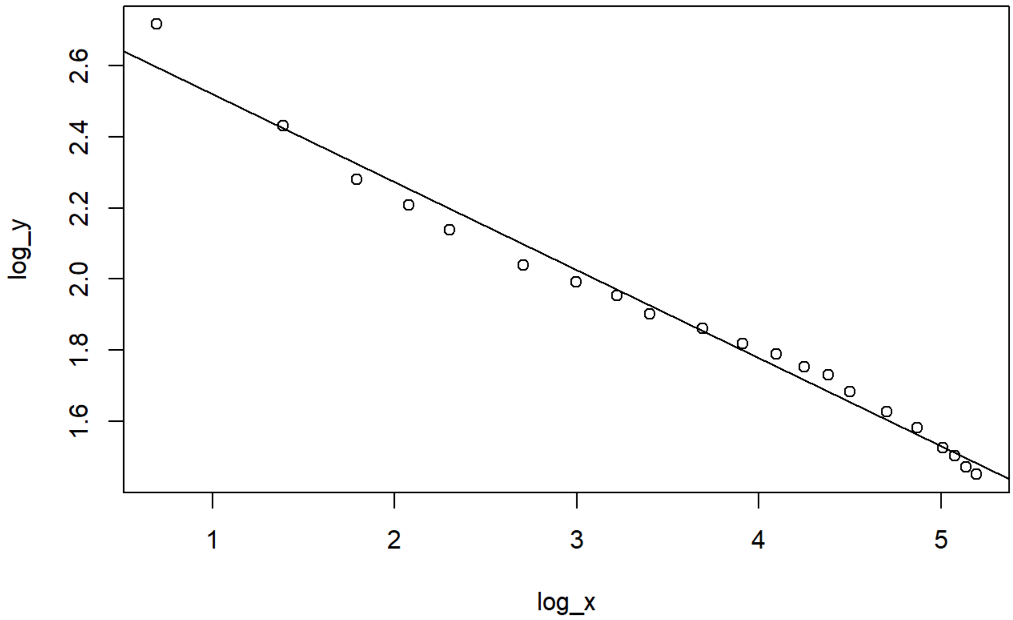
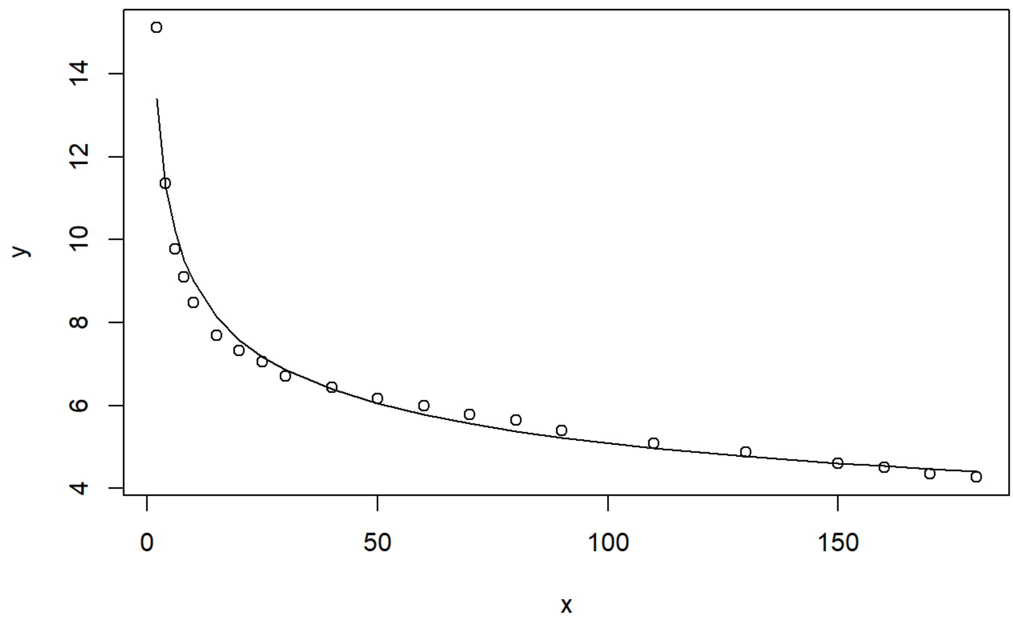
7.9 (a) data points and regression line in log-log coordinate



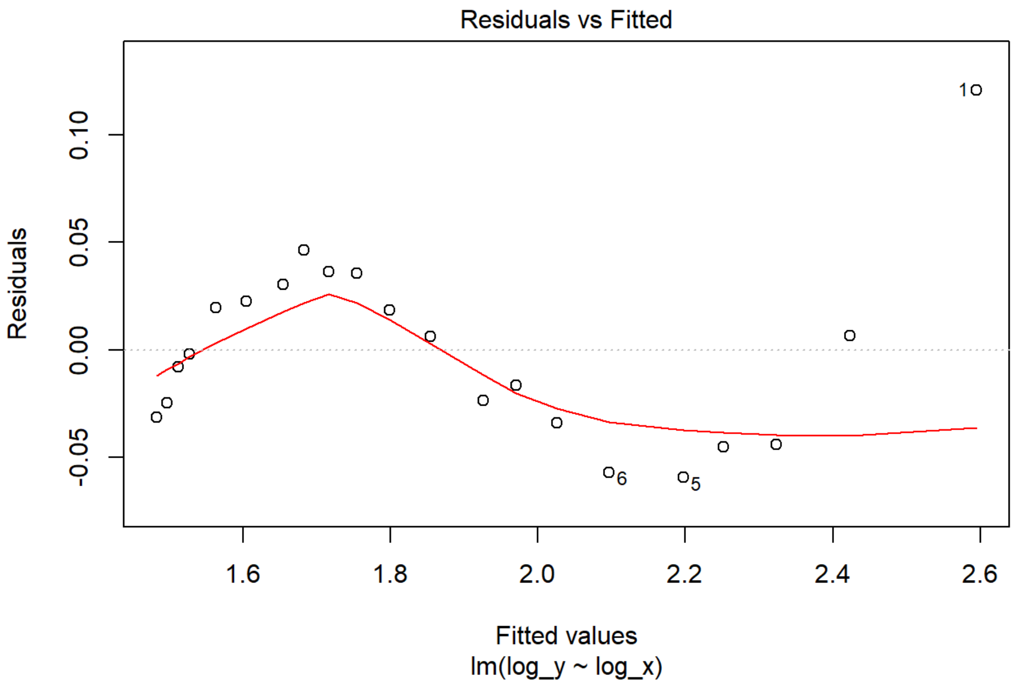
R squared value

```
## [1] 0.9839251
```

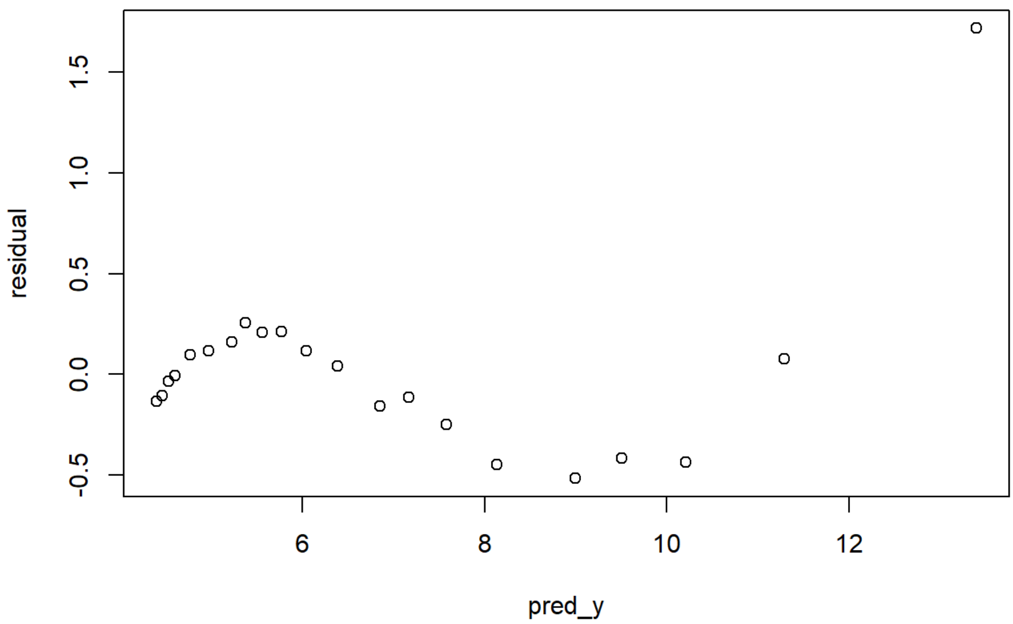
b. data points and regression line in original coordinate



c. plot residual against fitted values in log-log coordinates

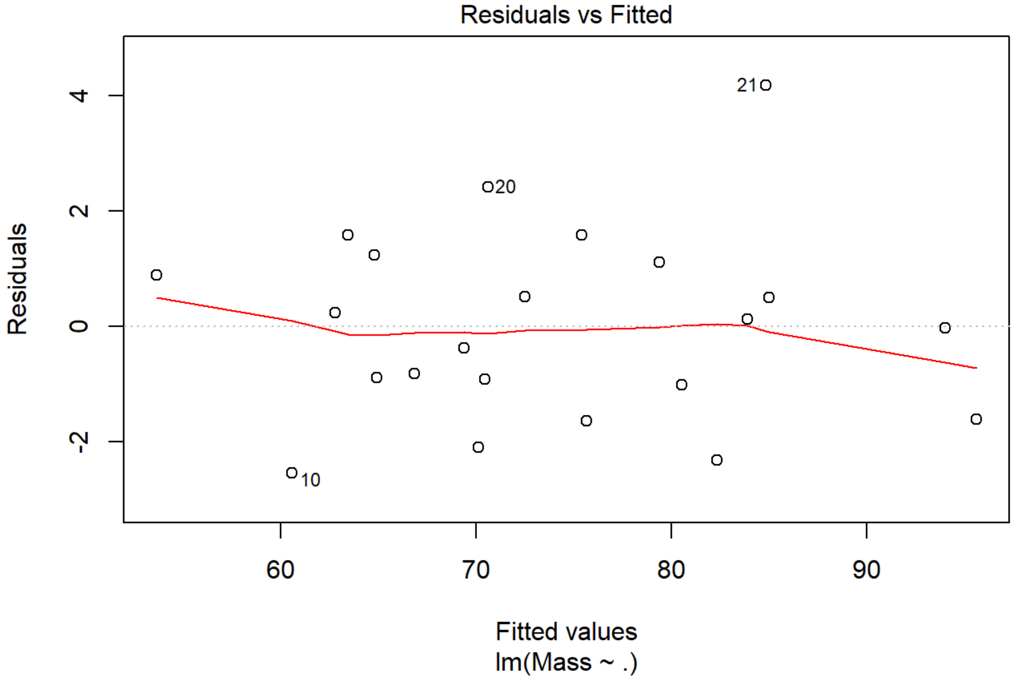


c. plot residual against fitted values in original coordinates



d. It's a good regression as the residual is random and the line (curves) fits the data(R squared should be high even though we are not supposed to measure it).

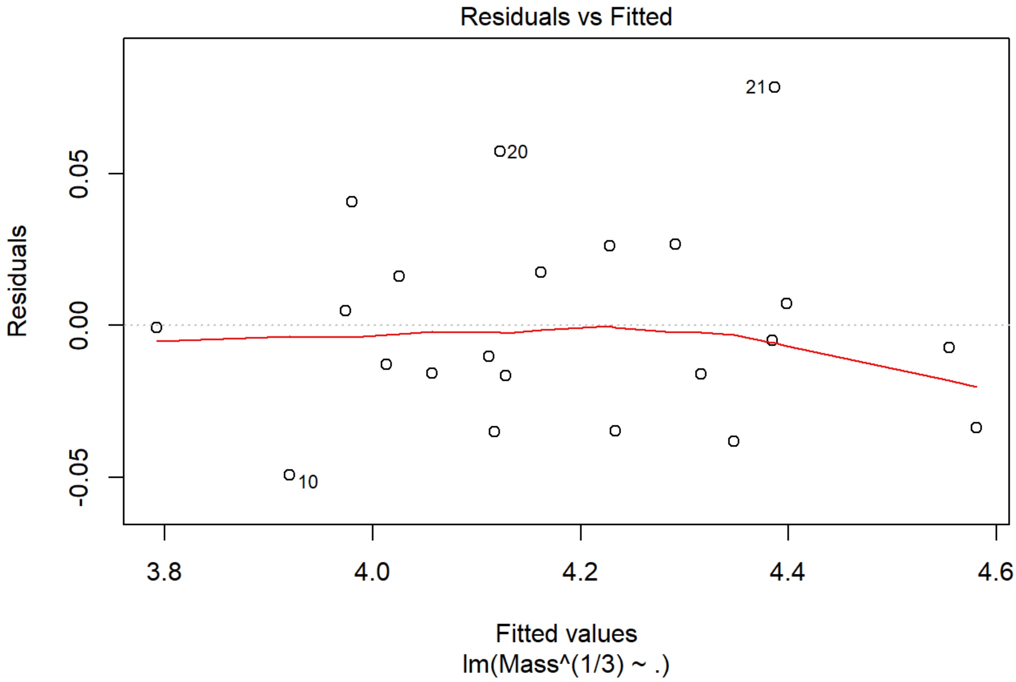
7.10 (a) plot the residual against the fitted values



R squared

```
## [1] 0.9772107
```

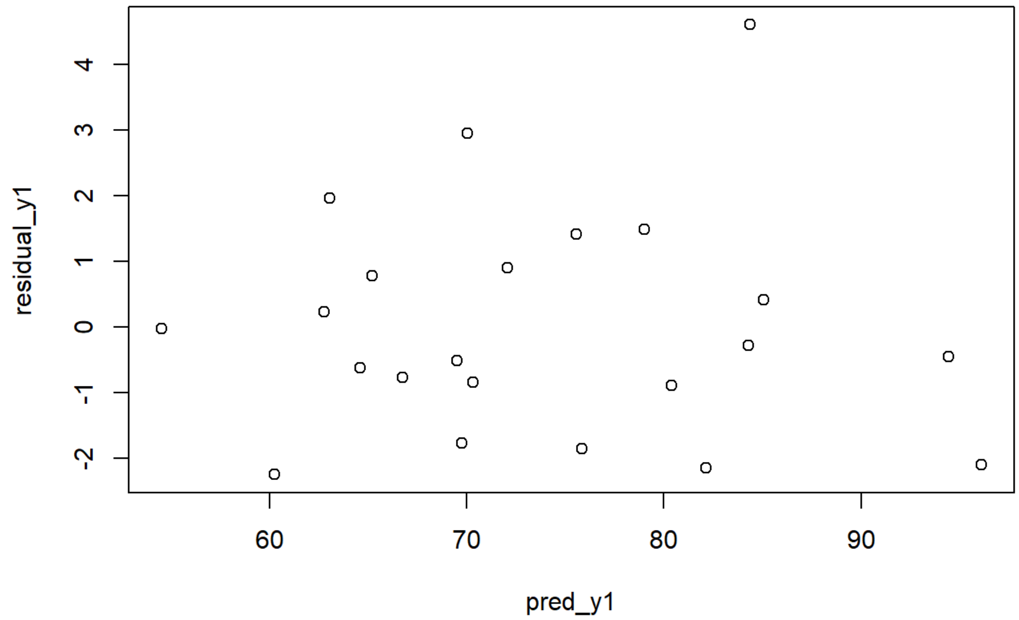
b. build model on cube root of mass



r squared

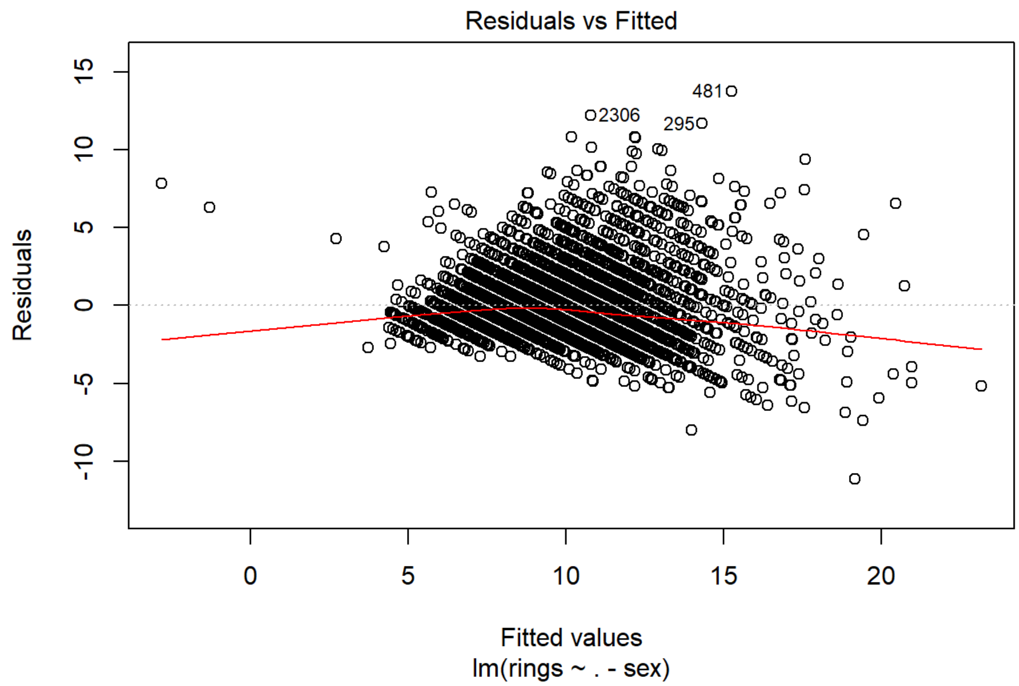
```
## [1] 0.9758476
```

residual vs fitted value on the original coordinated



c. By eyeballing the residual plots c and a, they tend to be very similar; by checking R square values, original model is slightly higher than cubic model. To answer the question " use your plots to explain which regression is better“, I would say both are good because they have random residual plot and the sparsity is the same. But essentially if comparing the R squared, original model is better.

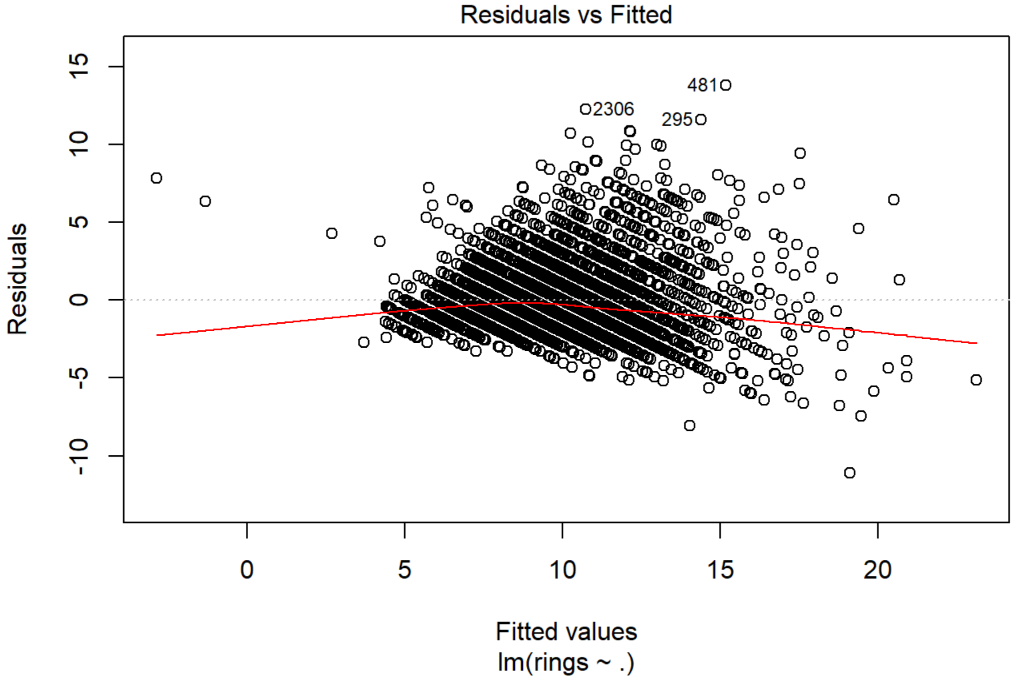
7.11 (a) residual against fitted values(gender excluded)



r squared

```
## [1] 0.5276299
```

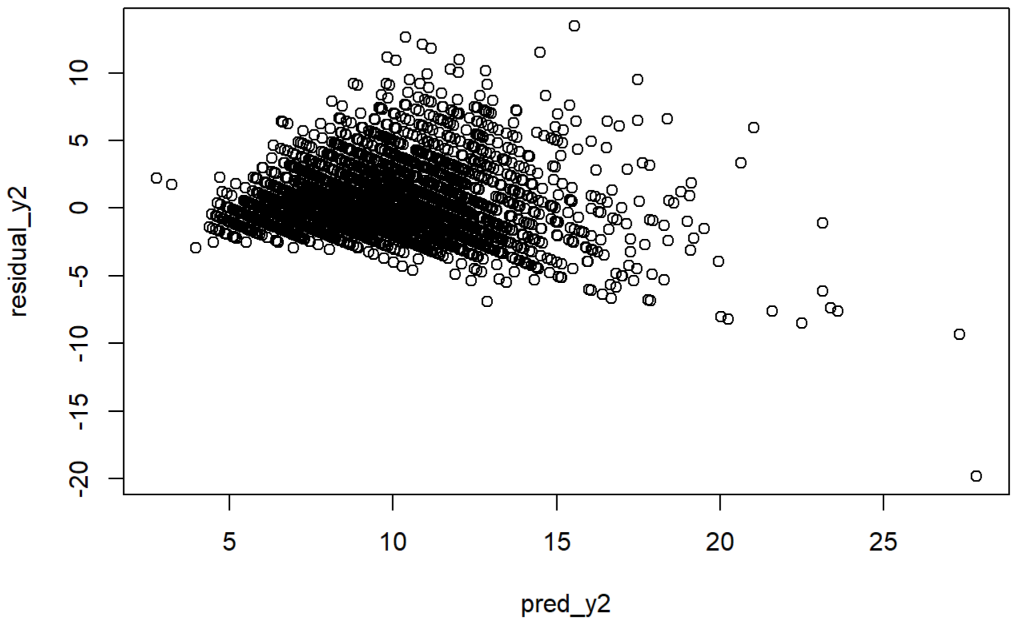
b. residual against fitted values(gender included)



r squared

```
## [1] 0.5278909
```

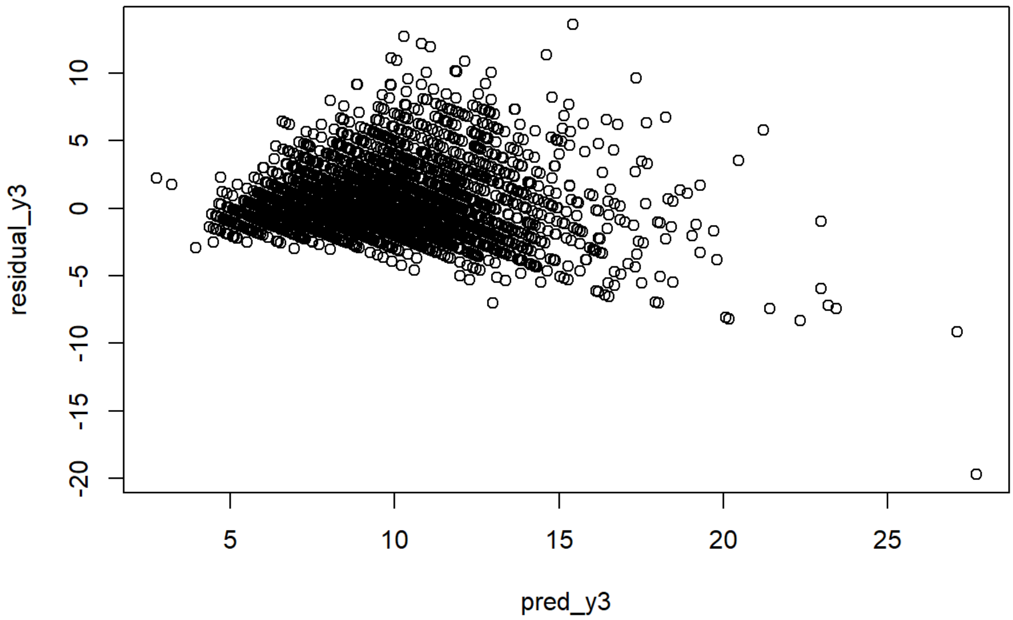
c. plot residual against the fitted values for predicting log age model(sex excluded) in the original coordinate.



r squared

```
## [1] 0.5854699
```

d. plot residual against the fitted values for predicting log age model(sex included) in the original space

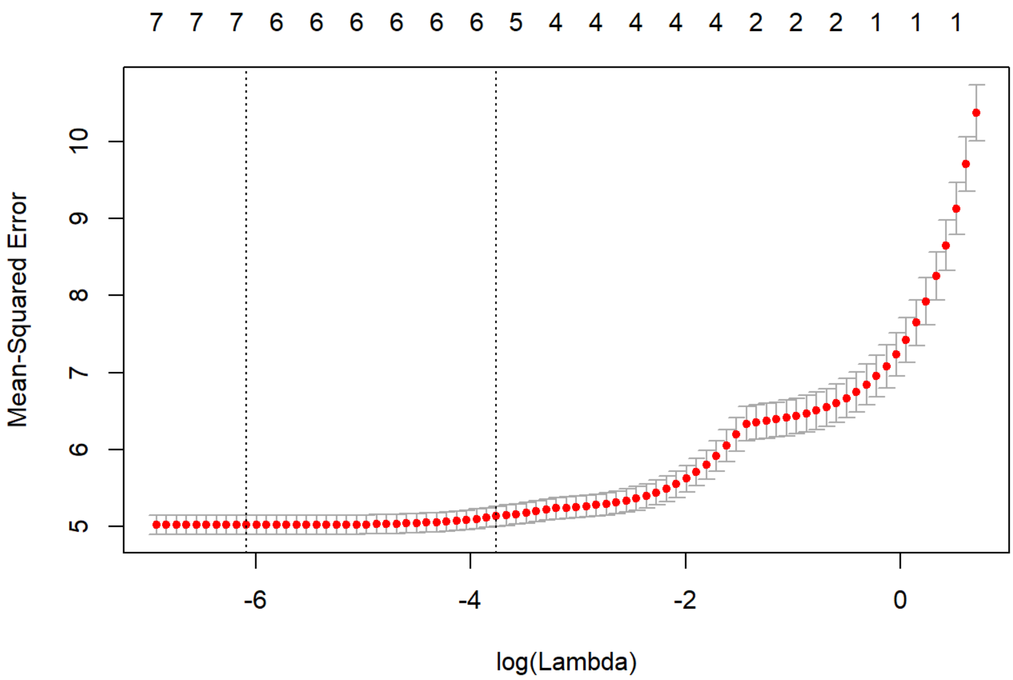


r squared

```
## [1] 0.5859342
```

e. By eyeballing the residual plot, a and b are similar, c and d are similar, c and d tend to be more sparsed than a and b, the residuals are more random, so c and d are better. I will choose model c (log age without gender) because comparing to a and b, c model is better fitted; Comparing to d, they are equally well fitted (residual plot is very similar), but using c will make this process (determin the age) easier, because you don't have to physically check the gender and get the data (gender is not considered in this model). I also check r squared, c,d are equal while a,b are equal but lower. This also proves my conclusion.

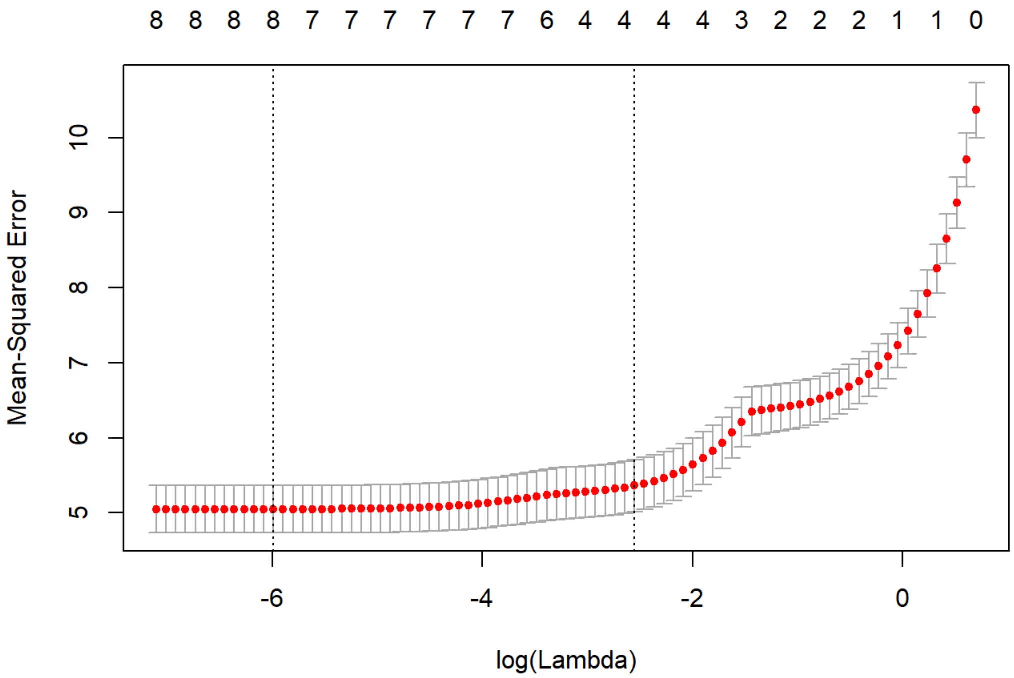
f. I didn't specify alpha runing regularizer. cross-validated error for regularized model a



deviance percentage. when using elastic net,this value equals to r squared

```
## [1] 0.5274003
```

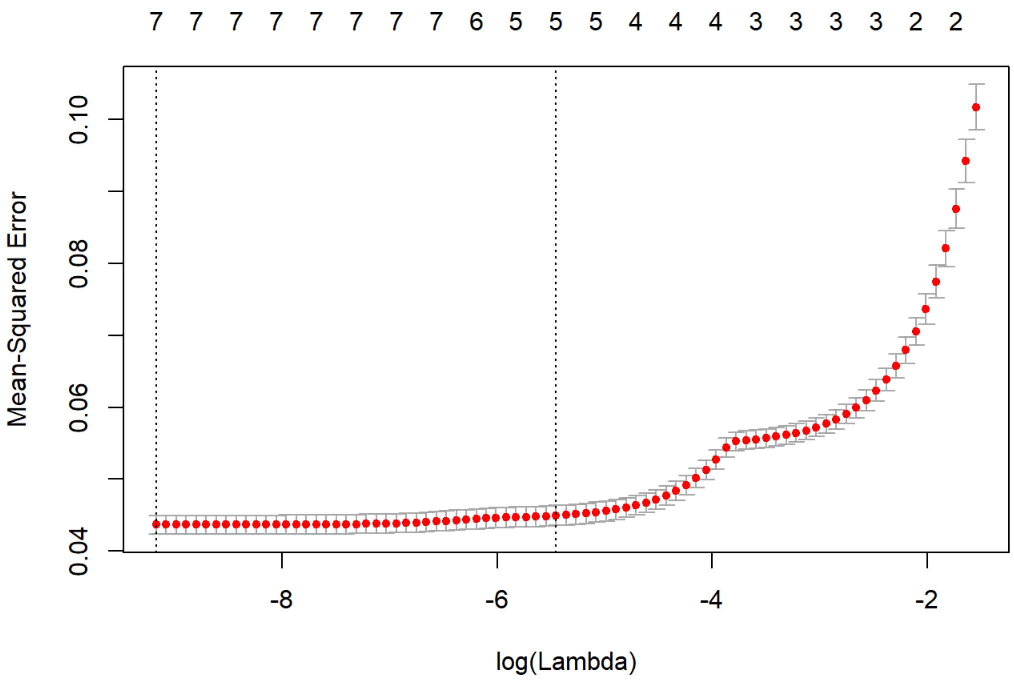
cross-validated error for regularized model b



deviance percentage. when using elastic net,this value equals to r squared

```
## [1] 0.5276227
```

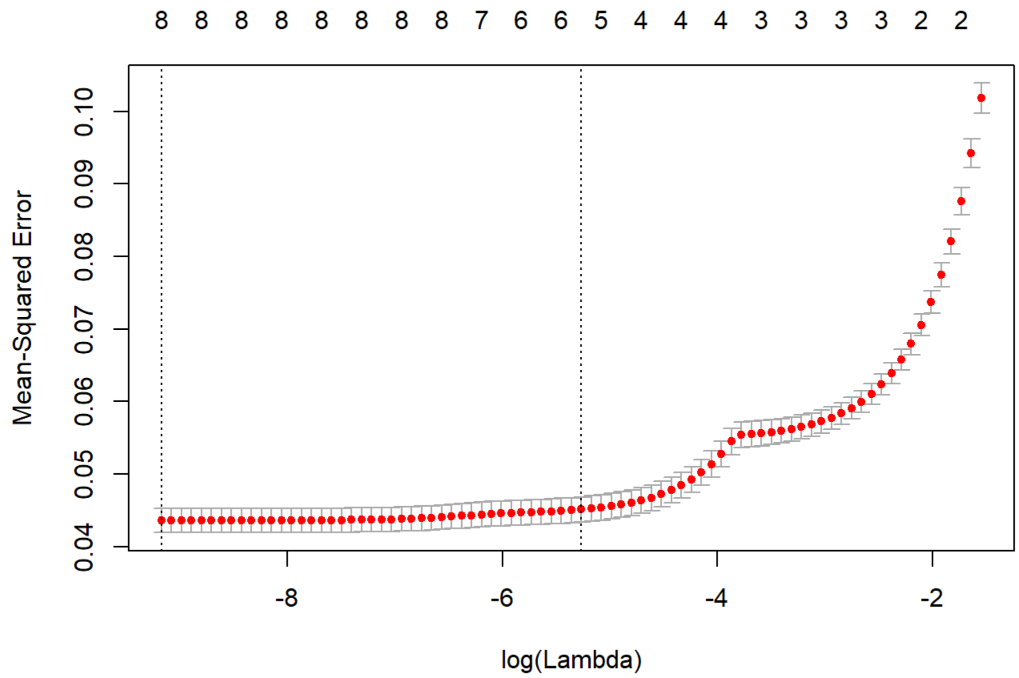
cross-validated error for regularized model c



deviance percentage. when using elastic net,this value equals to r squared

```
## [1] 0.5854198
```

cross-validated error for regularized model d



deviance percentage. when using elastic net,this value equals to r squared

```
## [1] 0.5858841
```

For each plot, imagine on the far left where all variables are used and the lambda closes to zero, can be treated as unregularized. So the error of the unregularized model is closes to the left red dot of the plot.

the left vertical line is where the lambda with the minimum deviance, I choose this value and check the error,it is almost the same as the unregularized model(the very left red dot).

This means if I choose this lambda to do regularization, the training error is the same. So it doesn't help improving the model. But it may generate better results on test data.

By checking the r squared, for each model, it tends to be the same after applying regularization. This proves my conclusion that a regularizer didn't help these regressions as it didn't help improve r squared