Mathematics Gains

Lucy Tai

Classroom

Modeling

```
classroom <- read.csv("classroom.csv",header=TRUE)
model <- lmer(mathgain ~ (1 | schoolid) + (1 | schoolid:classid),
              data = classroom, REML = FALSE)

summary(model)
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: mathgain ~ (1 | schoolid) + (1 | schoolid:classid)
##    Data: classroom
##
##      AIC      BIC   logLik deviance df.resid
##  11779.3  11799.7  -5885.7  11771.3     1186
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6477 -0.5962 -0.0339  0.5350  5.6354
##
## Random effects:
##  Groups           Name        Variance Std.Dev.
##  schoolid:classid (Intercept)   99.14   9.957
##  schoolid         (Intercept)   75.37   8.682
##  Residual                     1028.30  32.067
## Number of obs: 1190, groups:  schoolid:classid, 312; schoolid, 107
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   57.429      1.436      40
```

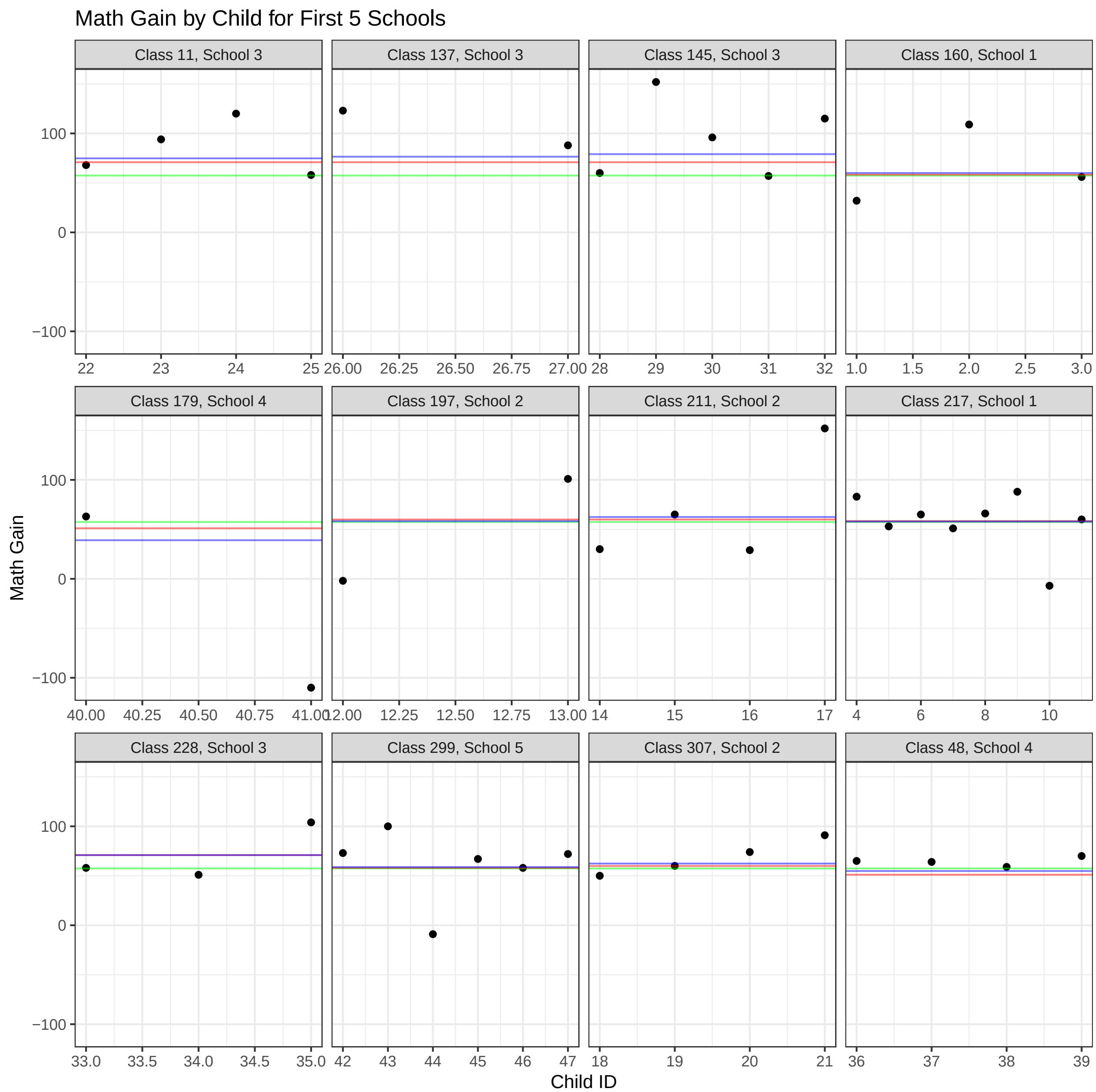
- The model summary is shown above.
- gamma0 hat is the fixed intercept estimate, 57.429. The SE(gamma 0) is the std error of that intercept, 1.443.
- sigma hat is the std dev of the residual, 32.067.
- tau 1 hat is the std dev of the etas distribution for each classroom, 9.957.
- tau 2 hat is the std dev of the zetas distribution for the schools, 8.682.

Plotting

```
# extract line estimations from the model

classplot <- classroom %>% filter(schoolid <= 5)
```

```r
classplot$class.schoolid <- paste(paste("Class", classplot$classid, sep = " "),
                                  classplot$schoolid,
                                  sep = ", School ")
classplot$schoolclass <- paste(classplot$schoolid,
                       classplot$classid,
                       sep = ":")

schoolclass <- unique(classplot$schoolclass)
classes <- unique(classplot$classid)
schools <- rep(NA, length(classes))

for (i in 1:length(classes)){
  temp <- classplot %>% filter(classid == classes[i]) %>% head(1)
  schools[i] <- temp$schoolid
}

gamma_0 <- rep(NA, length(classes))
beta <- gamma_0
alpha <- gamma_0

for (j in 1:length(classes)) {
  # overall average math gain
  gamma_0[j] <- fixef(model)[1]
  # avg math gain at school k
  beta[j] <- gamma_0[j] + ranef(model)$schoolid[schools[j],1]
  # avg math gain at classroom j
  alpha[j] <- beta[j] + ranef(model)$`schoolid:classid`[schoolclass[j],1]
}

params <- data.frame(class.schoolid = unique(classplot$class.schoolid),
                     gamma_0 = gamma_0,
                     beta_k = beta,
                     alpha_j = alpha)
```

```r
ggplot(data = classplot, aes(x = childid, y = mathgain)) +
  geom_point() +
  facet_wrap(~class.schoolid,
             scales = "free_x") +
  geom_abline(data = params, aes(intercept = gamma_0,
                                 slope = 0),
              color = "green", alpha = 0.5) +
  geom_abline(data = params, aes(intercept = beta_k,
                                 slope = 0),
              color = "red", alpha = 0.5) +
  geom_abline(data = params, aes(intercept = alpha_j,
                                 slope = 0),
              color = "blue", alpha = 0.5) +
  labs(x = "Child ID",
       y = "Math Gain",
       title = "Math Gain by Child for First 5 Schools")
```

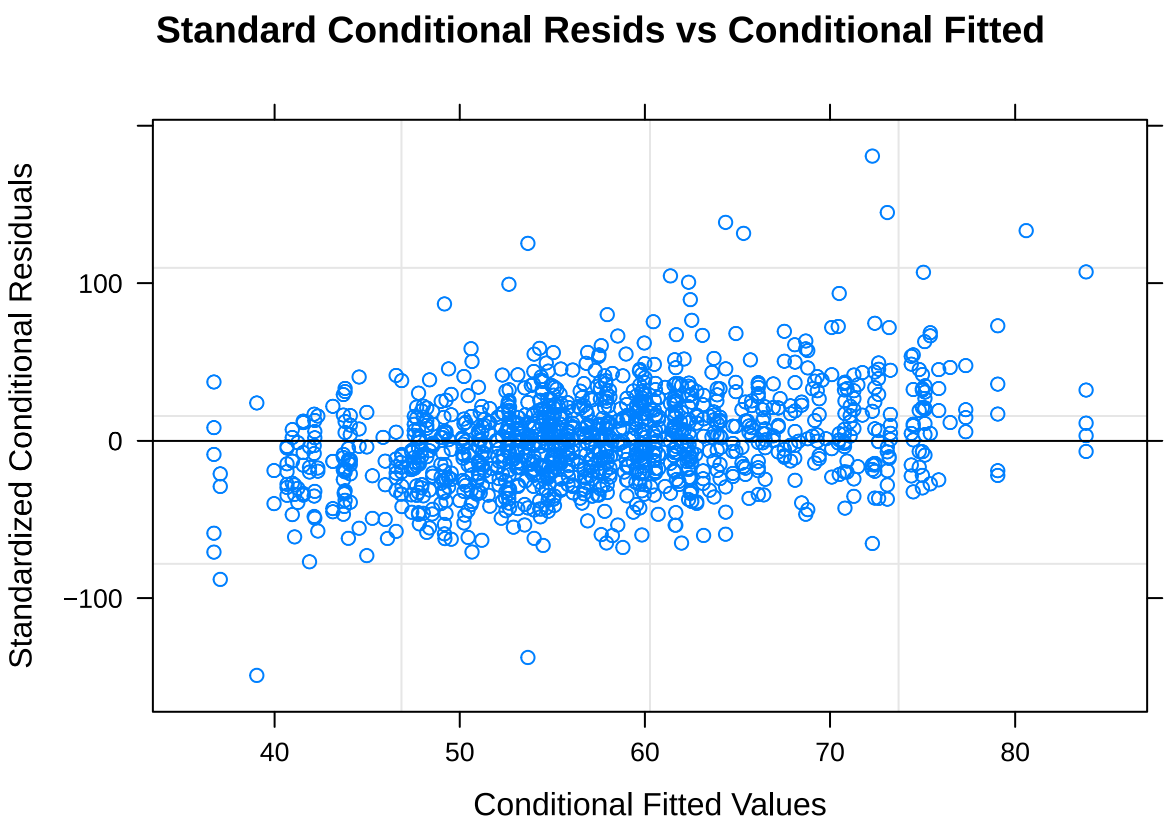


(green: estimated overall avg math gain. red: estimated average math gain for school. blue: estimated average math gain for classroom)

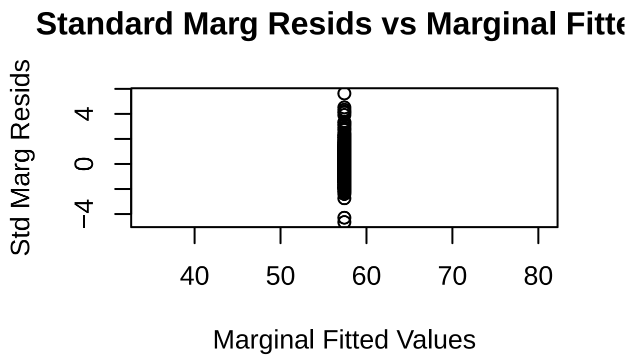
Residuals and standardized params

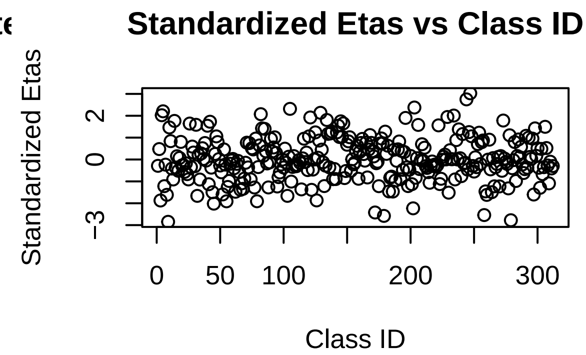
```
r.1s <- hlm_resid(model,level=1,include.ls=FALSE,standardize= TRUE)
r.2s <- hlm_resid(model,level="schoolid:classid",include.ls=FALSE,
                  standardize=TRUE)
r.3s <- hlm_resid(model,level="schoolid",include.ls=FALSE,
                  standardize=TRUE)

par(mfrow=c(2, 2))

plot(model, xlab = "Conditional Fitted Values",
```

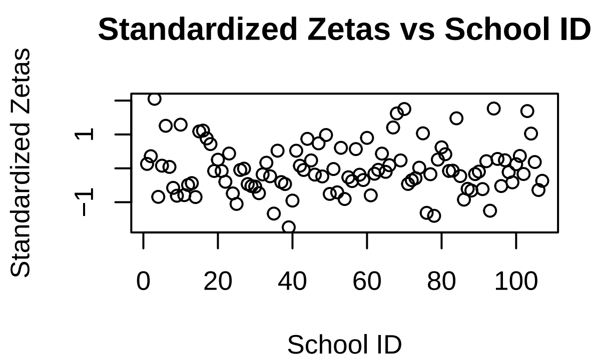
```
     ylab = "Standardized Conditional Residuals",
     main = "Standard Conditional Resids vs Conditional Fitted")
```




```
plot(y = r.1s$.std.resid, x = r.1s$.mar.fitted,
     xlab = "Marginal Fitted Values",
     ylab = "Std Marg Resids",
     main = "Standard Marg Resids vs Marginal Fitted")

# the code for r.2s which should have given the etas
# returns many NA values, so I manually standardized them...
eta_df <- ranef(model)$`schoolid:classid`
eta_df <- tibble::rownames_to_column(eta_df, "schoolclass")
colnames(eta_df) <- c("schoolclass", "eta")
eta_df$class <- sub(".*:", "", eta_df$schoolclass) #extract class only
eta_df$std_eta <- scale(eta_df$eta)

plot(y = eta_df$std_eta, x = eta_df$class,
     xlab = "Class ID",
     ylab = "Standardized Etas",
     main = "Standardized Etas vs Class ID")
plot(y = r.3s$.std.ranef.intercept, x = r.3s$schoolid,
     xlab = "School ID",
     ylab = "Standardized Zetas",
     main = "Standardized Zetas vs School ID")
```







- The standard marginal residuals and marginal fitted plot looks a bit strange. I'm not sure why the model has the same marginal fitted value for every student. Otherwise, the other plots don't look too bad. I don't see a discernible pattern for the etas or zetas. The conditional residuals look like they might be trending slightly upwards as conditional fitted values increase, so that might be something to keep in mind as we look at the model going forward.

Make school

```
school <- rep(NA, 312)
for (j in 1:length(school)){
  # assumes each classroom j is only in one school
  # get first student in classroom j
  temp <- classroom %>%
    filter(classid == j) %>%
    filter(row_number() == 1)
  # set their school as the school for jth class
  school[j] <- temp$schoolid
}
```

Set phasers to STAN (bad joke)

```
classdata <- list(N_students = nrow(classroom),
                  N_classes = length(unique(classroom$classid)),
                  N_schools = length(unique(classroom$schoolid)),
                  classid = classroom$classid,
```

```r
                   school = school,
                   mathgain = classroom$mathgain)

classmodel <- stan(file = "classroom.stan",
                        data = classdata)
classresult <- stan(fit = classmodel, data = classdata,
                     iter = 2000)
```

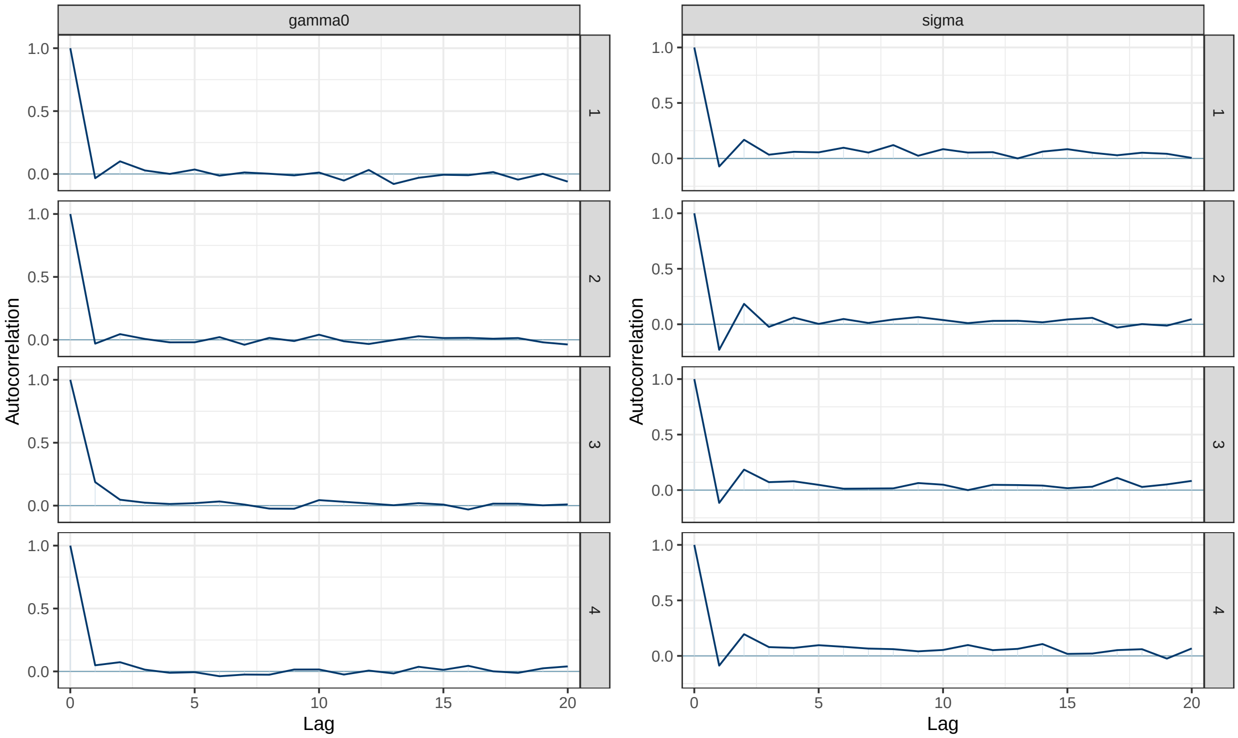
```r
print(classresult,pars=c("gamma0","sigma","tau1","tau2"))
```

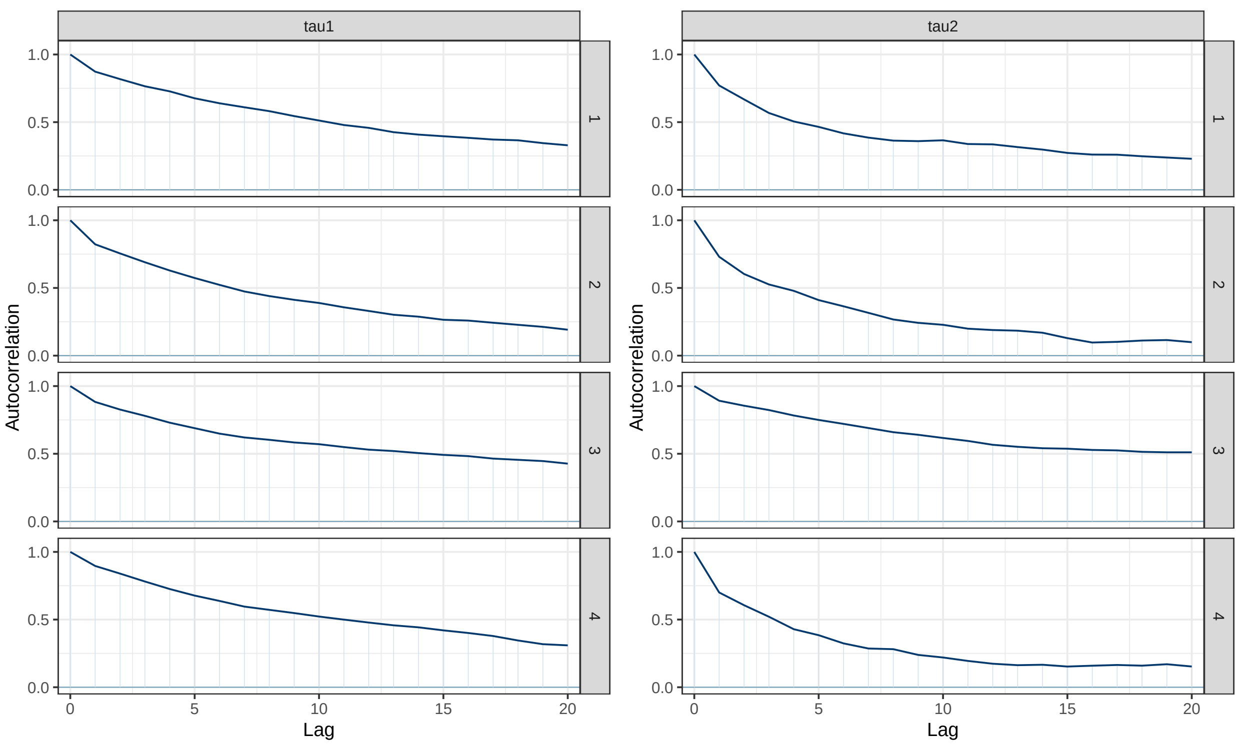
```
## Inference for Stan model: classroom.
## 4 chains, each with iter=2000; warmup=1000; thin=1;
## post-warmup draws per chain=1000, total post-warmup draws=4000.
##
##         mean se_mean   sd  2.5%   25%   50%   75% 97.5% n_eff Rhat
## gamma0 57.38    0.03 1.46 54.52 56.43 57.40 58.36 60.18  3133 1.00
## sigma  31.55    0.02 0.73 30.15 31.07 31.53 32.04 33.00  1254 1.00
## tau1   10.72    0.22 2.12  6.30  9.33 10.76 12.16 14.74    95 1.01
## tau2    8.53    0.27 2.24  2.93  7.31  8.69 10.04 12.42    68 1.04
##
## Samples were drawn using NUTS(diag_e) at Tue Jul 19 17:05:00 2022.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

- The results are shown above and are indeed quite similar to what we got in using lmer.
- The R hats are very close to 1 and all smaller than 1.05, which is good.
- Our n_eff for tau1 and tau2 are not looking so good, they're both less than 10% of the 4000 draws, so our correlations might be dubious. Should do more draws when given more time.

Autocorrelation plots for MCMC chains

```r
g1 <- mcmc_acf(classresult, pars = "gamma0")
g2 <- mcmc_acf(classresult, pars = "sigma")
g3 <- mcmc_acf(classresult, pars = "tau1")
g4 <- mcmc_acf(classresult, pars = "tau2")

grid.arrange(g1,g2,g3,g4,ncol=2)
```

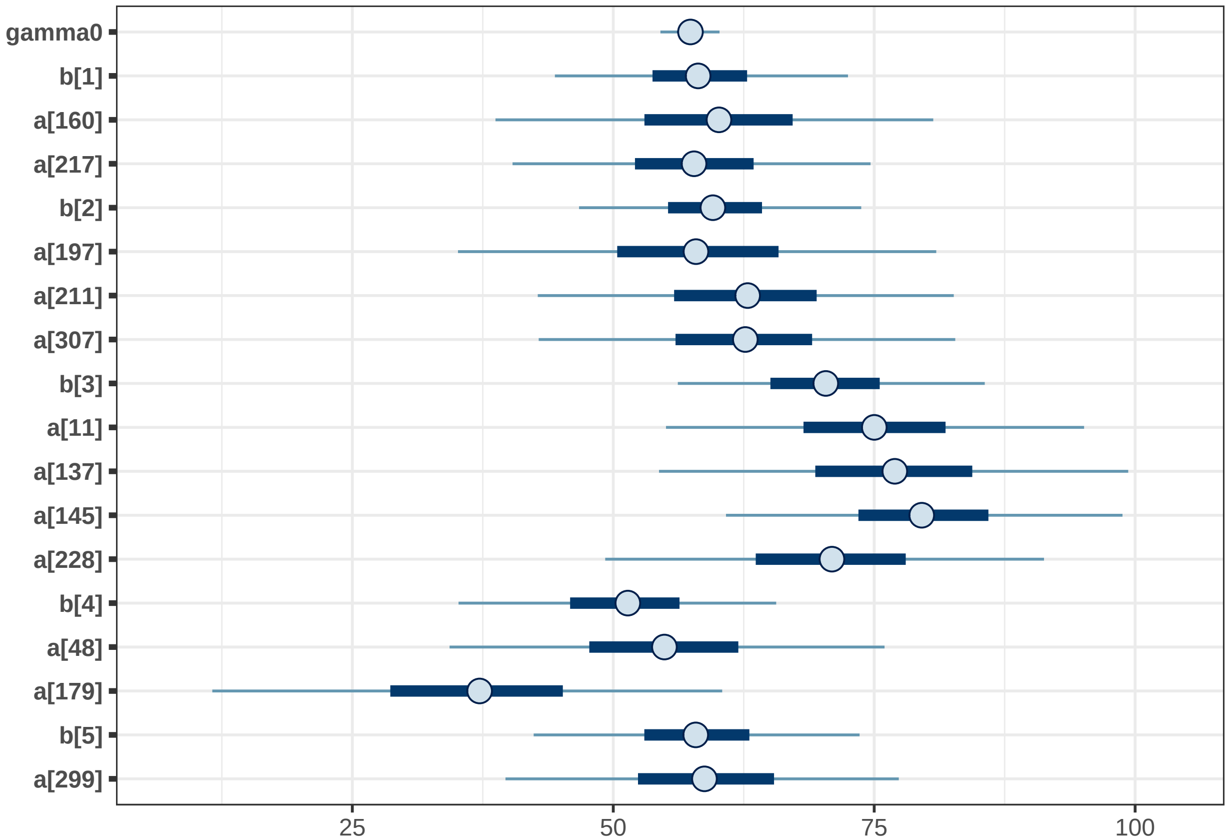




- The correlations, though decreasing, are not great, especially for tau 1 (and to a lesser degree tau 2). They should look more like those of sigma and gamma0, but instead they don't really converge for any of the 4 markov chains, though they almost do for the 2nd one.
- The acf plots suggest that both taus may be experiencing this problem– we can see that from the aforementioned very slow convergence in both their graphs.

Posterior 95% CIs for first 5 schools alphas and betas

```
# classroom %>% filter(schoolid == 5)

params <- c("gamma0",
            "b[1]", # school 1
            "a[160]", # classrooms in school 1
            "a[217]",

            "b[2]", # school 2
            "a[197]",
            "a[211]",
            "a[307]",

            "b[3]", # s3
            "a[11]",
            "a[137]",
            "a[145]",
            "a[228]",

            "b[4]", # s4
            "a[48]",
            "a[179]",

            "b[5]", # s5
            "a[299]"
            )

mcmc_intervals(classresult, prob_outer=0.95,
               pars=grep("lp__", params,invert=T,value=T))
```



- Above is a plot of posterior 95% CIs for gamma 0 (the first row), followed by each school's beta[school id] and that school's corresponding classes' alpha[class id].
- Shrinkage is evident for both the school and classroom effect estimates, where beta CIs are clustered near gamma0, and generally the alpha CIs are near their corresponding school betas.