ST 437/537: Applied Multivariate and Longitudinal Data Analysis

Longitudinal Data Analysis: Random Coefficient Model (RCM)

Arnab Maity
NCSU Department of Statistics
SAS Hall 5240 919-515-1937 amaity[at]ncsu.edu

References:

- Modeling Longitudinal Data by Robert E. Weiss. New York: Springer.
- Linear Mixed Models for Longitudinal Data by Geert Verbeke and Geert Molenberghs. New York: Springer.
- Applied Longitudinal Analysis by Fitzmaurice by G.M., Laird, N.M., and Ware, J.H. New York: Wiley (on reserve at NCSU library)

Introduction

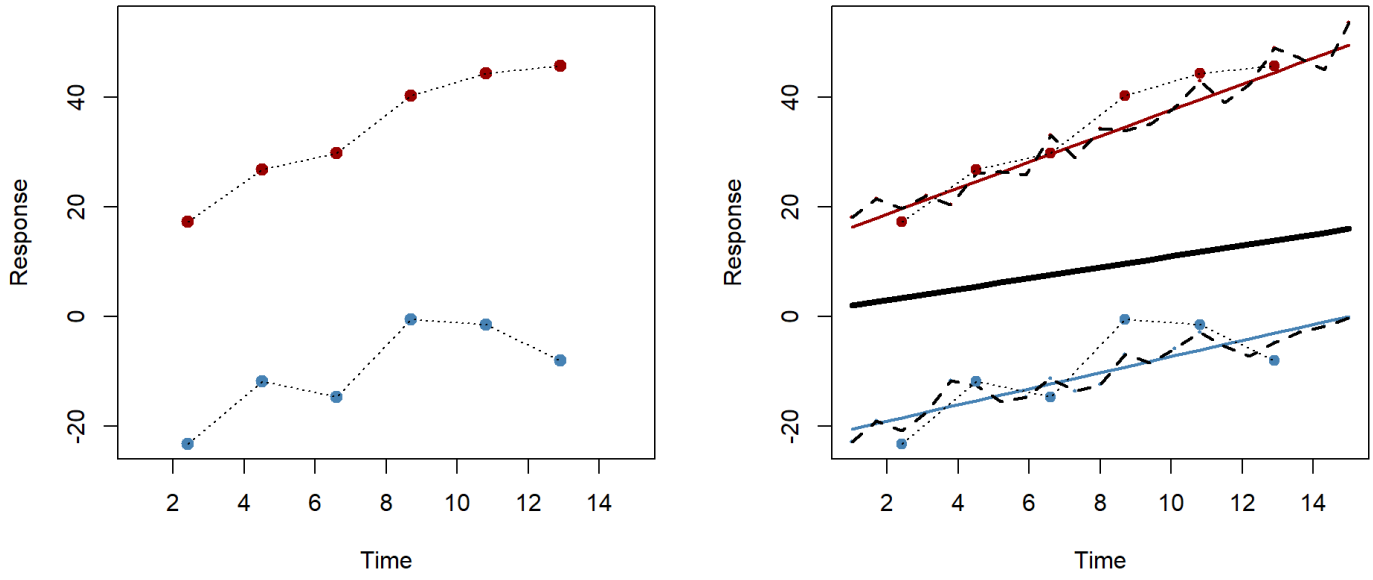
The general linear model approach to longitudinal data discusses last chapter has two main disadvantages:

- The modeling of the covariance matrix aggregates the two sources of variation and does not allow the analysts to understand the two sources separately
- The main focus is modeling of the mean trajectories over time; the reconstruction of the individual trajectories is not considered. Characterizing the subject trajectories may be of interest but the general linear model framework does not allow such study.

Let us discuss the two points mentioned above in more detail.

Sources of variation

Let us consider the hypothetical situation shown below.



The left panel shows hypothetical observations of two subjects. The solid circles are the observed responses.

The right panel shows four features of the data and data generating process:

- The "true" **population mean trajectory** (in thick solid line) over time. We can view this as the mean of *all* subject trajectories in the population. This is of course unknown (but fixed), and we want to estimate this mean trajectory.
- The **inherent trend for each subject** (the thin solid lines). Even if the trend is linear, we can not expect the observed data for a particular subject would exactly fall on the line. Thus we can envision this trend as a general pattern of the subject trajectories. The position of this trend tells us if the subjects are higher or lower compared to the mean trajectory. Thus this trend represents the **biological deviation among units**.
- The **error free subject observatons** (the thin dashed lines). These are observations for each subject *if there was no measurement error*. We can view these as fluctuations around the smother trend and represent how responses for that subject may evolve.
- The actual **observed responses** of the subjects (solid circles). The observed values are the error free trends with added measurement errors.

Thus we have the conceptul model:

$$\text{Data}_{ij} = \text{Mean}_j + \text{SubjSpecific}_{ij} + \text{BiologicalDev}_{1ij} + \text{Error}_{2ij},$$

where

- $\text{Mean}_j = \mu_j$ is the overall (population) mean at time point $t_{ij} = t_j$,
- SubjSpecific_{ij} represents the biological variation of the ith unit; this deviation dictates the inherent trend of the ith subject,
- $\text{BiologicalDev}_{1ij}$ is the component of deviation from the subject's trend that is due to the biological variation over time within the subject,
- Error_{2ij} is the component of the deviation that is due to the measurement error.

Thus we can identify two sources of variation:

- The SubjSpecific_{ij} term represnt **among (between) units variaton**.
- The sum of biological deviation and measurement errors ($\text{BiologicalDev}_{1ij}$ + Error_{2ij}) represent **within unit variation**.

The general linear model approach models the total variation (variation due to SubjSpecific_{ij} + $\text{BiologicalDev}_{1ij}$ + Error_{2ij}) but we can not identify individual components of the variation. In contrast the modeling we study in this chapter focuses primarily on modeling the subject/unit trajectory.

Our approach

The intuition behind this approach is that we consider the subject/unit trajectory itself and model its behavior using 2 stages:

- **Stage 1:** Describe the trend of each subject trajectory by using some sort of parametric model and subject-specific parameters (*subject-level stage*),
- **Stage 2:** Describe how the subject-specific parameters vary across subjects (*population stage*).

This modeling approach explicitly acknowledges the two sources of variation: within-unit and between-unit. The perspective offer more flexible models that do not require balanced/regular designs across units/subjects, allow for more general covariance structures, and can easily accommodate additional covariate information. In this chapter we will discuss interpretation of the model components/estimation and inference as well as prediction of the full subject trajectory.

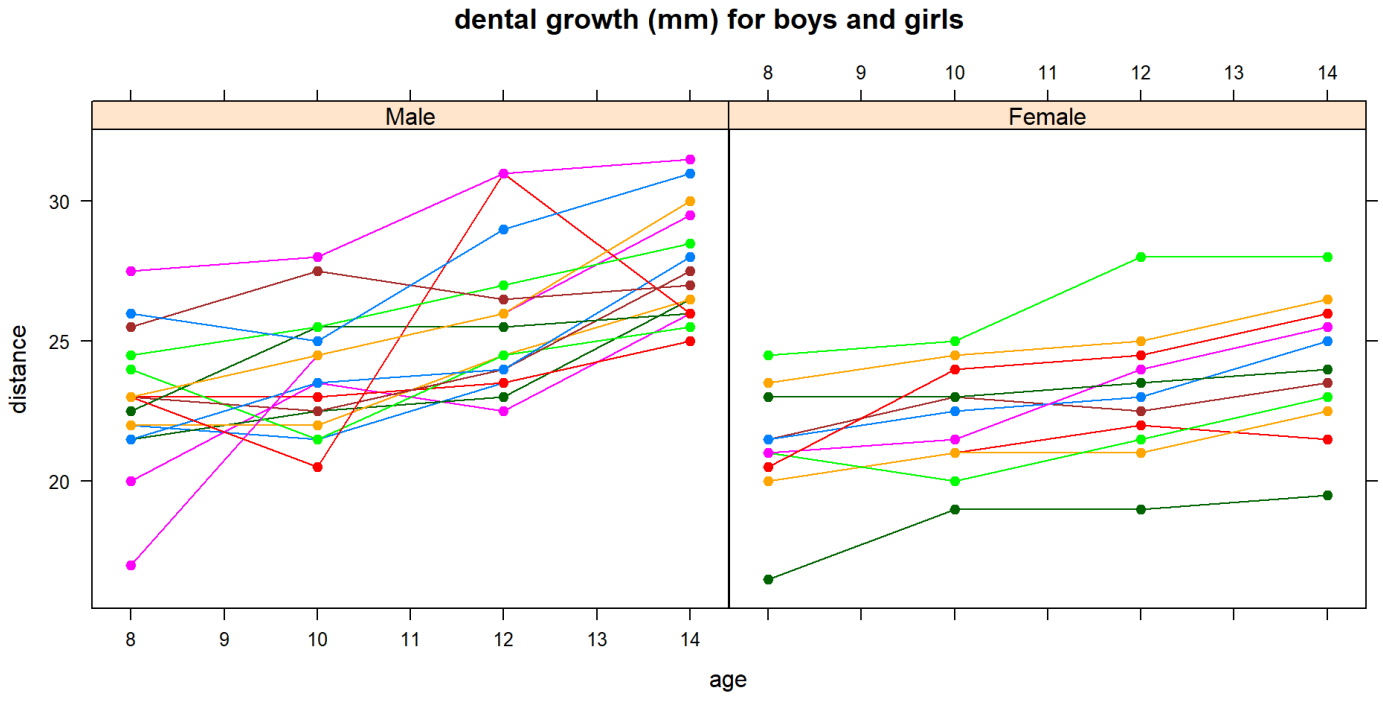
Random Coefficient Model (RCM)

Recall the Dental study; see Example A in the lecture on **[models for mean and covariance] (Lecture08_LDA_Modeling_and_Estimation.html)**. In summary, researchers collected dental growth measurements of the distance (mm) from the center of the pituitary gland to the pterygomaxillary fissure for 27 children (11 girls and 16 boys) at ages 8, 10, 12, and 14.

```
library(nlme)
library(latticeExtra)
head(Orthodont)
```

```
## Grouped Data: distance ~ age | Subject
##   distance age Subject  Sex
## 1     26.0   8     M01 Male
## 2     25.0  10     M01 Male
## 3     29.0  12     M01 Male
## 4     31.0  14     M01 Male
## 5     21.5   8     M02 Male
## 6     22.5  10     M02 Male
```

```
xyplot(distance ~ age | Sex, data=Orthodont, type="b", groups = Subject,
       pch=19, main = "dental growth (mm) for boys and girls")
```



Consider the following questions:

- Examine the response trajectory for each subject in part. What features do you observe (look at how the distance varies over time and how it changes)?

- Discuss a model that could be used to describe mathematically the way that the distance (response) for a subject varies over time. Make sure that whatever parameters you use are specific to the subject.

- Using this perspective, what type of variation do you think you can quantify at this step (recall we discuss two types of variations: within and between).

- What is the other type of variation that we need to describe. Could you suggest possible ways to describe this other type of variation (in an intuitive way)?

Subject-level stage

Recall our set up: $Y_{ij} = Y_i(t_{ij})$ denotes the response observed for the ith subject at time t_{ij}. Suppose we specify a *linear mean trend* model for the response trajectory of the ith subject:

$$Y_i(t_{ij}) = \beta_{0i} + \beta_{1i} t_{ij} + e_{ij},$$

where

- β_{0i} is the *subject-specific intercept* and
- β_{1i} is the *subject-specific slope*.

The line $\beta_{0i} + \beta_{1i} t_{ij}$ describes the response trajectory for the ith subject, in an average way; it is called the **subject-specific mean trajectory**.

The departure of the response Y_{ij} from the ith subject-mean trajectory is considered random and is attributable to either biological fluctuations about the subject-mean trajectory or measurement error (or both). Here e_{ij} denotes the random deviation from the subject-mean trajectory; it is assumed to have zero mean. The model assumed for the variation of e_{ij} describes the **within-subject variation**.

Population stage model

The parameters included in the specification of the subject-mean trajectory depend on the subject and thus are assumed random; let $\beta_i = (\beta_{0i}, \beta_{1i})^T$ denotes the full parameter for the subject i. The model assumed for the variation of β_i describes the **between-unit** variation.

Because the subject level parameters β_i are random, the modeling approach is called **random coefficient model** (RCM). RCM are a particular case of the wider class of models *linear mixed effects (LME) models* which we study later in this chapter.

Effectively, each subject trajectory has been summarized by a intercept (β_{0i}) and a slope (β_{1i}) parameter. Thus we can view the population of trajectories as the collection of all such values of intercept and slope parameters. We assume that the **overall mean slope and intercapets** are β_0 and β_1, respectivelt, ans that each β_{0i} is varying around the population mean β_0 (and similarly β_{1i} varying around β_1).

Thus we model

$$\beta_{0i} = \beta_0 + b_{0i} \quad \text{and} \quad \beta_{1i} = \beta_1 + b_{1i},$$

for $i = 1, \ldots, n$, where b_{0i} are b_{1i} are random. The joint-distribution of (b_{0i}, b_{1i}) essentially models the between-subjects variation. Generally, we might assume that the joint distribution is multivariate normal:

$$\begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \right), \text{ independent over } i.$$

Here the matrix D is unknown, and has to be estimated from the data.

In the formulation described above:

- β_0 and β_1 are fixed (but unknown) quantities that need to be **estimated**; they are called **fixed effects**.
- b_{0i} and b_{1i} are random deviations (also unobserved) that can be **predicted**; they are called **random effects**.

Since the model has both fixed and random effects, we often call it a **mixed efffects model**.

Combining two stages

Putting the models from the two stages together, we obtain the following:

$$Y_{ij} = \underbrace{(\beta_0 + b_{0i})}_{\beta_{0i}} + \underbrace{(\beta_1 + b_{1i})}_{\beta_{1i}} t_{ij} + e_{ij}.$$

The model above clearly shows that each child has intercept and slope that varies about the mean intercept and slope, β_0 and β_1.

- Let us re-write the model for Y_{ij} to separate the systematic trend (or population mean) from the random trend (or subject specific deviation).

- Write down the population mean trajectory. Discuss whether this makes sense.

Note that we can write this model in a matrix form. Recall that $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{im_i})^T$ is the response vector of the i-th subject. So we can write

$$\begin{pmatrix} Y_{i1} \\ \vdots \\ Y_{im_i} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & t_{i1} \\ \vdots & \vdots \\ 1 & t_{im_i} \end{pmatrix}}_{\boldsymbol{X}_i} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & t_{i1} \\ \vdots & \vdots \\ 1 & t_{im_i} \end{pmatrix}}_{\boldsymbol{Z}_i} \begin{pmatrix} b_{i0} \\ b_{i1} \end{pmatrix} + \begin{pmatrix} e_{i1} \\ \vdots \\ e_{im_i} \end{pmatrix}$$

In general, we can write the random coefficient model in the matrix form

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\beta + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{e}_i$$

Specification of within-subject variation: $cov(\boldsymbol{e}_i)$

Recall that the errors $\boldsymbol{e}_i$ represent the total effect of *biological deviation* and *measurement errors*. Define $\boldsymbol{R}_i = cov(\boldsymbol{e}_i)$.

The **default choice** in many software is to assume independent errors, that is, $\boldsymbol{R}_i = \sigma^2 I_{m_i}$. This assumption might be a reasonable one if the observation times t_{ij} are far away from each other; otherwise, this assumption **should be examined carefully**.

In general, the *measurement errors* are assumed to be independent with constant variance, but sme covariance structure is put on the biological deviations.Thus R_i is specified as $R_i = \sigma_1^2 \Gamma_i + \sigma_2^2 I_{m_i}$, where Γ_i has some known correlation pattern model (AR1, exchangeable, unstructured, etc as described in the previous chaper). The two components describe the two sources of variability at the subject level:

- biological variation about the subject mean trend (which is quantified by $\sigma_1^2 \Gamma_i$) and
- the measurement error (which is quantified by $\sigma_2^2 I_{m_i}$.

In practice the structure assumed for R_i is related to which of the two sources is believed to dominate. Specifically an assumption like $R_i = \sigma^2 I_{m_i}$ has the interpretation that the measurement error dominates, while an assumption like $R_i = \sigma_1^2 \Gamma_i$ means that the biological variation dominates.

Including covariates

So far we have only looked at the subject trajectory but did not include the grouping factor (`Sex`). This can be easily done by just adding another term in the specification of subject-specific coefficients. Specifically, suppose G_i is a dummy variable (0 if male, 1 if female). Then we have:

Stage 1 model:

$$y_{ij} = \beta_{0i} + \beta_{1i} t_{ij} + e_{ij},$$

Stage 2 model:

$$\beta_{0i} = G_i \beta_{0G} + (1 - G_i)\beta_{0M} + b_{0i}$$

$$\beta_{1i} = G_i \beta_{1G} + (1 - G_i)\beta_{1M} + b_{1i}$$

Combined model:

$$y_{ij} = G_i\beta_{0G} + G_i t_{ij}\beta_{1G} + (1 - G_i)\beta_{0M} + (1 - G_i)t_{ij}\beta_{1M} + b_{0i} + t_{ij}b_{1i} + e_{ij}.$$

Interpretation of the parameters

- Write the model for the "girls" group ($G_i = 1$)

- Write the model for the "boys" group ($G_i = 0$)

- What are we actually assuming about the random deviations b_{0i} and b_{1i}? [Hint: does their covariance depend on G_i] Is this assumption realistic?

Fitting RCM in R

There are two popular libraries in R for fitting RCMs, and in general mixed effects model:

- The `lmer` function in the **`lme4` library (https://cran.r-project.org/web/packages/lme4/)**; see also their "Vignettes" section for a **comprehensive reference (https://cran.r-project.org/web/packages/lme4/vignettes/lmer.pdf)**
- The `lme` function **`nlme` library (https://cran.r-project.org/web/packages/nlme/)**

We will use the `lme` function in `nlme` package for our demonstration. The function call looks like

> *lme(formula, random, correlation, weights, data)*

where

- `formula` specifies the mean model

- `random` specifies the random effect structure (including grouping etc), that is, **between dubject** variation. This argument essentially determines what D matrix to use.
- `correlation` and `weights` are similar to those in `gls` function discussed before; however, these arguments now refer to the **within-subject** covariance structure $\boldsymbol{R}_i$

Consider the dental data again, and let us look at the following model:

$$\text{(Model A)} \quad y_{ij} = G_i\beta_{0G} + G_i t_{ij}\beta_{1G} + (1-G_i)\beta_{0M} + (1-G_i)t_{ij}\beta_{1M} + b_{0i} + t_{ij}b_{1i} + e_{ij},$$

where

- the **errors** ($\boldsymbol{e}_i$) has **diagonal within-subject covariance** ($\sigma^2 I$), that is same for each gender
- the **random effects** (b_{0i}, b_{1i}) have a **unstructured** 2×2 covariance matrix that is same for each gender.

```
# Create covariates
id <- Orthodont$Subject
age <- Orthodont$age
distance <- Orthodont$distance
sex <- Orthodont$Sex

# Dummy variables
G <- (sex == "Female") + 0 #Girls=1, Boys=0
M <- 1 - G #Defining (1-G) as a separate var for simplicity

# Formula for fixed and random effects
model.formula <- distance ~ -1 + G + G:age + M + M:age
random.formula <- ~ age | id

fit.a <- lme(model.formula,
             random = random.formula,
             method="ML")
```

Notice that

- we specify the random effects as `~ age | id`. This specifies an unstructured covariance matrix for the random coefficients for intercept (automatically included in the formula) and `age`.
- we do not specify any argiment for correlation and weights. By default, it takes $\sigma^2 I$ as the within subject covariance, as we intended in Model A.

- The argument `Method = "ML"` specifies a maximum likelihood fit. You can use REML as well.

```
summary(fit.a)
```

```
## Linear mixed-effects model fit by maximum likelihood
##  Data: NULL
##       AIC     BIC   logLik
##   443.806 465.263 -213.903
##
## Random effects:
##  Formula: ~age | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev   Corr
## (Intercept) 2.134688 (Intr)
## age         0.154139 -0.603
## Residual    1.310040
##
## Fixed effects: list(model.formula)
##           Value Std.Error DF   t-value p-value
## G     17.372727 1.2045404 25 14.422702       0
## M     16.340625 0.9987521 25 16.361042       0
## G:age  0.479545 0.1017051 80  4.715058       0
## age:M  0.784375 0.0843294 80  9.301321       0
##  Correlation:
##       G     M     G:age
## M      0.00
## G:age -0.88  0.00
## age:M  0.00 -0.88  0.00
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -3.33603362 -0.41539842  0.01039175  0.49169519  3.85819292
##
## Number of Observations: 108
## Number of Groups: 27
```

We can use the following commands to to extract different parts of the output:

1. **Fixed effects estimate**

```
beta.hat <- fixed.effects(fit.a)
beta.hat
```

```
##          G          M      G:age      age:M
## 17.3727273 16.3406250  0.4795455  0.7843750
```

2. **Predicted random efffects**

```
ran.eff <- random.effects(fit.a)
head(ran.eff)
```

```
##     (Intercept)         age
## M16 -0.89466782 -0.07698874
## M05 -1.56171573 -0.01440218
## M02 -1.12527775 -0.02469778
## M11  0.05507812 -0.11501024
## M07 -0.91118908 -0.01413118
## M08  0.12368747 -0.10101179
```

3. The D matrix

```
D <- getVarCov(fit.a, type="random.effects")
D
```

```
## Random effects variance covariance matrix
##             (Intercept)       age
## (Intercept)     4.55690 -0.198250
## age            -0.19825  0.023759
##   Standard Deviations: 2.1347 0.15414
```

4. Within subject variance component σ^2

```
sigmasq <- sigma(fit.a)^2
sigmasq
```

```
## [1] 1.716204
```

5. The within-subject covariance matrix $\boldsymbol{R}_i = cov(\boldsymbol{e}_i)$

```
# Has to be computed for a particular individual
# i = 1 (male)
R_1 <- getVarCov(fit.a, type="conditional", individual=1)
R_1
```

```
## id M01
## Conditional variance covariance matrix
##        1      2      3      4
## 1 1.7162 0.0000 0.0000 0.0000
## 2 0.0000 1.7162 0.0000 0.0000
## 3 0.0000 0.0000 1.7162 0.0000
## 4 0.0000 0.0000 0.0000 1.7162
##   Standard Deviations: 1.31 1.31 1.31 1.31
```

```
# i = 17 (female)
R_17 <- getVarCov(fit.a, type="conditional", individual=17)
R_17
```

```
## id F01
## Conditional variance covariance matrix
##        1      2      3      4
## 1 1.7162 0.0000 0.0000 0.0000
## 2 0.0000 1.7162 0.0000 0.0000
## 3 0.0000 0.0000 1.7162 0.0000
## 4 0.0000 0.0000 0.0000 1.7162
##   Standard Deviations: 1.31 1.31 1.31 1.31
```

6. **The total covariance matrix ($cov(\boldsymbol{Y}_i)$)**

```
# Has to be computed for a particular individual
# i = 1 (male)
V_1 <- getVarCov(fit.a, type="marginal", individual=1)
V_1
```

```
## id M01
## Marginal variance covariance matrix
##        1      2      3      4
## 1 4.6216 2.8891 2.8727 2.8563
## 2 2.8891 4.6839 3.0464 3.1251
## 3 2.8727 3.0464 4.9363 3.3938
## 4 2.8563 3.1251 3.3938 5.3788
##   Standard Deviations: 2.1498 2.1642 2.2218 2.3192
```

```
# i = 17 (female)
V_17 <- getVarCov(fit.a, type="marginal", individual=17)
V_17
```

```
## id F01
## Marginal variance covariance matrix
##        1      2      3      4
## 1 4.6216 2.8891 2.8727 2.8563
## 2 2.8891 4.6839 3.0464 3.1251
## 3 2.8727 3.0464 4.9363 3.3938
## 4 2.8563 3.1251 3.3938 5.3788
##   Standard Deviations: 2.1498 2.1642 2.2218 2.3192
```

7. **Standard errors of the fixed effects estimate**

This is a little tricky. By default, the standard errors reported in the t-table in the summary of the output (obtained by `summary` command above) are not quite correct. The correct model-based standard errors need to be computed as below.

```
sqrt(diag(fit.a$varFix))
```

```
##          G          M      G:age      age:M
## 1.18202362 0.98008221 0.09980390 0.08275303
```

8. **Approximate confidence intervals**

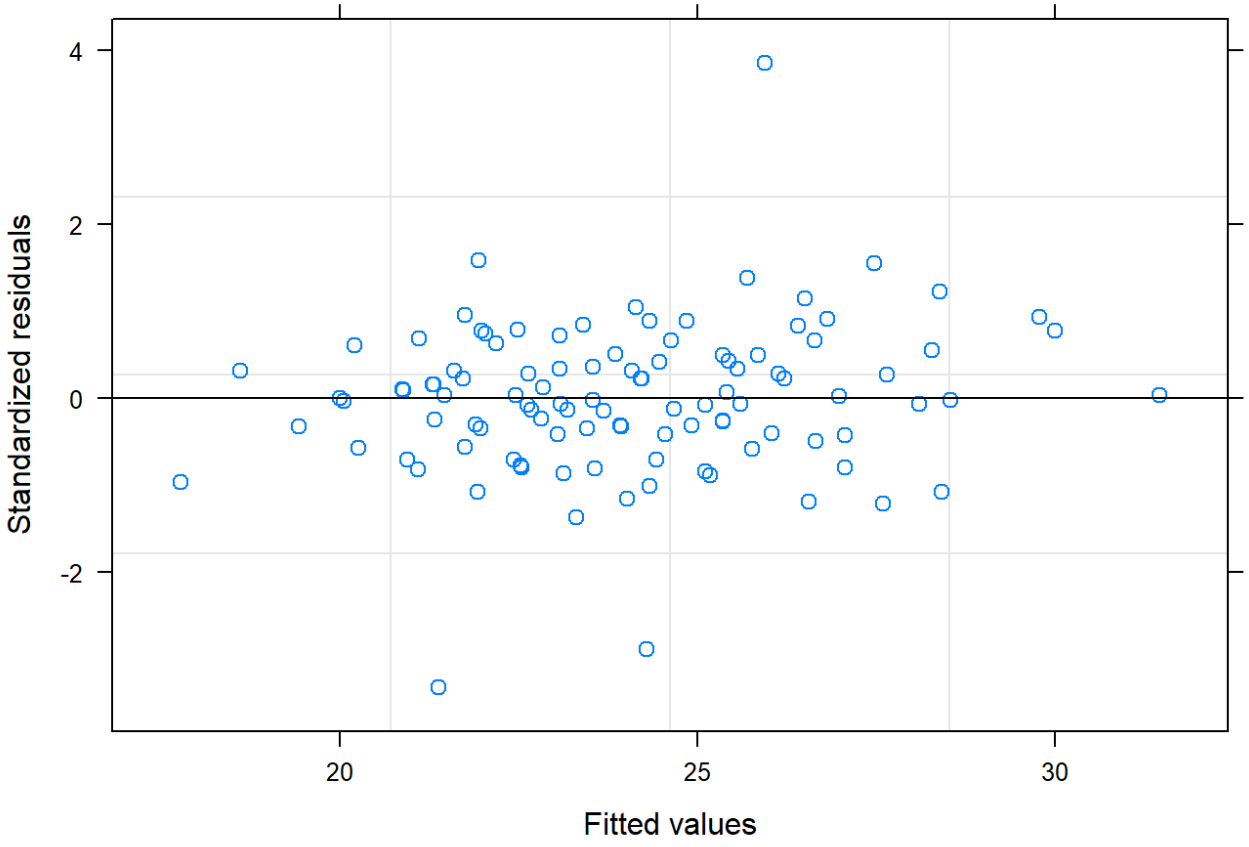
```
# Default is 95% intervals
int <- intervals(fit.a)

# Intervals for fixed effects
int$fixed
```

```
##             lower       est.      upper
## G     14.9383041 17.3727273 19.8071505
## M     14.3221079 16.3406250 18.3591421
## G:age  0.2809294  0.4795455  0.6781616
## age:M  0.6196912  0.7843750  0.9490588
## attr(,"label")
## [1] "Fixed effects:"
```

We can also produce various diagnostic plots:

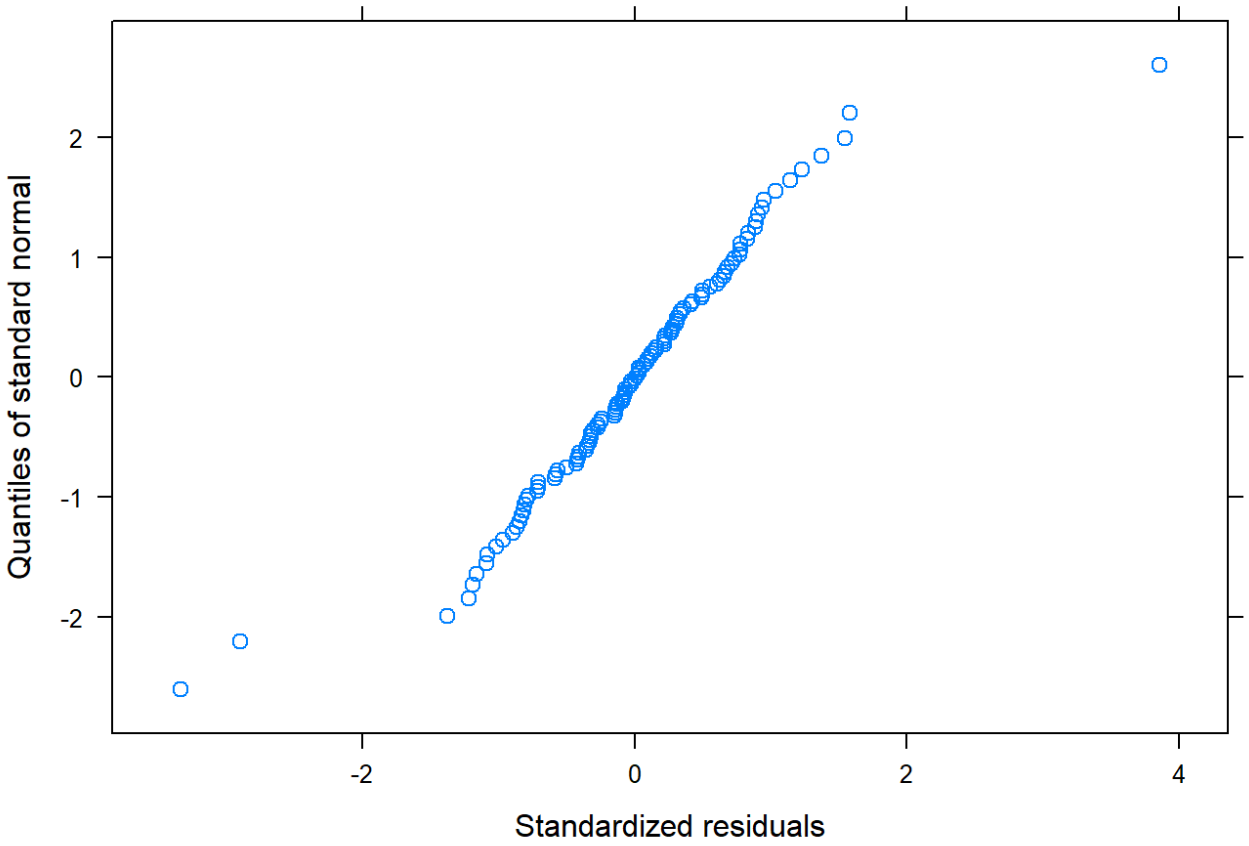
9. **Standardized residuals vs Fitted values**

```
plot(fit.a)
```



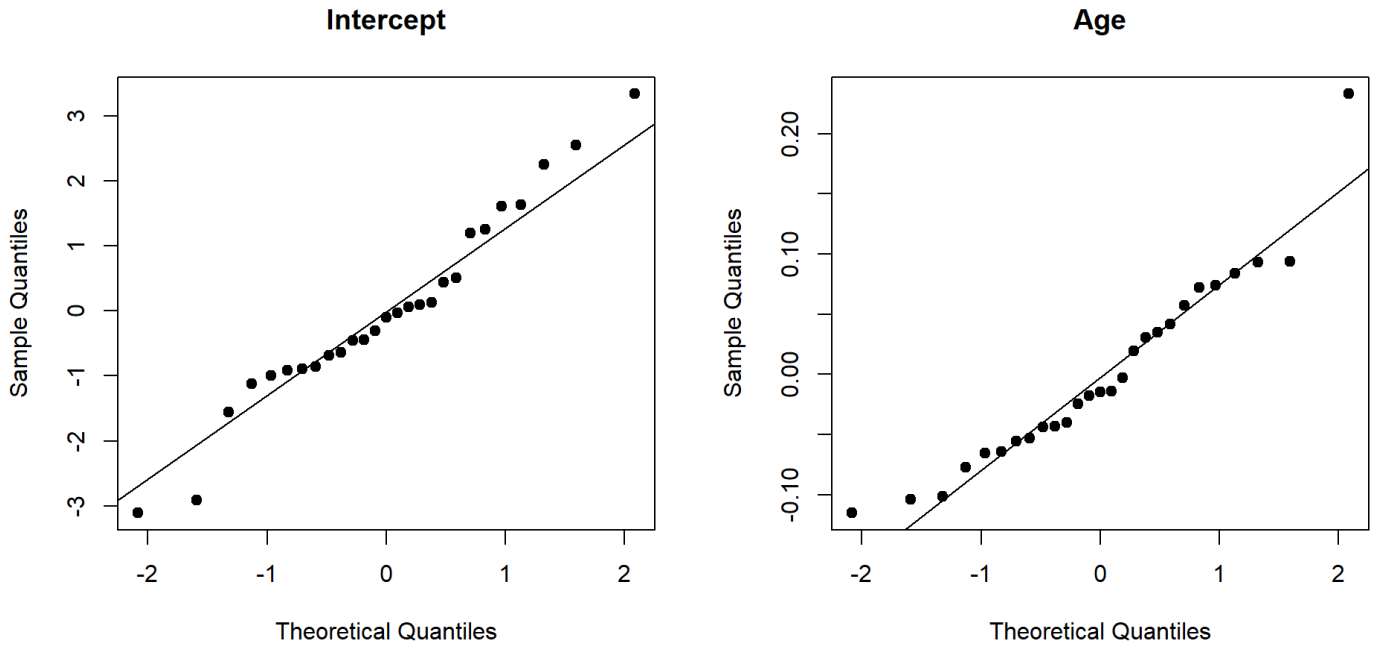
10. **Normal Q-Q plot of the residuals**

```
qqnorm(fit.a)
```



11. **Normal Q-Q plot of the predicted random effects**

```
par(mfrow = c(1,2))
qqnorm(ran.eff[, 1], pch=19, main = "Intercept")
qqline(ran.eff[, 1])
qqnorm(ran.eff[, 2], pch=19, main = "Age")
qqline(ran.eff[, 2])
```



Specifying different random effects covariance structures

There are several options in `lme` regarding covariance structures of the random effects. Generally, the `random` argument can take a formula (as we have done above in fitting (Model A)) or it can take a covariance matrix specification directly. A list possible options is shown below:

- `pdIdent` : a multiple of identity of the form dI.
- `pdDiag` : a diagonal matrix, that is, the random effects are uncorrelated.
- `pdCompSymm` : compound symmetry
- `pdLogChol` , `pdSymm` , `pdNatural` : general positive-definite matrix using the log-Cholesky, SVD and natural (in terms of standard deviations and correlations) parameterizations, respectively.
- `pdBlocked` : a blocked-diagonal matrix, when there are multiple random effects with some are correlated while others are not. For example,

$$D = \begin{pmatrix} D_{11} & D_{12} & 0 \\ D_{21} & D_{22} & 0 \\ 0 & 0 & D_{33} \end{pmatrix}.$$

Thus (Model A), where we have

$$D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix},$$

can be also fitted using the following command.

```
fit.a.alt <- lme(model.formula,
             random = list(id = pdLogChol(form = ~ age)),
             method="ML")
```

Note the `random` argument takes a *list* as value. In the list, we have the statement `id = pdLogChol(form = ~ age)`. T

- he first part `id =` specifies the subjects; the name `id` should be the same as the variable/column name containing the subject index in the data set.
- Next the type of covariance is defined; here we chose `pdLogChol` (unstructured covariance matrix).
- Finally, in the `pdLogChol` command the argument `form` specifies the random effect formula; here `~ age` mean we have a random intercapt (automatically included) and a random slope for age. **Use `~ 0 + age` to just have the random slope but not the random intercept.**

Let us now look at a few examples of random effects covariance structures.

$$\text{(Model B)} \quad y_{ij} = G_i\beta_{0G} + G_i t_{ij}\beta_{1G} + (1-G_i)\beta_{0M} + (1-G_i)t_{ij}\beta_{1M} + b_{0i} + t_{ij}b_{1i} + e_{ij},$$

where

- the **errors** ($\boldsymbol{e}_i$) has **diagonal within-subject covariance** ($\sigma^2 I$), that is same for each gender
- the **random effects** (b_{0i}, b_{1i}) have a **diagonal** 2×2 covariance matrix that is same for each gender

$$D = \begin{pmatrix} D_{11} & 0 \\ 0 & D_{22} \end{pmatrix},$$

We need to specify: `random = list(id = pdDiag(form = ~ age))`

```
fit.b <- lme(model.formula,
             random = list(id = pdDiag(form = ~ age)),
             method="ML")
getVarCov(fit.b)
```

```
## Random effects variance covariance matrix
##             (Intercept)       age
## (Intercept)      2.2492 0.0000000
## age              0.0000 0.0067576
##   Standard Deviations: 1.4997 0.082205
```

$$\text{(Model C)} \quad y_{ij} = G_i(\beta_{0G} + t_{ij}\beta_{1G} + b_{0i,G} + t_{ij}b_{1i,G}) + (1 - G_i)(\beta_{0M} + t_{ij}\beta_{1M} + b_{0i,M} + t_{ij}b_{1i,M}) + e_{ij},$$

where

- the **errors** ($\boldsymbol{e}_i$) has **diagonal within-subject covariance** ($\sigma^2 I$), that is **different for each gender**
- the **random effects** have a **unstructured** 2×2 covariance matrix that is **different for each gender**

$$cov(\begin{bmatrix} b_{0i,G} \\ b_{1i,G} \end{bmatrix}) = D_{girls} = \begin{pmatrix} D_{11,g} & D_{12,g} \\ . & D_{22,g} \end{pmatrix},$$

$$cov(\begin{bmatrix} b_{0i,M} \\ b_{1i,M} \end{bmatrix}) = D_{boys} = \begin{pmatrix} D_{11,b} & D_{12,b} \\ . & D_{22,b} \end{pmatrix},$$

Write the model for the random effects and errors in this situation:

We need to specify: `pdBlocked`

```
# Specify blocks
rlist <- pdBlocked( list(~ 0 + M + M:age,
                         ~ 0 + G + G:age)
                    )

fit.c <- lme(model.formula,
             random = list(id = rlist), # Gender specific D
             weights = varIdent( form = ~ 1 | sex), # Gender specific error variances
             method="ML")

# Random effects covariance D
getVarCov(fit.c)
```

```
## Random effects variance covariance matrix
##               M     M:age         G     G:age
## M      5.64660 -0.282710  0.000000  0.000000
## M:age -0.28271  0.025301  0.000000  0.000000
## G      0.00000  0.000000  2.971700 -0.075389
## G:age  0.00000  0.000000 -0.075389  0.021513
##   Standard Deviations: 2.3763 0.15906 1.7238 0.14667
```

```
# Error covariance for boys
# Individual 1 is male
getVarCov(fit.c, type = "conditional", individuals = 1)
```

```
## id M01
## Conditional variance covariance matrix
##        1      2      3      4
## 1 2.5891 0.0000 0.0000 0.0000
## 2 0.0000 2.5891 0.0000 0.0000
## 3 0.0000 0.0000 2.5891 0.0000
## 4 0.0000 0.0000 0.0000 2.5891
##   Standard Deviations: 1.6091 1.6091 1.6091 1.6091
```

```
# Error covariance for girls
# Individual 17 is male
getVarCov(fit.c, type = "conditional", individuals = 17)
```

```
## id F01
## Conditional variance covariance matrix
##         1       2       3       4
## 1 0.44659 0.00000 0.00000 0.00000
## 2 0.00000 0.44659 0.00000 0.00000
## 3 0.00000 0.00000 0.44659 0.00000
## 4 0.00000 0.00000 0.00000 0.44659
##   Standard Deviations: 0.66827 0.66827 0.66827 0.66827
```

The random coefficients model is a special case of the Linear Mixed Effects Models (LMM). We will discuss this in the next lecture.

Main page: **ST 437/537: Applied Multivariate and Longitudinal Data Analysis (https://maityst537.wordpress.ncsu.edu/)**