add HV to the models. We can test two models here easily: one is that HV and WMH are independent, while the other is that they interact. HV shall interact only with linear time (we have no reason to believe it’s not linear at this stage in SVD):

# 

# {r warnings=F} # fit1 <- lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject) + # lnwmh + lnwmh\*timesq + hv + hv\*time, data = dfLong) # fit2 <- lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject) + # lnwmh + lnwmh\*timesq + hv + hv\*time + lnwmh\*hv, data = dfLong) # summary(fit1) # summary(fit2) # anova(fit1, fit2) #

# 

# Interesting eh? It makes sense that WMH and HV become non-significant when adding respective interaction terms with time (as we expect both WMH progression and HV atrophy to be “symptoms” of age-related cognitive decline). Interestingly it seems like the interaction with WMH renders HV non-significant. Perhaps Wallerian degeneration driven by WMH does this.

# 

# Let’s try it with mice!:

# ```{r message=F}

# library(mice)

# dfTemp <- mice(dfLong)

# fit1 <- with(dfTemp, lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject) +

# lnwmh + lnwmh*timesq + hv + hv*time))

# fit2 <- with(dfTemp, lmer(cognitiveindex ~ 1 + age06 + time + timesq + (1 + time|subject) +

# lnwmh + lnwmh*timesq + hv + hv*time + lnwmh*hv)) # fit3 <- with(dfTemp, lmer(cognitiveindex ~ 1 + age06 + poly(time, 2) + (1 + time|subject) + # lnwmh + lnwmh*timesq + hv + hv\*time)) # # summary(pool(fit1)) # summary(pool(fit2)) # summary(pool(fit3))