Session 5.1: Hierarchical models: longitudinal data

Imperial College London

Learning Objectives

After this session you should be able to:

- Specify hierarchical models for longitudinal data
- Distinguish between random intercept and random slope models and recognise when each is more appropriate
- Be able to run the above models in R-INLA

The topics covered in this lecture are covered in Chapter 4 of Gómez-Rubio (2020a).

Outline

There is huge scope for elaborating the basic hierarchical models discussed in the previous lecture to reflect additional structure and complexity in the data, e.g.

- Adding covariates at different levels of the hierarchy
- Adding further levels to the hierarchy (patients within wards within hospitals, pupils within schools within local authorities, . . .)
- Adding non-nested (cross-classified) levels (patients within GPs crossed with hospitals, . . .)
- **Repeated observations on some/all units (longitudinal data - we will see it in this lecture)**
- Modelling temporal or spatial structure in data, . . . (we will see it from next week)

Outline

What are longitudinal data

What are longitudinal data?

- Arise in studies where individual (or units) are measured repeatedly over time
- For a given individual, observations over time will be typically dependent
- Longitudinal data can arise in various forms:
- continuous or discrete response; discrete response can be binary/binomial, categorical or counts
- equally spaced or irregularly spaced
- same or different time points for each individual
- with or without missing data
- many or few time points, T
- many or few individuals or units, n

Analysing longitudinal data

- There are many different ways to analyse longitudinal data
- This is a very big field, so we have to be selective
- The key feature of longitudinal data is the need to account for the dependence structure of the data
- Two common methods:
  - random effects (hierarchical) models
  - autoregressive models
- Here, we will focus on random effects models

Example: sleep study

Sleep study

- Belenky, Wesensten, Thorne, Thomas, Sing, Redmond, Russo, and Balkin (2003) describes a study of reaction time in patients under sleep deprivation up to 10 days.
- 18 subjects followed for 10 days
- Subjects rated on average reaction time (in ms) for different activities at each measurement

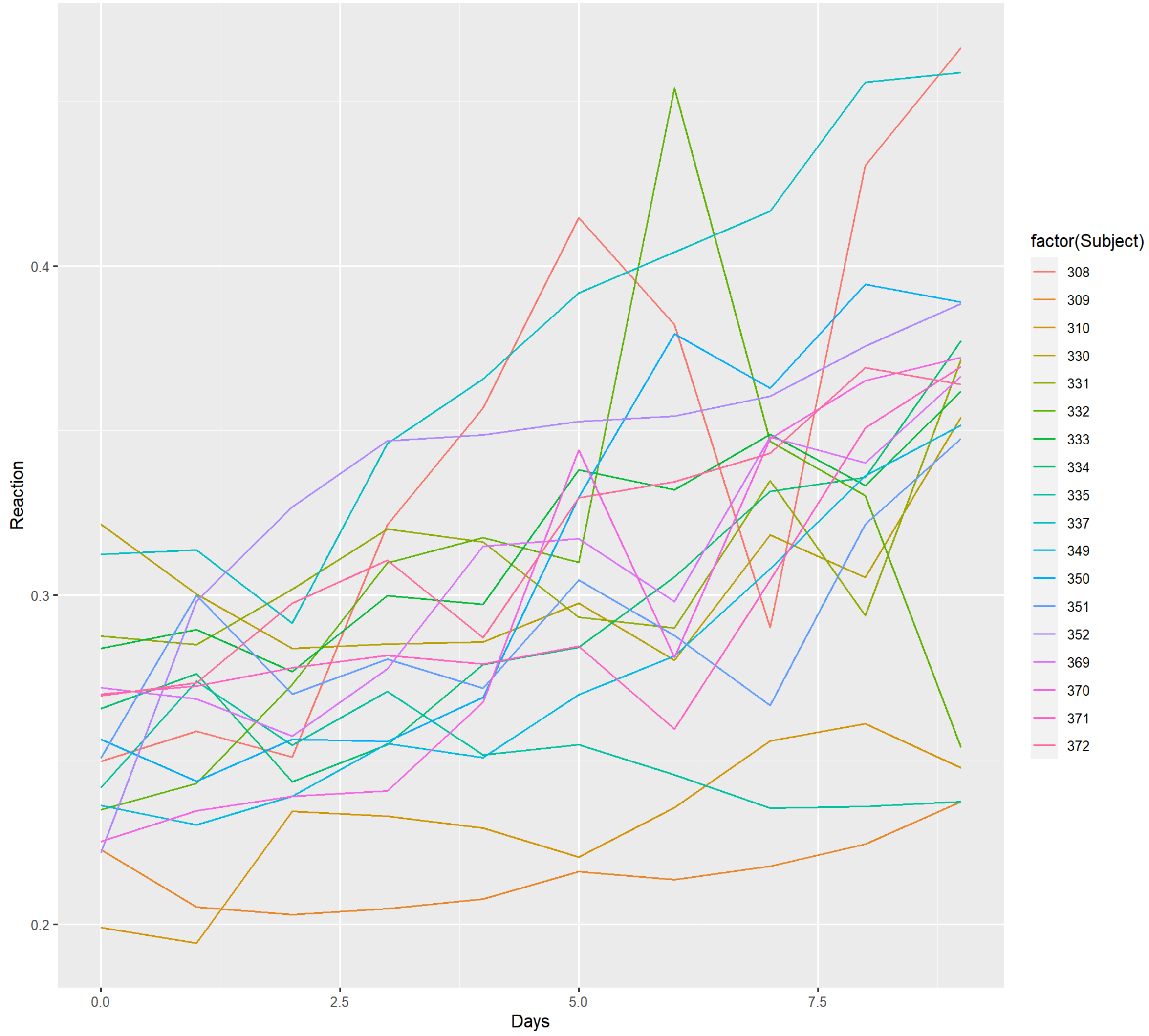
```
> library(lme4)
> data(sleepstudy)
> head(sleepstudy)
```

```
  Reaction Days Subject
1 249.5600    0     308
2 258.7047    1     308
3 250.8006    2     308
4 321.4398    3     308
5 356.8519    4     308
6 414.6901    5     308
```

- Reaction time will be rescaled by dividing by 1000 to have the reaction time in seconds

```
> sleepstudy$Reaction <- sleepstudy$Reaction / 1000
```

Sleep Example: data



Sleep Example: objective

- Study objective: is the length of sleep deprivation a determinant of reaction time?

- The variables we will use are:
  - y: Reaction time (in s)
  - t: Days

- For simplicity we will
  - assume a linear relationship

- The models we will consider are:
  - a non-hierarchical model (standard linear regression) (LM)
  - a hierarchical model with random intercepts (LMM)
  - a hierarchical model with random intercepts and random slopes (LMM2)

Model specification

Sleep Example: a Bayesian (non-hierarchical) linear model (LM)

- Specification: 1. probability distribution for responses:

$$y_{it} \sim \text{Normal}(\mu_{it}, \sigma^2)$$

  - y_{it} = the reaction time for individual i on day $t(\text{days } 0, \ldots, 9)$
  - linear predictor: $\mu_{it} = \alpha + \beta t$
  - t = the day of the measurement

2. In this model no account is taken of the repeated structure (observations are nested within individuals)

3. Assume vague priors for all parameters:

$$\alpha, \beta \sim \text{Normal}(0, 10000)$$
$$\frac{1}{\sigma^2} \sim \text{Gamma}(1, 0.001)$$

Sleep Example: a Bayesian hierarchical linear model

- Modify LM to allow a separate intercept for each individual:

$$\begin{aligned} y_{it} &\sim \text{Normal}(\mu_{it}, \sigma^2) \\ \mu_{it} &= \alpha_i + \beta t \end{aligned}$$

We are assuming that *conditionally* on α_i, $\{y_{it}, t = 0, \ldots, 9\}$ are independent

- Assume that all the $\{\alpha_i\}$ follow a *common* prior distribution, e.g.

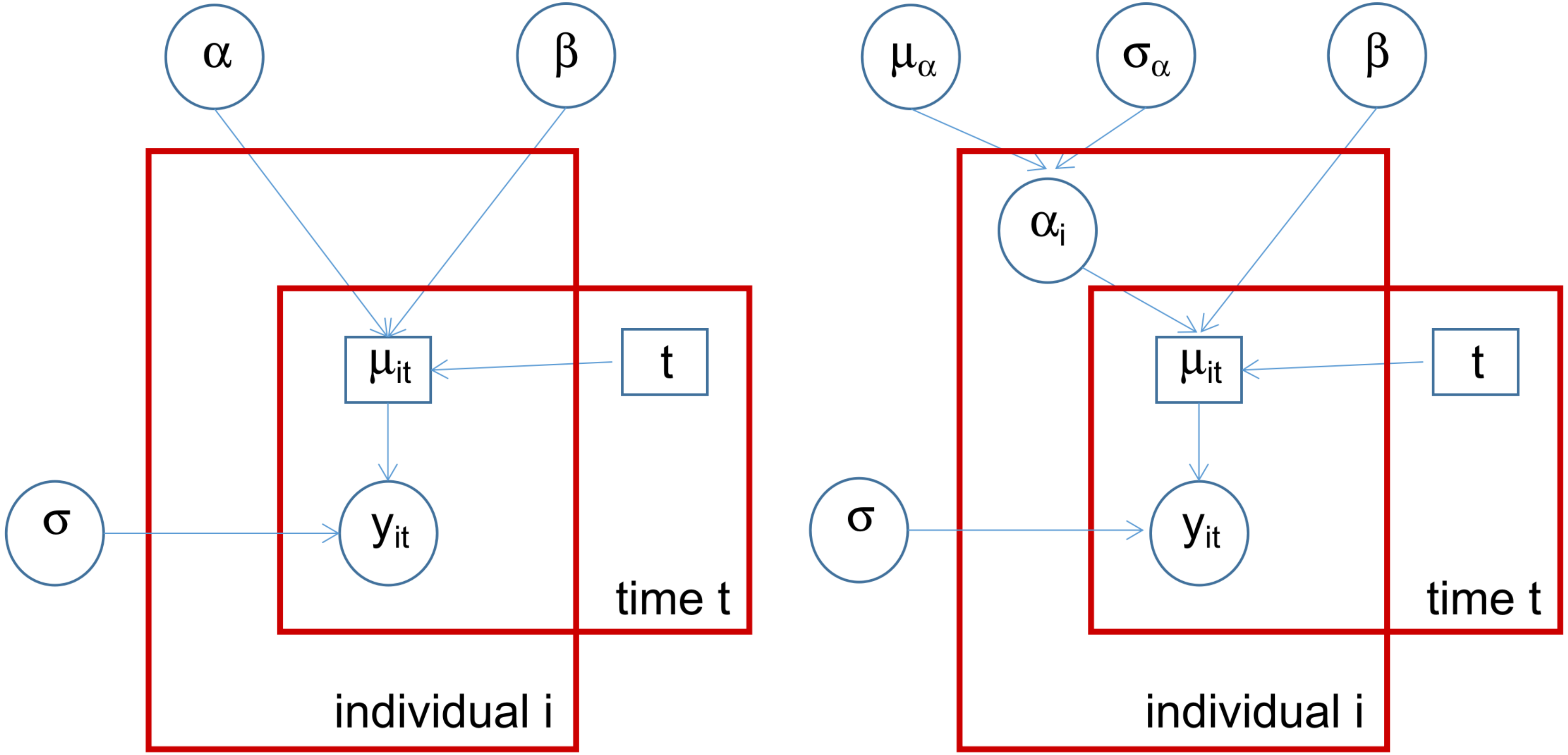
$$\alpha_i \sim \text{Normal}(\mu_\alpha, \sigma^2_\alpha) \quad i = 1, \ldots, 18$$

- Here we are assuming exchangeability between all the individuals
- We may then assume vague priors for the *hyperparameters* of the population distribution:

$$\begin{aligned} \mu_\alpha &\sim \text{Normal}(0, 10000) \\ \frac{1}{\sigma^2_\alpha} &\sim \text{Gamma}(1, 0.001) \end{aligned}$$

- This is an example of a *Hierarchical LM* or *Linear Mixed Model (LMM)* or *Random Intercepts* model

Comparing the two models



(left): LM

(right): LMM - random intercept

HAMD example: long vs wide format

- In `R-INLA` is extremely easy to work with longitudinal data, as long as the dataset is in *long format*

Long format | Wide format

```
    Reaction Days Subject
1  0.2495600    0     308
2  0.2587047    1     308
3  0.2508006    2     308
4  0.3214398    3     308
5  0.3568519    4     308
6  0.4146901    5     308
7  0.3822038    6     308
8  0.2901486    7     308
9  0.4305853    8     308
10 0.4663535    9     308
```

- Note that a score is present only if the corresponding subject has been observed at that time point

HAMD example: long vs wide format

- In `R-INLA` is extremely easy to work with longitudinal data, as long as the dataset is in *long format*

Long format | **Wide format**

```
   Subject Reaction.0 Reaction.1 Reaction.2 Reaction.3 Reaction.4
1      308  0.2495600  0.2587047  0.2508006  0.3214398  0.3568519
11     309  0.2227339  0.2052658  0.2029778  0.2047070  0.2077161
21     310  0.1990539  0.1943322  0.2343200  0.2328416  0.2293074
31     330  0.3215426  0.3004002  0.2838565  0.2851330  0.2857973
41     331  0.2876079  0.2850000  0.3018206  0.3201153  0.3162773
51     332  0.2348606  0.2428118  0.2729613  0.3097688  0.3174629
61     333  0.2838424  0.2895550  0.2767693  0.2998097  0.2971710
71     334  0.2654731  0.2762012  0.2433647  0.2546723  0.2790244
81     335  0.2416083  0.2739472  0.2544907  0.2708021  0.2514519
91     337  0.3123666  0.3138058  0.2916112  0.3461222  0.3657324
```

- Note that with this format should a subject not have measurement for the entire set of days, R would pad these out with NAs

Sleep example: R-INLA code for LM

Code Plot

```
> library(INLA)
> formula_LM <- Reaction ~ Days
> lm <- inla(formula_LM, data=sleepstudy, family="gaussian", control.compute = list(waic=TRUE))
>
> #Plot
> ggplot(data = sleepstudy, aes(x=Days, y=Reaction)) + geom_point(aes(colour = Subject), alpha = .9) +
+   #geom_smooth(method="lm")
+   geom_abline(intercept=lm$summary.fixed[1,1], slope=lm$summary.fixed[2,1]) + labs(title = "Linear Mode
+   theme(legend.position = "none")
```

Sleep example: R-INLA code for LM

Code | Plot

- all the subjects share the same intercept and slope

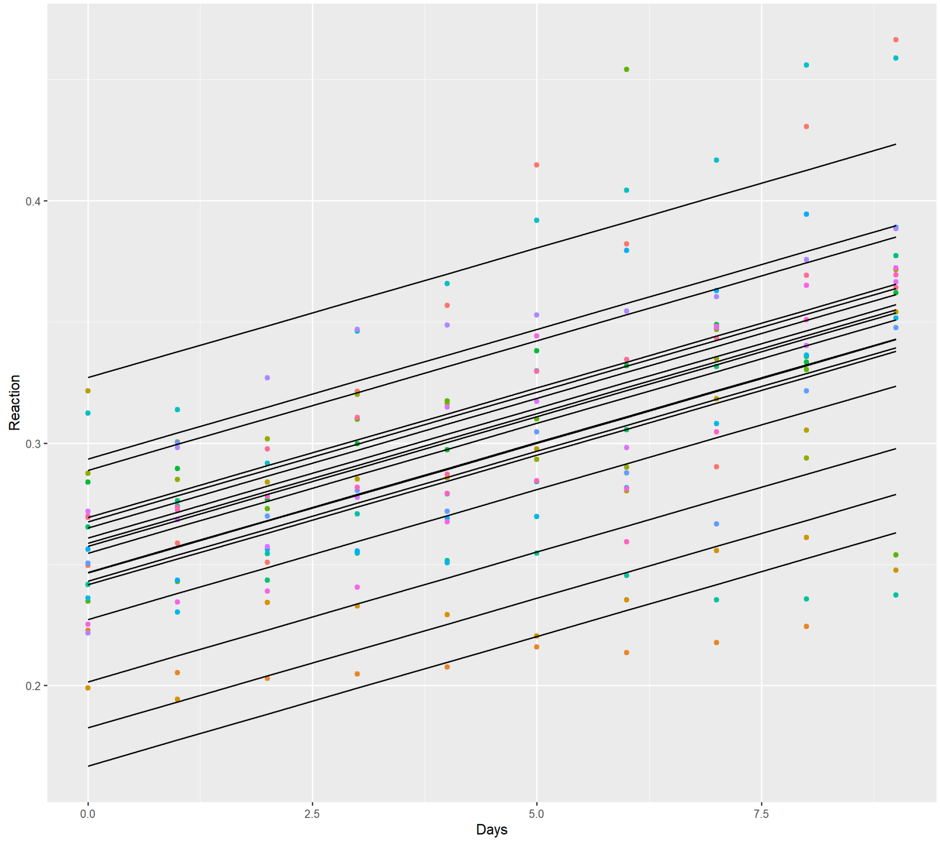
Sleep example: R-INLA code for LMM

Code Plot

```
> library(INLA)
> formula_LMM <- Reaction ~ -1 + Days + f(Subject,model="iid")
> lmm <- inla(formula_LMM, data=sleepstudy, family="gaussian", control.compute = list(waic=TRUE))
> #Plot
> p <- ggplot(data = sleepstudy, aes(x=Days, y=Reaction, group=Subject))
> p + geom_point(aes(colour = Subject), alpha = .9) + geom_abline(slope=lmm$summary.fixed[2,1], intercept
+   theme(legend.position = "none")
```

Sleep example: R-INLA code for LMM

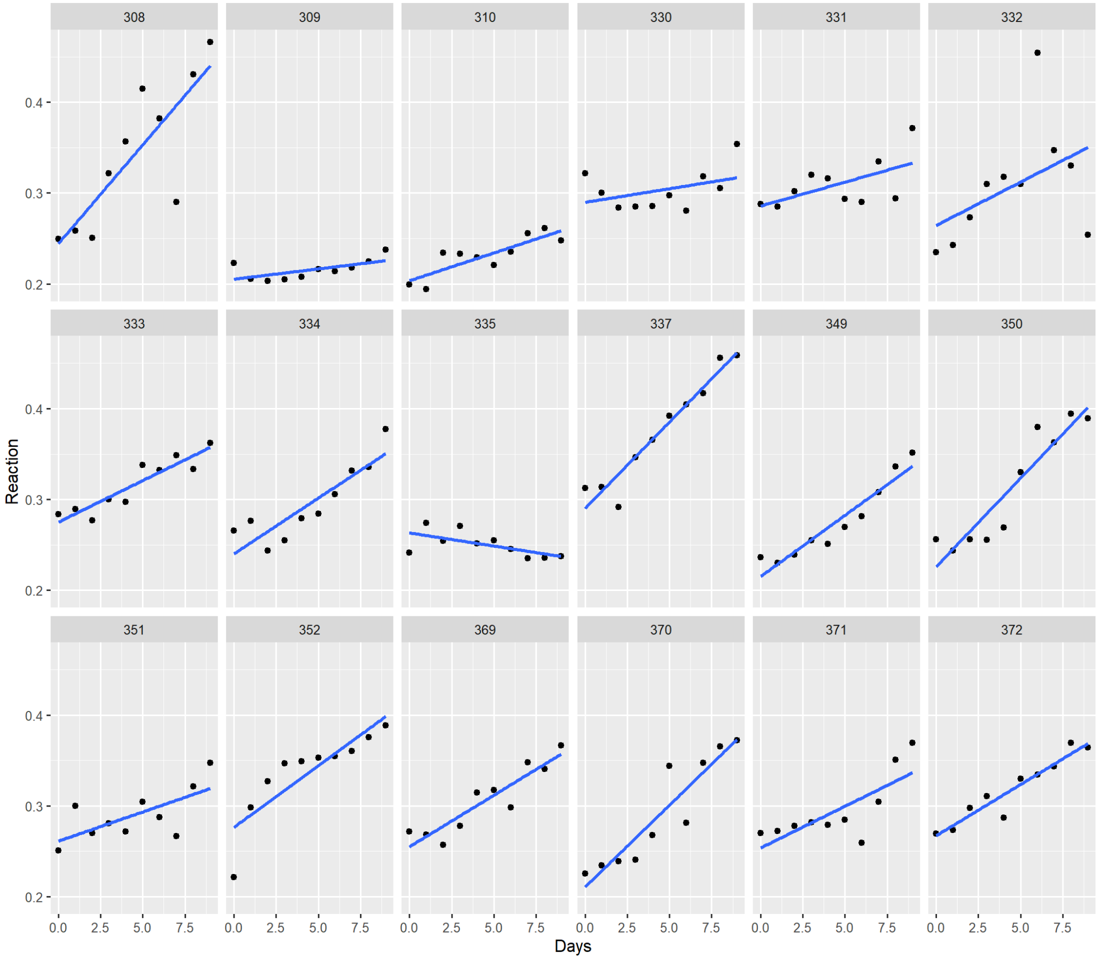
Code | Plot



- each individual has a different regression line
- but all individuals have the same slope (parallel lines)

Interpretation

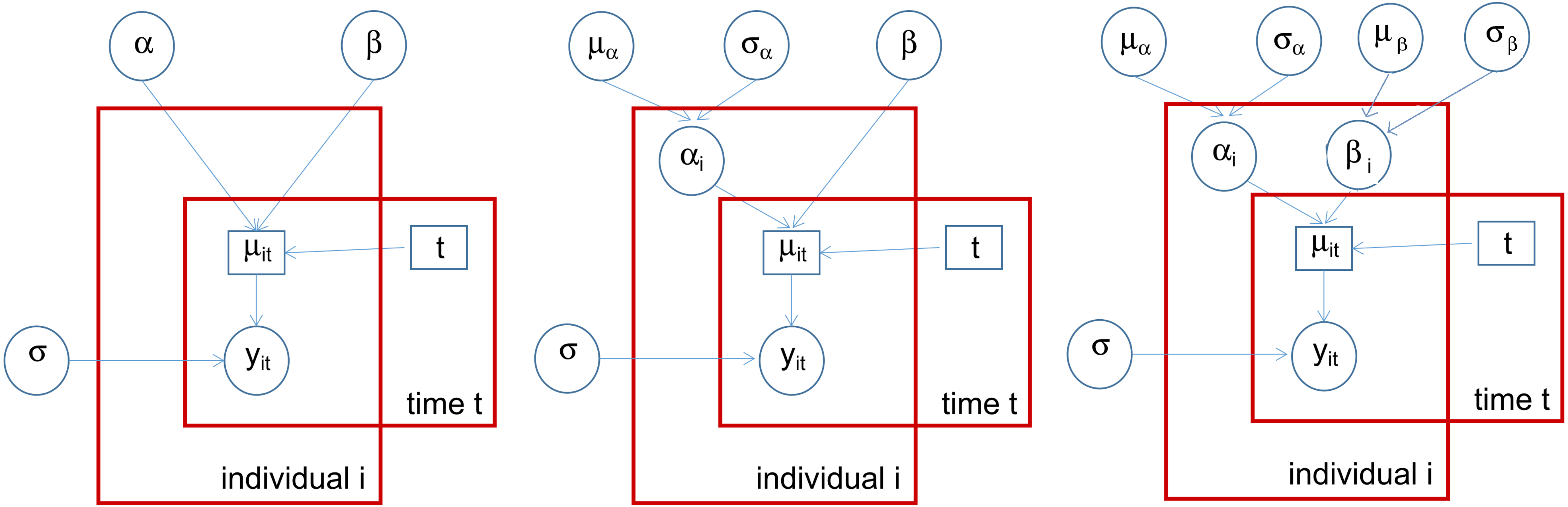
Sleep Example: revisiting the data



- blue: regression line for each subject;
- black: regression line from the LMM

- Note that some of the subjects do not fit well in the *parallel lines* scenario

- So we add random slopes to the hierarchical model



(left): LM

(center): LMM - random intercept

(right): LMM - random intercept and slope

Sleep Example: adding random slopes

- Modify LMM to allow a separate slope for each individual:

$$
\begin{aligned}
y_{it} &\sim \text{Normal}(\mu_{it}, \sigma^2) \\
\mu_{it} &= \alpha_i + \beta_i t
\end{aligned}
$$

- As for the $\{\alpha_i\}$, assume that the $\{\beta_i\}$ follow *common* prior distributions with vague priors on their *hyperparameters*

Sleep Example: adding random slopes

- Modify LMM to allow a separate slope for each individual:

$$y_{it} \sim \text{Normal}(\mu_{it}, \sigma^2)$$
$$\mu_{it} = \alpha_i + \beta_i t$$

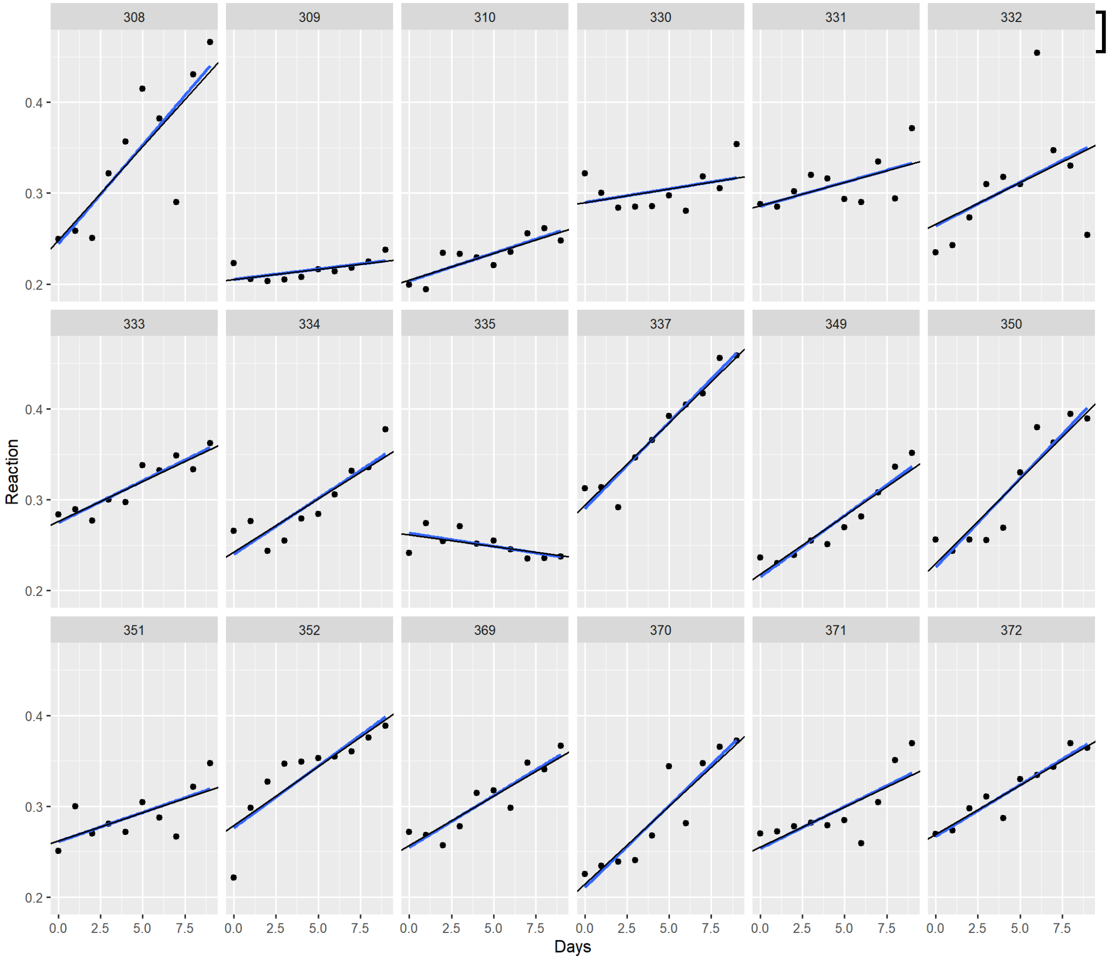
- As for the $\{\alpha_i\}$, assume that the $\{\beta_i\}$ follow *common* prior distributions with vague priors on their *hyperparameters*

- In `R-INLA`

```
> Subject2<-sleepstudy$Subject
> formula_LMM2 <- Reaction ~ -1 + f(Subject, model="iid", constr=FALSE) + f(Subject2, Days, model="iid",
> lmm2 <- inla(formula_LMM2, data=sleepstudy, family="gaussian", control.compute = list(waic=TRUE))
> #To access the summaries of the beta_i
> lmm2$summary.random$Subject2[1:5,]
```

```
   ID        mean          sd    0.025quant    0.5quant  0.975quant        mode          kld
1 308 0.020670507 0.002755625  0.0152490733 0.020674366 0.026070095 0.020674410 3.306600e-09
2 309 0.002253256 0.002725803 -0.0030979042 0.002252969 0.007606057 0.002252971 3.069934e-09
3 310 0.005886435 0.002727833  0.0005291753 0.005886891 0.011241114 0.005886895 3.063615e-09
4 330 0.003006663 0.002726323 -0.0023450831 0.003006223 0.008360930 0.003006226 3.073045e-09
```

HAMD Example: random intercepts and slopes



- LMM with random intercepts only:
  - each individual has a different regression line
  - but for each treatment, only intercept varies by individual
- LMM with random intercepts and random slopes:
  - now intercepts and slopes both vary
  - better fit for each individual

Which model is the best?

- Comparing WAIC

```
> #Linear model
> lm$waic$waic
```

```
[1] -580.1704
```

```
> #Random intercept
> lmm$waic$waic
```

```
[1] -716.7855
```

```
> #Random intercept and slope
> lmm2$waic$waic
```

```
[1] -763.4362
```

- It looks like the model with random intercept and slope fits the data best

- We can look at the variance hyperparameters of the lmm2 model to see how variable the intercept and slope are across individuals

```
> #Random slope model
> #Variability of the random intercept
> inla.emarginal(fun=function(x) 1/x,
+                lmm2$marginals.hyperpar[[2]])
```

```
[1] 0.06369514
```

```
> #Variability of the random slope
> inla.emarginal(fun=function(x) 1/x,
+                lmm2$marginals.hyperpar[[3]])
```

```
[1] 0.0001472959
```

- Some variability for both intercept and slope.

References

Belenky, G., N. J. Wesensten, D. R. Thorne, et al. (2003). "Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study". In: *Journal of sleep research* 12.1, pp. 1-12.

Gómez-Rubio, V. (2020a). *Bayesian inference with INLA*. CRC Press.