

R: Computer Lab Session 2 Tuesday March 9, 2021

The goals of this session are to learn how to

- More practice creating data sets in R with centered and/or scaled variables.
- Fit random intercept and slope models.
- Detect problems in estimated solutions and fix them.

For this computer lab, we'll use the USA TIMSS data and model science scores again; however, I've added some micro and macro variables. A list and definition of all the variables in this data set are on the last page.

1. Read in the data set "lab2.txt"
2. Create a school centered math scores (i.e., `school_center_math`), and the school math means (i.e., `school_mean_math`).
3. Make gender, id and grade factor variables.
4. We will change the optimizer to the Nelder-Mead (the default is bound optimization by quadratic approximation or `bobyqa`). A nice page for this is at <http://svmiller.com/blog/2018/06/mixed-effects-models-optimizer-checks>.

For example,

```
# model j
modelj <- lmer(science ~ 1 + school_center_math + third + gender
              + school_mean_math + (1 | idschool), lab2,
              REML=FALSE, control = lmerControl(optimizer = "Nelder-Mead"))
```

5. Fit the following models to the data:
 - (j) Random intercept model with `school_center_math`, `gender`, `grade`, `school_mean_math`. Verify that you get the same results as model (j) from computer lab/homework #1. (i.e., the one just above).
 - (m) Random intercept model with `school_center_math`, `gender`, `third`, `school_mean_math` (i.e., the model from part (j)) but add the following micro level variables to the model: `hoursTV`, `hourscomputergames`. This model should include all micro level variables, as well as `school_mean_math`, which is an explanatory variable for the random intercept.

- (n) The model part (m) but add `community` as an additional explanatory variable for the intercept.
- (o) Model (n) but add random slope for `school_center_math`.
Note that this model does not converge if you use the default optimizer!
- (p) Model (o) with random slope for `school_center_math` but remove type of `community`.
- (q) Model (p) but random slope for only `hoursTV`.
- (r) Model (p) with a random slope for only `hourscomputergames`.
- (s) Model (p) with a random slope for `school_center_math` where `school_mean_math` is an explanatory variable for the random slope.
- (s)-rescaled When fitting model (s) you had a lot of warning messages about scaling. This will take care of them.
 - (a) Compute standard deviations and variances of micro and macro variables (and interaction) that are in model (u).
 - (b) Rescale `school_center_math` and `grpMmath`. The following will do the trick:


```
> lab2$school_center_math <- scale(school_center_math,
  center=FALSE,scale=TRUE)
> lab2$grpMmath <- scale(grpMmath, center=FALSE,scale=TRUE)
```
 - (c) Re-fit model (s) but used re-scaled variables.
 - (d) **NOTE: From here on out, used the re-scaled variables.**
- (t) Model (s)-rescaled but add `short2` and `short3` as an explanatory variable for the random intercept.
- (u) Model (t) but add `short2` and `short3` as explanatory variables for the random slope for `school_center_math`.
- (v) Fit a really complex model with `school_center_math`, `gender`, `third`, `hoursTV`, `hourscomputergames`, `grpMmath`, `shortages`, `isolate`, `rural`, `suburban`, `school_center_math*grpMmath`, `school_center_math*shortages`, `school_center_math*isolate` + `school_center_math*rural` and random intercept, and random slopes for `school_center_math` and `hoursTV`
- (w) Any other model that you want to try or any that you may need to help with homework.

Variable in Data Set for Computer Lab 2:

TIMSS	SAS data	Values	Description
idschool	idschool	10–263	School ID
idstud	idstud		Student ID
idgrade	grade	3 or 4	Student’s grade in school
assnrsc	science	103.4 – 185	Science score based on ability estimates using the RASCH IRT model.
asmnrsc	math	104.3 – 189	Math scores based on ability estimates using the RASCH model.
itsex	gender	girl, gender	Student’s gender
acbgst01	genshortages		General school shortages of instructional materials
		none a little some a lot	
—	shortages	0,1,2,3	Re-coded genshort as a numerical variable.
asbgday1	hoursTV		Time spent watching TV or videos
		1 = 2 = 3 = 4 = 5 =	no time < 1 hr. 1-2 hrs. 3-4 hrs. >4 hrs.
asbgday2	hourscomputergames		Time spent playing computer games
		1 = 2 = 3 = 4 = 5 =	no time less than 1 hour” 1–2 hours 3–4 hours more than 4 hours
acbgcomm	typecommunity		Type of community where school is located
		1 = 2 = 3 = 4 =	a geographically isolated area village or rural (farm) area one on the outskirts of a town/city one close to the center of a town/city

Dummy codes created to help with fitting discrete effects

<code>third</code>	= 1 if 3rd grade, 0 if 4th grade
<code>short0</code>	= 1 if shortages “none”, = 0 otherwise
<code>short1</code>	= 1 if shortages “a_little”, = 0 otherwise
<code>short2</code>	= 1 if shortages “some”, = 0 otherwise
<code>short3</code>	= 1 if shortages “a_lot”, = 0 otherwise
<code>isolate</code>	= 1 if type community is isolate, = 0 otherwise
<code>rural</code>	= 1 if type community is rural, = 0 otherwise
<code>suburban</code>	= 1 if type community is suburban, = 0 otherwise
<code>urban</code>	= 1 if type community is urban, = 0 otherwise
