Ridge Regression

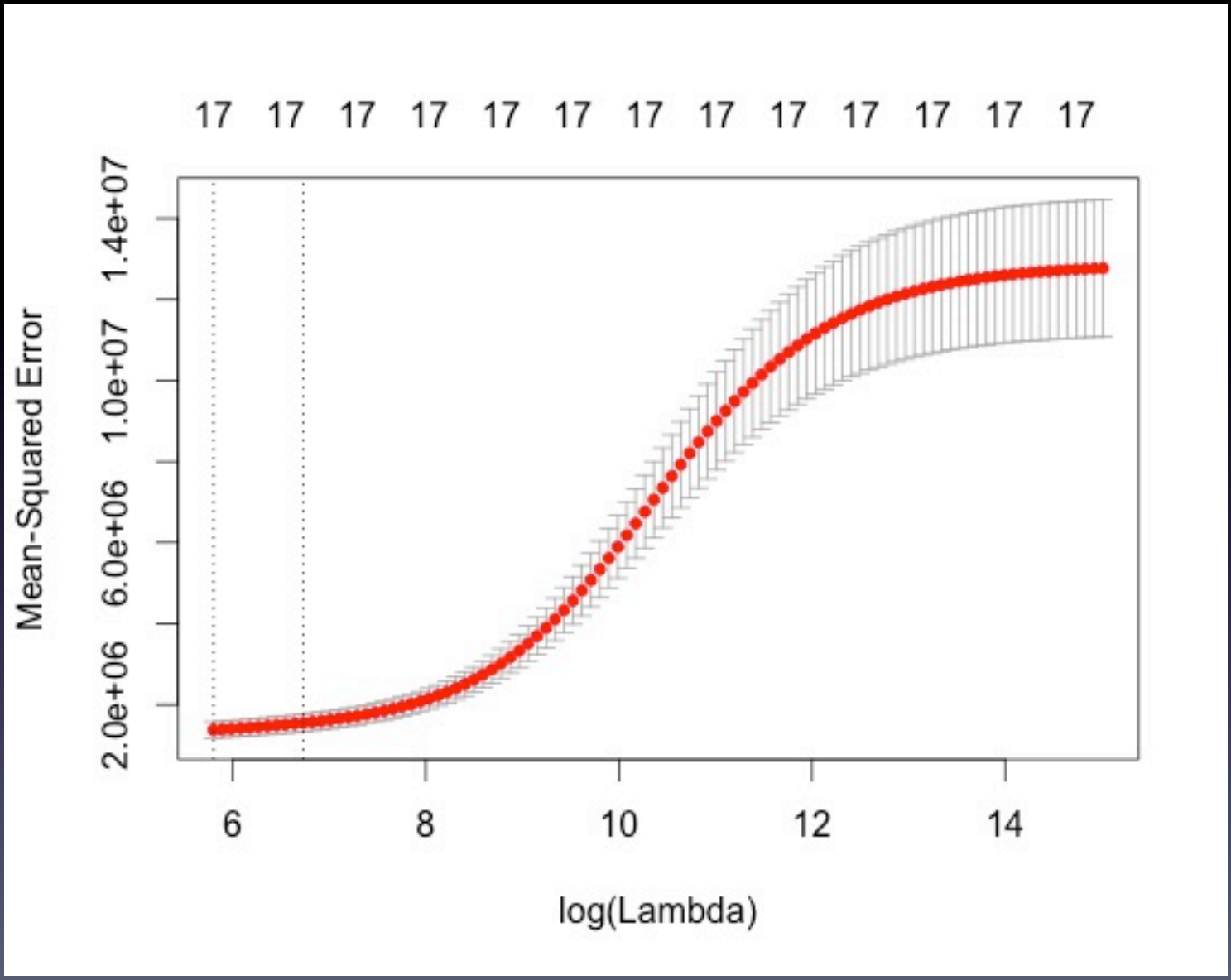
Arshbir Bhullar, Emily Duncan, Kevin Hoang, Brett Mitchell

What it does

- Shrinkage method to prevent over fitting
- Similar to least squares
- Forces estimates to be small but not zero
- Ridge does not give us a good interpretation, nothing on what variables are important
- Better fit than other methods
- Performs well in terms of prediction accuracy
- λ is the parameter we use to find the smallest coefficients
- Finding λ is hard, that's why we use R!

R-Code Example

```
set.seed(100)
cv.out=cv.glmnet(x[train,],y[train],alpha=0)
plot(cv.out)
bestlam=cv.out$lambda.min
bestlam
[1] 331.2333
#The best lambda from cross validation is 331
ridge.pred=predict(ridge.mod,s=bestlam,newx=x[test,])
mean((ridge.pred-y.test)^2)
[1] 1278350
#Which gives us this MSE
out=glmnet(x,y,alpha=0)
> predict(out,type="coefficients",s=bestlam)[1:18,]
(Intercept)    PrivateYes        Accept        Enroll     Top10perc
-1.468326e+03 -5.278781e+02  1.004588e+00  4.313442e-01  2.580619e+01
Top25perc   F.Undergrad   P.Undergrad      Outstate    Room.Board
5.501092e-01  7.258520e-02  2.420595e-02 -2.407454e-02  1.987732e-01
Books      Personal           PhD      Terminal     S.F.Ratio
1.285477e-01 -8.146130e-03 -4.028284e+00 -4.811071e+00  1.302180e+01
perc.alumni        Expend     Grad.Rate
-8.544783e+00  7.589013e-02  1.126699e+01
baseMSE=mean((mean(y[train])-y.test)^2)
baseMSE
[1] 13503050
1278350/13503050
[1] 0.0946712
```


17 17 17 17 17 17 17 17 17 17 17 17 17
Mean-Squared Error
2.0e+06
6.0e+06
1.0e+07
1.4e+07
6
8
10
12
14
log(Lambda)

Results

- While Ridge Regression fits the model well, it literally tells us nothing about what is significant
- The ratio between MSE when $\lambda = 331$ and the base MSE is 0.09 which is good but not interpretable (does not perform variable selection)

ioq