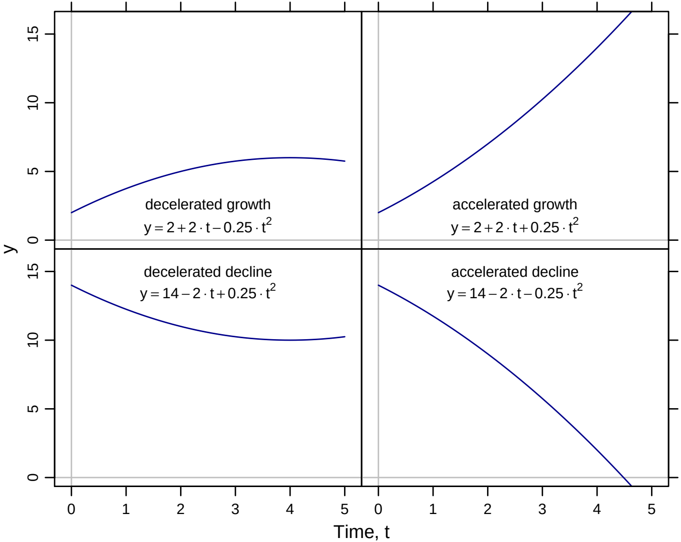
Growth curve models

Nora Wickelmaier

Last modified: January 8, 2025

Quadratic time trends

- A lot of times the assumption of a linear time trend is too simple
- Change is not happening unbraked linearly but flattens out



Quadratic time trends

- Quadratic regression model

$$
\begin{aligned}
y_{ij} &= b_{0i} + b_{1i}\, t_{ij} + b_{2i}\, t_{ij}^2 + \varepsilon_{ij} \\
&= b_{0i} + (b_{1i} + b_{2i}\, t_{ij}) t_{ij} + \varepsilon_{ij}
\end{aligned}
$$

- The linear change depends on time t

$$
\frac{\partial y}{\partial t} = b_{1i} + 2 b_{2i}\, t
$$

- The intercept $t = -b_{1i}/(2b_{2i})$ is the point in time when a positive (negative) trend becomes negative (postive)

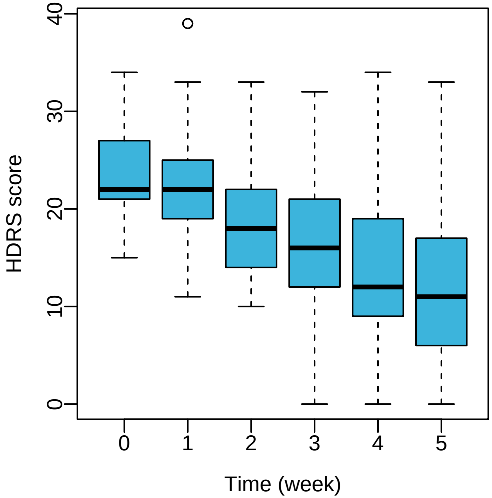
Outline

1 Depression and Imipramin

Depression and Imipramin (Reisby et al., 1977)

- Reisby et al. (1977) studied the effect of Imipramin on 66 inpatients treated for depression
- Depression was measured with the Hamilton depression rating scale (HDRS)
- Additionally, the concentration of Imipramin and its metabolite Desipramin was measured in their blood plasma
- Patients were classified into endogenous and non-endogenous depressed
- Depression was measured weekly for 6 time points; the effect of the antidepressant was observed starting at week 2 for four weeks

Descriptive statistics



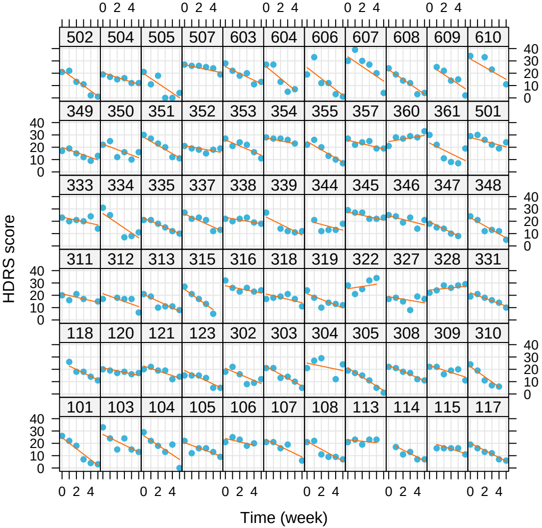
HDRS score

t	W0	W1	W2	W3	W4	W5
M	23.44	21.84	18.31	16.42	13.62	11.95
SD	4.53	4.70	5.49	6.42	6.97	7.22
n	61	63	65	65	63	58

Empirical correlation matrix of HDRS score

	W0	W1	W2	W3	W4	W5
Week 0	1	.49	.41	.33	.23	.18
Week 1	.49	1	.49	.41	.31	.22
Week 2	.41	.49	1	.74	.67	.46
Week 3	.33	.41	.74	1	.82	.57
Week 4	.23	.31	.67	.82	1	.65
Week 5	.18	.22	.46	.57	.65	1

Predictions random slope model



$$y_{ij} = \beta_0 + \beta_1 time + \upsilon_{0i} + \upsilon_{1i} time + \varepsilon_{ij}$$

with

$$\begin{pmatrix} \upsilon_{0i} \\ \upsilon_{1i} \end{pmatrix} \overset{iid}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0 \upsilon_1} \\ \sigma_{\upsilon_0 \upsilon_1} & \sigma^2_{\upsilon_1} \end{pmatrix} \right)$$

$$\varepsilon_i \overset{iid}{\sim} N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i})$$

Model with quadratic trend

- Model with quadratic individual and quadratic group trend

$$y_{ij} = \beta_0 + \beta_1\, t_{ij} + \beta_2\, t_{ij}^2 + \upsilon_{0i} + \upsilon_1\, t_{ij} + \upsilon_2\, t_{ij}^2 + \varepsilon_{ij}$$

with

$$\begin{pmatrix} \upsilon_{0i} \\ \upsilon_{1i} \\ \upsilon_{2i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\upsilon_0}^2 & \sigma_{\upsilon_0\upsilon_1} & \sigma_{\upsilon_0\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma_{\upsilon_1}^2 & \sigma_{\upsilon_1\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_2} & \sigma_{\upsilon_1\upsilon_2} & \sigma_{\upsilon_2}^2 \end{pmatrix} \right) \text{ i.i.d.}$$

$$\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}) \text{ i.i.d.}$$

Depression and Imipramin

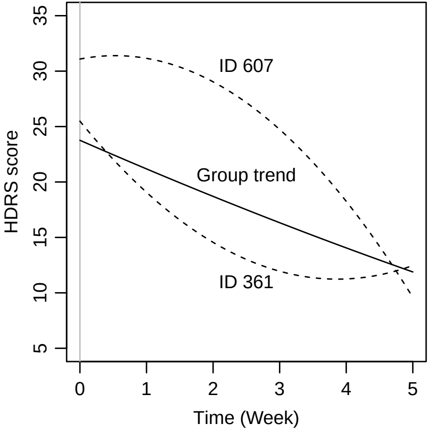
```r
dat      <- read.table("data/reisby.dat", header = TRUE)
dat$id   <- factor(dat$id)
dat$diag <- factor(dat$diag, levels = c("nonen", "endog"))
dat      <- na.omit(dat)     # drop missing values

# random intercept model
lme1 <- lmer(hamd ~ week + (1 | id), data = dat, REML = FALSE)

# random slope model
lme2 <- lmer(hamd ~ week + (week | id), data = dat, REML = FALSE)

# model with quadratic time trend
lme3 <- lmer(hamd ~ week + I(week^2) + (week + I(week^2) | id),
             data = dat, REML = FALSE)
```

Model predictions

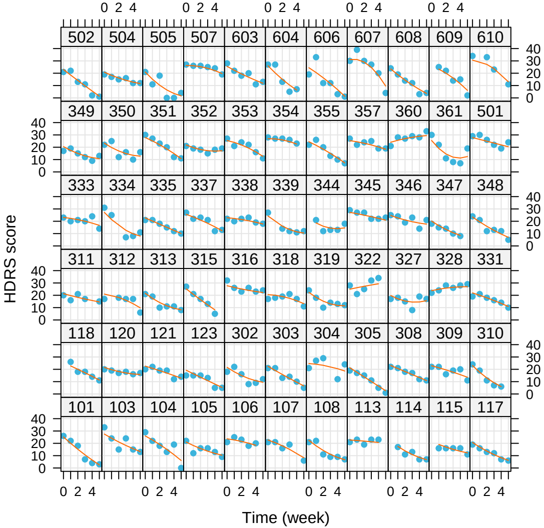


- Averaged over persons an approximately linear trend is obtained, $\hat{\beta}_1 = -2.63$, $\hat{\beta}_2 = 0.05$
- Some of the predicted individual trends are strongly nonlinear

- Test against a model without individual quadratic trends

$$\mathrm{H}_0\colon\ \sigma^2_{\upsilon_2} = \sigma_{\upsilon_0\upsilon_2} = \sigma_{\upsilon_1\upsilon_2} = 0 \qquad G^2(3) = 10.98,\ p = .012$$

Model predictions




```r
xyplot(
  hamd + predict(lme3)
      ~ week | id,
  data = dat,
  type = c("p", "l", "g"),
  pch = 16,
  distribute.type = TRUE,
  ylab = "HDRS score",
  xlab = "Time (Week)")
```

Implied marginal covariance matrix

Predicted

$$\mathbf{Z}_i \hat{\boldsymbol{\Sigma}}_\upsilon \mathbf{Z}_i' + \hat{\sigma}^2 \mathbf{I}_{n_i} = \begin{pmatrix} 20.96 & 9.41 & 8.16 & 6.68 & 4.98 & 3.06 \\ 9.41 & 23.86 & 15.57 & 16.08 & 14.88 & 11.97 \\ 8.16 & 15.57 & 31.07 & 23.11 & 23.26 & 20.98 \\ 6.68 & 16.08 & 23.11 & 38.31 & 30.12 & 30.09 \\ 4.98 & 14.88 & 23.26 & 30.12 & 45.98 & 39.29 \\ 3.06 & 11.97 & 20.98 & 30.09 & 39.29 & 59.11 \end{pmatrix}$$

Observed

$$\widehat{Cov}(\mathbf{y}_i) = \begin{pmatrix} 20.55 & 10.11 & 10.14 & 10.09 & 7.19 & 6.28 \\ 10.11 & 22.07 & 12.28 & 12.55 & 10.26 & 7.72 \\ 10.14 & 12.28 & 30.09 & 25.13 & 24.63 & 18.38 \\ 10.09 & 12.55 & 25.13 & 41.15 & 37.34 & 23.99 \\ 7.19 & 10.26 & 24.63 & 37.34 & 48.59 & 30.51 \\ 6.28 & 7.72 & 18.38 & 23.99 & 30.51 & 52.12 \end{pmatrix}$$

Centering variables

- If multiples of the time variables (t, t^2, t^3, etc.) are entered into the regression equation, multicollinearity can become a problem
- For example, $t = 0, 1, 2, 3$ and $t^2 = 0, 1, 4, 9$ correlate almost perfectly
- By centering the variables, this problem can be diminished: $(t - \bar{t}) = -1.5, -0.5, 0.5, 1.5$ and $(t - \bar{t})^2 = 2.25, 0.25, 0.25, 2.25$ are uncorrelated
- By centering variables the interpretation of the intercept in a linear model changes:
  - Uncentered intercepts represent the difference to the first time point ($t = 0$)
  - Centered intercepts represent the difference after half of the time

Analysis with centered time variable

```
dat$week_c <- dat$week - mean(dat$week)
cor(dat$week, dat$week^2)        # 0.96
cor(dat$week_c, dat$week_c^2)    # 0.01

# random slope model
lme2c <- lmer(hamd ~ week_c + (week_c | id), data = dat, REML = FALSE)

# model with quadratic time trend
lme3c <- lmer(hamd ~ week_c + I(week_c^2) + (week_c + I(week_c^2) | id),
              data = dat, REML = FALSE)
```

- When comparing the estimated parameters, it becomes obvious that not only the intercept changes but the estimates for the (co)variances do as well
- Why?

Analysis with centered time variable

```
dat$week_c <- dat$week - mean(dat$week)
cor(dat$week, dat$week^2)       # 0.96
cor(dat$week_c, dat$week_c^2)   # 0.01

# random slope model
lme2c <- lmer(hamd ~ week_c + (week_c | id), data = dat, REML = FALSE)

# model with quadratic time trend
lme3c <- lmer(hamd ~ week_c + I(week_c^2) + (week_c + I(week_c^2) | id),
              data = dat, REML = FALSE)
```

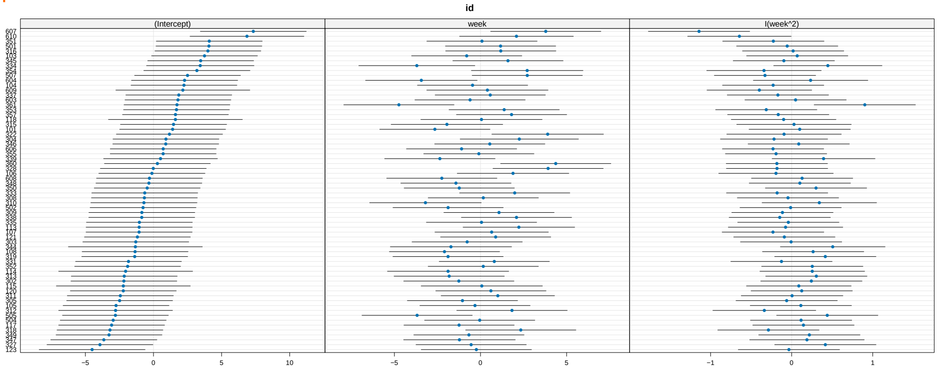
- When comparing the estimated parameters, it becomes obvious that not only the intercept changes but the estimates for the (co)variances do as well
- Why? Be sure to make an informed choice when centering your variables!

Investigating random effects structure

- In order to get a better understanding of the necessary random effects it might be a good idea to take a closer look at them
- Two plots often used are the so-called caterpillar and shrinkage plots
- Play around with different models and compare how, e. g., the caterpillar plots change with and without covariances in the model!

Investigating random effects structure

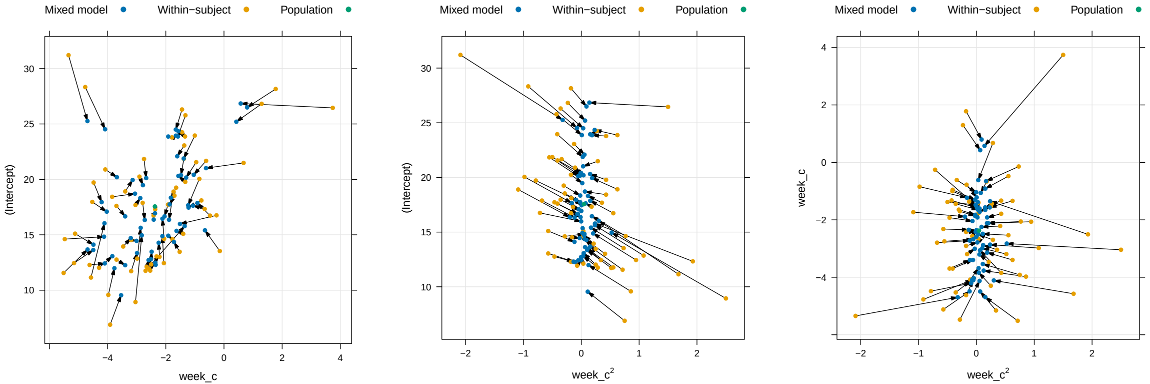
Caterpillar plot




```r
library(lattice)
dotplot(ranef(lme3), scales = list( x = list(relation = "free")))$id
```

Investigating random effects structure

Shrinkage plots




```
# model without covariances
lme4 <- lmer(hamd ~ week_c + I(week_c^2) + (week_c + I(week_c^2) || id),
             data = dat, REML = FALSE)
```

Investigating random effects structure

- The shrinkage plots suggest that we do not need the quadratic effect for each subject (on the population level it does not show anyway)
- Let us dive a little deeper and look at the profiles for the random effects
- We see that the parameter for the individual quadratic effect (σ_3) cannot be estimated sensibly

```
pm4 <- profile(lme4, which = "theta_")

xyplot(log(pm4))
densityplot(log(pm4))
splom(log(pm4))

confint(pm4)
# --> a model with the covariances will not even be estimated by the
# profile function in R
```

Intermediate conclusions

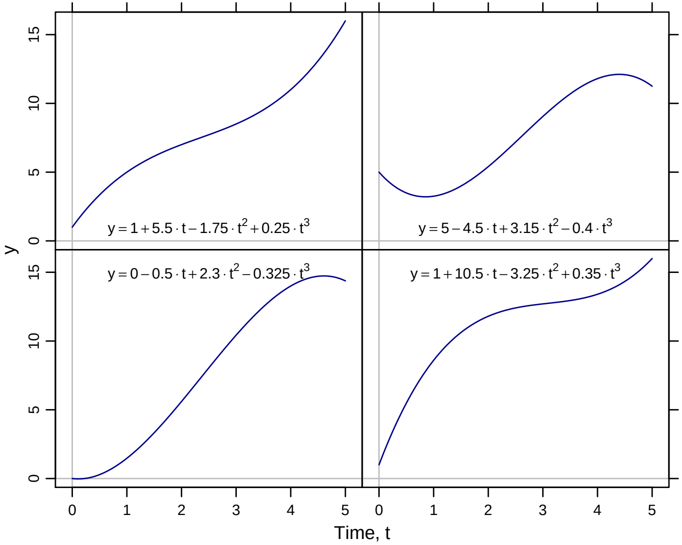
- After closer investigation of the random effects structure, I am not so sure that the story with the quadratic individual trends still holds
- This is only important if I want to interpret the random effects
- If I only include them for modeling dependency in my data, this is not of so much relevance
- In this case, I am (usually) conducting a more conservative test (more on this next session)

```
lme5 <- lmer(hamd ~ week_c + (week_c || id),
             data = dat, REML = FALSE)
lme6 <- lmer(hamd ~ week_c + I(week_c^2) + (week_c + I(week_c^2) || id),
             data = dat, REML = FALSE)
anova(lme5, lme6)
```

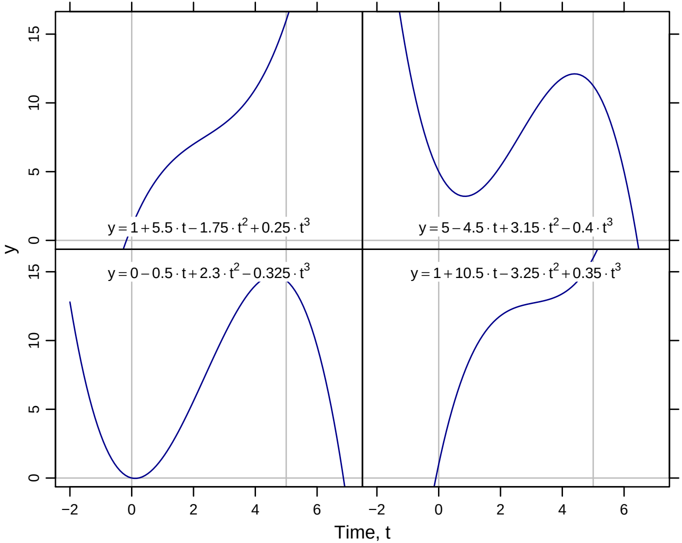
Higher-order polynomials

- Nonlinear time trends can be modelled in a flexibel and parsimonious way by using higher-order polynomials
- For example, saddle or reversal points in a time trend can be described
- Polynomials have the advantage that the regression model stays linear in its parameters
- They have the disadvantage that extrapolated values can quickly be outside of a range that can still be interpreted in a meaningful way

Cubic time trends



Polynomial regression: Extrapolation



2 Vocal charades

Vocal charades (Winter & Wieling, 2016)

This (simulated) example from linguistics is taken from Winter and Wieling (2016)

- Participants play a game of ‘vocal charades’
- At each round, a participant has to vocalize a meaning to the partner (e. g., ‘ugly’) without using language (e. g., through grunting or hissing)
- The partner has to guess the meaning of the vocalization
- This game is played repeatedly with the finding that over time, a dyad converges on a set of nonlinguistic vocalizations that assure a high degree of intelligibility between the two participants in the dyad

Example: Growth curve model

- Initially, participants may be struggling with the task and explore very different kinds of vocalizations
- Over time, they may converge on a more stable set of iconic vocalizations, that is vocalizations that resemble the intended referent (e. g., a high-pitched sound for ‘attractive’ and a low-pitched sound for ‘ugly’)
- Finally, after even more time, the dyad may conventionalize to idiosyncratic patterns that deviate from iconicity and become increasingly arbitrary

Example: Growth curve model

- 100 observations of three variables (simulated data set)

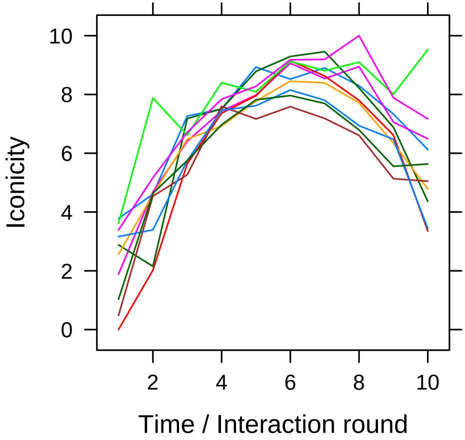
Variable	Description
dyad	different pairs of subjects playing the vocal charades game
t	sequential rounds for which the vocal charades game was played
iconicity	iconicity measure

- For better interpretation, t will be centered

```
dat <- read.csv("data/lzw003_supplementary-data/example2_dyads.csv")
dat$dyad <- as.factor(dat$dyad)

dat$t_c <- dat$t - mean(dat$t)
```

Visualization of data




```
xyplot(
  iconicity ~ t, dat,
  groups = dyad,
  type = "l",
  xlab = "Time/Interaction round",
  ylab = "Iconicity")
```

Mixed-effects model with quadratic trend

- We will now consider a model with uncorrelated random effects

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + \upsilon_{0i} + \upsilon_{1i} t_{ij} + \upsilon_{2i} t_{ij}^2 + \varepsilon_{ij}$$

with

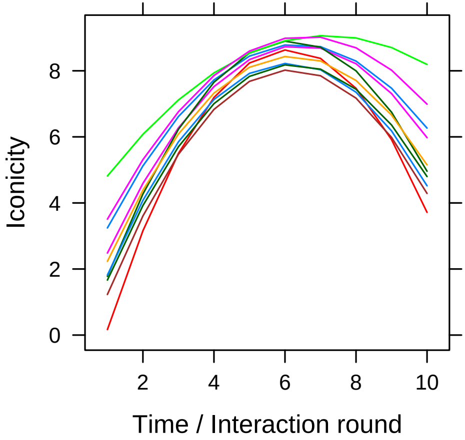
$$\begin{pmatrix} \upsilon_{0i} \\ \upsilon_{1i} \\ \upsilon_{2i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\upsilon_0}^2 & 0 & 0 \\ 0 & \sigma_{\upsilon_1}^2 & 0 \\ 0 & 0 & \sigma_{\upsilon_2}^2 \end{pmatrix} \right) \text{ i.i.d.}$$

$$\varepsilon_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}) \text{ i.i.d.}$$

- This model is fitted by

```
gcm1 <- lmer(iconicity ~ t_c + I(t_c^2) +
  (1 | dyad) + (0 + t_c | dyad) + (0 + I(t_c^2) | dyad),
  data = dat, REML = FALSE)
```

Visualization of model predictions



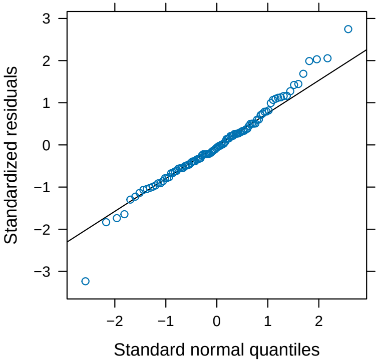

```
xyplot(
  predict(gcm1) ~ t, dat,
  groups = dyad,
  type = "l",
  xlab = "Time/Interaction round",
  ylab = "Iconicity")
```

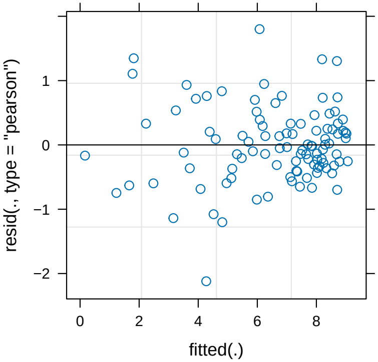
Interpretation of results

- There are now two slopes, one for the effect of linear time (t_c, $\beta_1 = 0.35$) and one for the effect of quadratic time (t_c^2, $\beta_2 = -0.23$), both of which are allowed to differ by dyad ($\sigma_{\upsilon_{1i}} = 0.05$ and $\sigma_{\upsilon_{2i}} = 0.06$)
- The negative value for the quadratic term indicates the inverse U-shape
- The point of reversal is $t = \bar{t} + \frac{-\hat{\beta}_1}{2 \cdot \hat{\beta}_2} = 5.5 + \frac{-.35}{2 \cdot (-.22)} = 6.29$
- The model assumes that the random intercept and slopes are all uncorrelated

Check model assumptions

```
hist(residuals(gcm1))    # o.k.
qqmath(gcm1)             # o.k.
plot(gcm1)               # o.k.
```





Exercise[1]

- Load the data set `elstongrizzle.dat` into R; data are from a dental study measuring the lengths of the ramus of the mandible (mm) in 20 boys at 8, 8.5, 9, and 9.5 years of age
- Plot the individual data points for each subject either as a spaghetti plot and/or as a panel plot
- Fit a random slope model to the data; how would you interpret the correlation parameter in the model?
- Recenter your time variable, so that zero means "8 years old"
- Refit your random slope model; try to explain why and how the correlation parameter changes
- Look at the caterpillar plots for the random slope model with and without recentered time variable; why do they look different?
- Create a shrinkage plot plotting the individual intercept as a function of the individual slopes
- Add individual and quadratic time effects to your model; test this model against the random slope model
- Look at the profiles for the random effects for the quadratic model; what would you conclude?

[1] Inspired by https://embraceuncertaintybook.com/longitudinal.html#the-elstongrizzle-data

References

Alday, P., Kliegl, R., & Bates, D. (2025). *Embrace uncertainty – Mixed-effects models with Julia*. https://embraceuncertaintybook.com/

Reisby, N., Gram, L. F., Bech, P., Nagy, A., Petersen, G. O., Ortmann, J., Ibsen, I., Dencker, S. J., Jacobsen, O., Krautwald, O., Sondergaard, I., & Christiansen, J. (1977).Imipramine: Clinical effects and pharmacokinetic variability. *Psychopharmacology*, *54*, 263–272.

Winter, B., & Wieling, M. (2016).How to analyze linguistic change using mixed models, growth curve analysis and generalized additive modeling. *Journal of Language Evolution*, *1*(1), 7–18. https://doi.org/10.1093/jole/lzv003