

WWD_4

zhou ziliang

2019/11/18

Load the data and prepare it for analysis

```
schoolboys <- read_xls('schoolboys.xls')
schoolboys <- rename(schoolboys,
                     Age = "Age centered about 13 years",
                     Height = "Height (cm)",
                     Season = "Season, months from start of year")
schoolboys$ID <- as.factor(schoolboys$ID)
```

question 2

```
mod1 <- lm(Height ~ 1 + Age + Season, schoolboys)
summary(mod1)

##
## Call:
## lm(formula = Height ~ 1 + Age + Season, data = schoolboys)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.6970  -5.1345   0.4539   4.7533  18.9030
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 149.26841    1.11753 133.570  < 2e-16 ***
## Age          6.54468     0.84926   7.706 3.81e-13 ***
## Season       0.01921     0.18263   0.105  0.916
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.098 on 231 degrees of freedom
## Multiple R-squared:  0.2155, Adjusted R-squared:  0.2087
## F-statistic: 31.72 on 2 and 231 DF,  p-value: 6.707e-13
```

The relation between height and age is a positive relationship.

question3

```
schoolboylmer1 <- lmer(Height ~ Age + (1 | ID), data = schoolboys)
summary(schoolboylmer1)

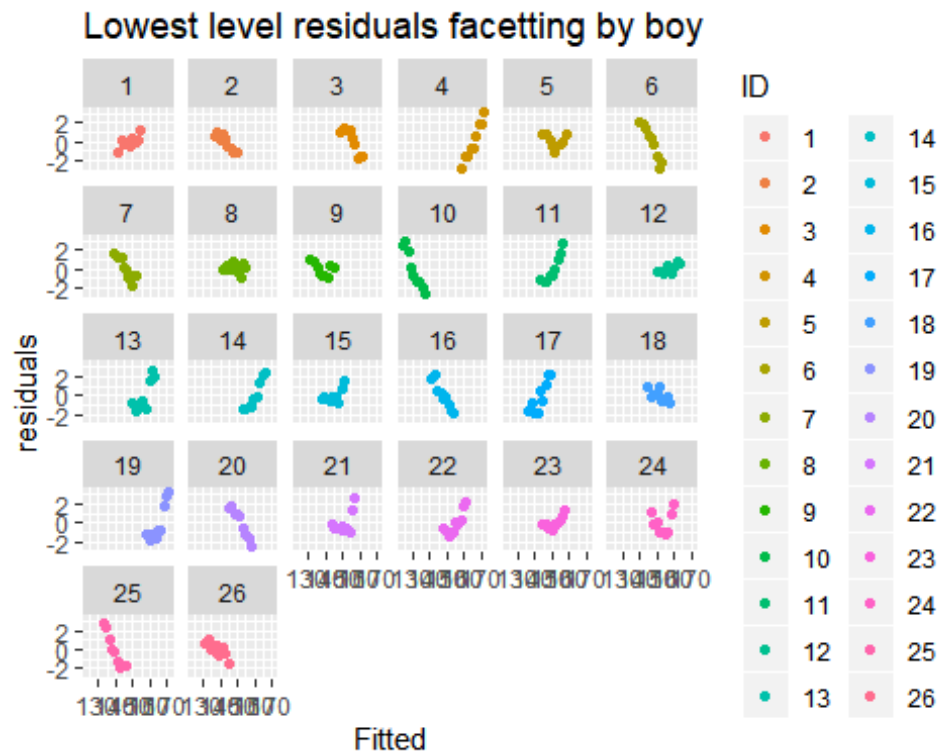
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Height ~ Age + (1 | ID)
## Data: schoolboys
##
## REML criterion at convergence: 940
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.1859 -0.6350 -0.1339  0.6253  2.5357
##
## Random effects:
## Groups Name Variance Std.Dev.
## ID      (Intercept) 65.555  8.097
## Residual              1.718  1.311
## Number of obs: 234, groups: ID, 26
##
## Fixed effects:
##              Estimate Std. Error    df t value Pr(>|t|)
## (Intercept) 149.3719    1.5902 25.0002  93.93  <2e-16 ***
## Age          6.5238     0.1325 207.0000  49.23  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## Age -0.002

performance::icc(schoolboylmer1)

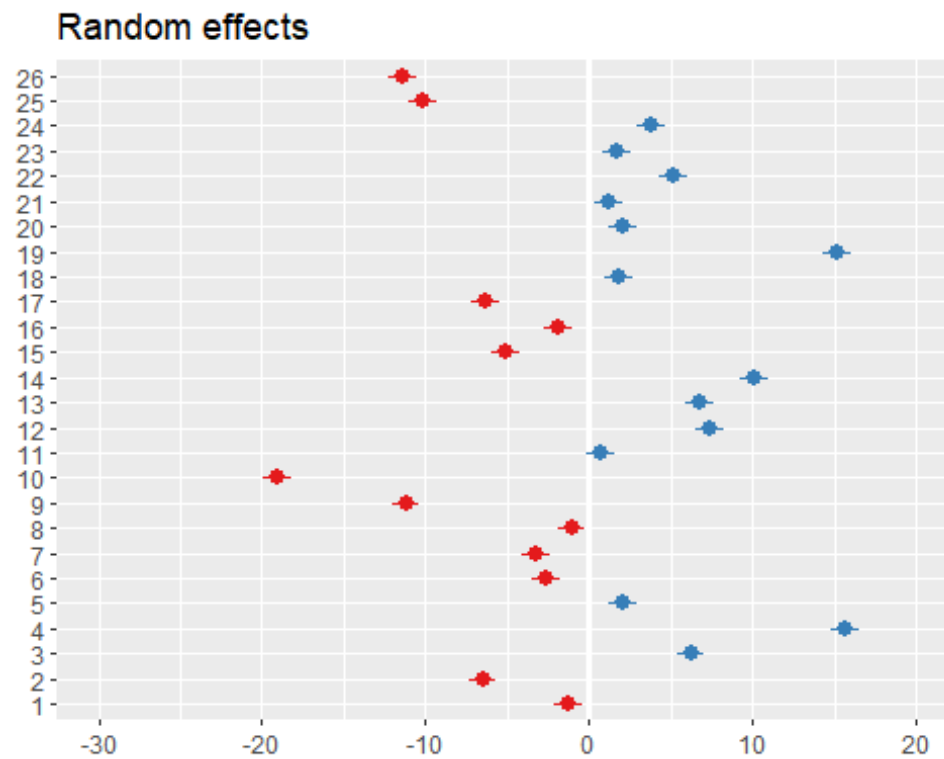
## # Intraclass Correlation Coefficient
##
## Adjusted ICC: 0.974
## Conditional ICC: 0.770

schboymerdiag <- data.frame(
  residuals = resid(schoolboylmer1),
  ID = schoolboys$ID,
  Fitted = fitted(schoolboylmer1)
)

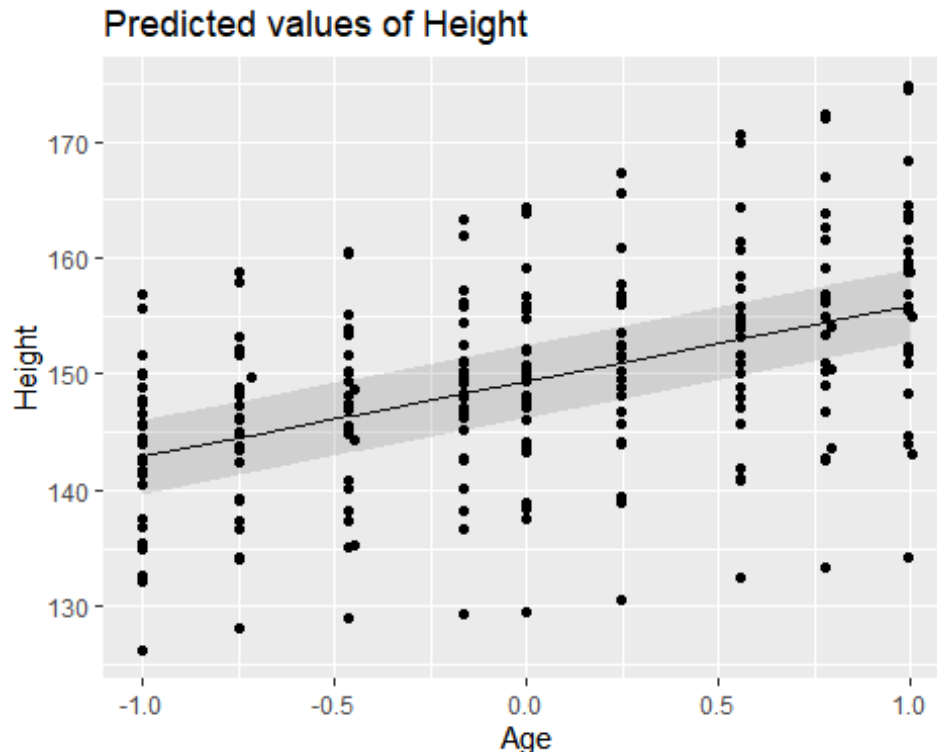
ggplot(schboymerdiag, aes(x = Fitted, y = residuals, col = ID)) +
  geom_point() +
  facet_wrap(~ ID) +
  ggtitle("Lowest level residuals facetting by boy")
```



```
plot_model(schoolboy1mer1, type = "re")
```



```
plot_model(schoolboy1mer1, type = "eff", terms = "Age")+
  geom_point(aes(x=Age,y=Height),data = schoolboys)
```



question 4

```
schoolboy1mer2 <- lmer(Height ~ Age + (1 + Age | ID), schoolboys)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.0128457
## (tol = 0.002, component 1)

summary(schoolboy1mer2)

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Height ~ Age + (1 + Age | ID)
## Data: schoolboys
##
## REML criterion at convergence: 724.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.65235 -0.57475 -0.02844  0.59675  2.60612
##
## Random effects:
## Groups   Name                Variance Std.Dev. Corr
## ID      (Intercept) 65.6377   8.1017
```

```

##           Age           2.8275  1.6815   0.64
## Residual           0.4351  0.6596
## Number of obs: 234, groups: ID, 26
##
## Fixed effects:
##           Estimate Std. Error      df t value Pr(>|t|)
## (Intercept) 149.3719     1.5895  24.8096   93.98  <2e-16 ***
## Age          6.5254      0.3364  24.9606   19.39  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##      (Intr)
## Age 0.628
## convergence code: 0
## Model failed to converge with max|grad| = 0.0128457 (tol = 0.002, componen
t 1)

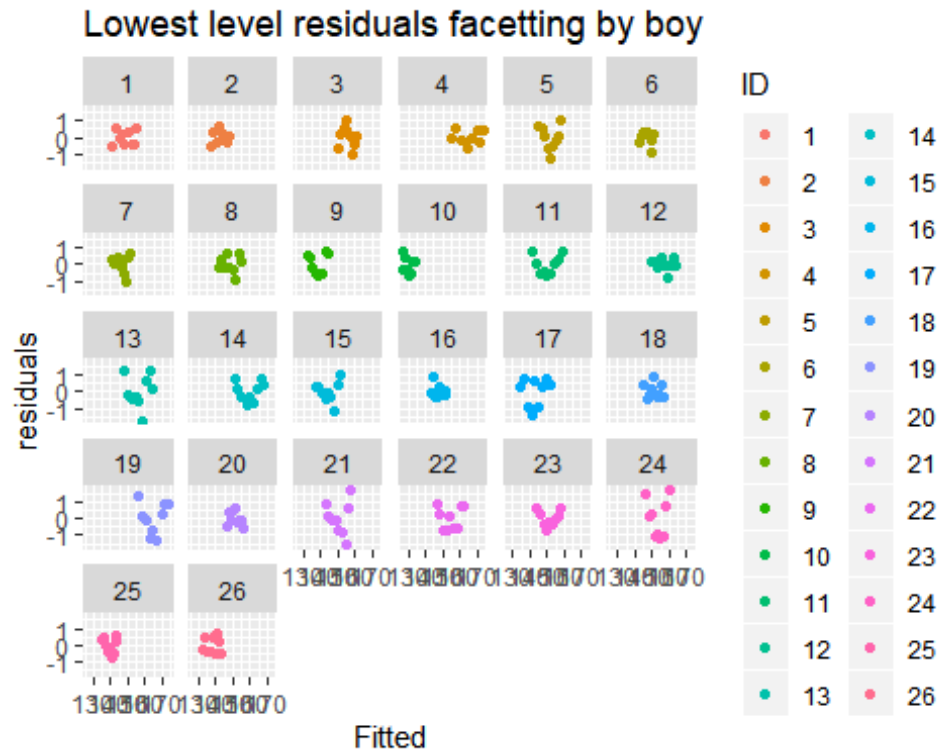
performance::icc(schoolboylmer2)

## # Intraclass Correlation Coefficient
##
## Adjusted ICC: 0.994
## Conditional ICC: 0.786

schboymer2diag <- data.frame(
  residuals = resid(schoolboylmer2),
  ID = schoolboys$ID,
  Fitted = fitted(schoolboylmer2)
)

ggplot(schboymer2diag, aes(x = Fitted, y = residuals, col = ID)) +
  geom_point() +
  facet_wrap(~ ID) +
  ggtitle("Lowest level residuals facetting by boy")

```



```
anova(schoolboylmer1, schoolboylmer2)

## refitting model(s) with ML (instead of REML)

## Data: schoolboys
## Models:
## schoolboylmer1: Height ~ Age + (1 | ID)
## schoolboylmer2: Height ~ Age + (1 + Age | ID)
##
```

	Df	AIC	BIC	logLik	deviance	Chisq	Chi	Df	Pr(>Chisq)
## schoolboylmer1	4	948.57	962.39	-470.28	940.57				
## schoolboylmer2	6	737.95	758.69	-362.98	725.95	214.61		2	< 2.2e-16

```
##
## schoolboylmer1
## schoolboylmer2 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The module two of random intercepts and slope is fitting better for it.