

S2208 MATH8050 Data Analysis - Section 001: Homework 10 Due on 11/30/22

Adithya Ravi, C09059838

2022-11-29

```
set.seed(123)
n <- 1000
p <- 2000
pred <- matrix(rnorm(n*p), nrow = n, ncol = p)
dv <- (rowSums(pred[,1:5]) + .8*rowSums(pred[,6:10]) +
      .6 * rowSums(pred[,11:15]) + .4*rowSums(pred[,16:20]) +
      .2 * rowSums(pred[,21:25]) + rnorm(n))
pred <- scale(pred)
train_rows <- sample(1:n, .8*n, replace = F)
pred.train <- pred[train_rows,]
dv.train <- dv[train_rows]
pred.test <- pred[-train_rows,]
dv.test <- dv[-train_rows]
```

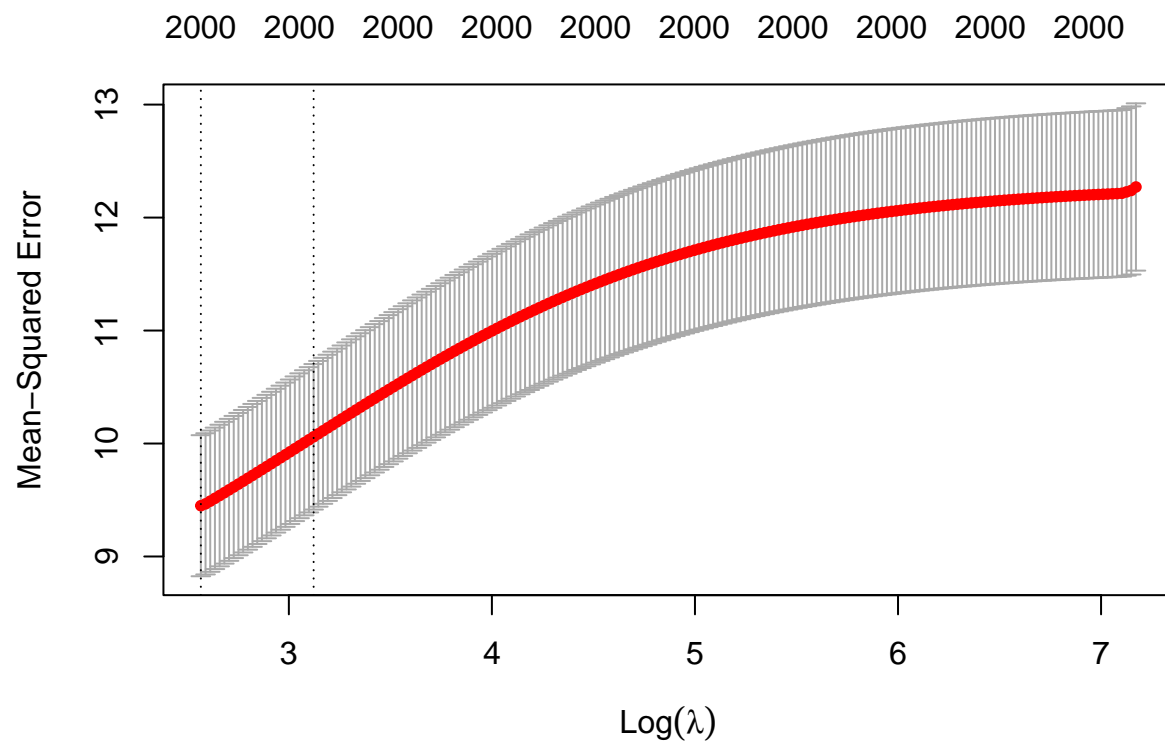
Solutions

question1

1a

```
ridge <- cv.glmnet(x=pred.train, y=dv.train, type.measure="mse",
alpha=0, family="gaussian", nlambd=200)
```

```
plot(ridge)
```



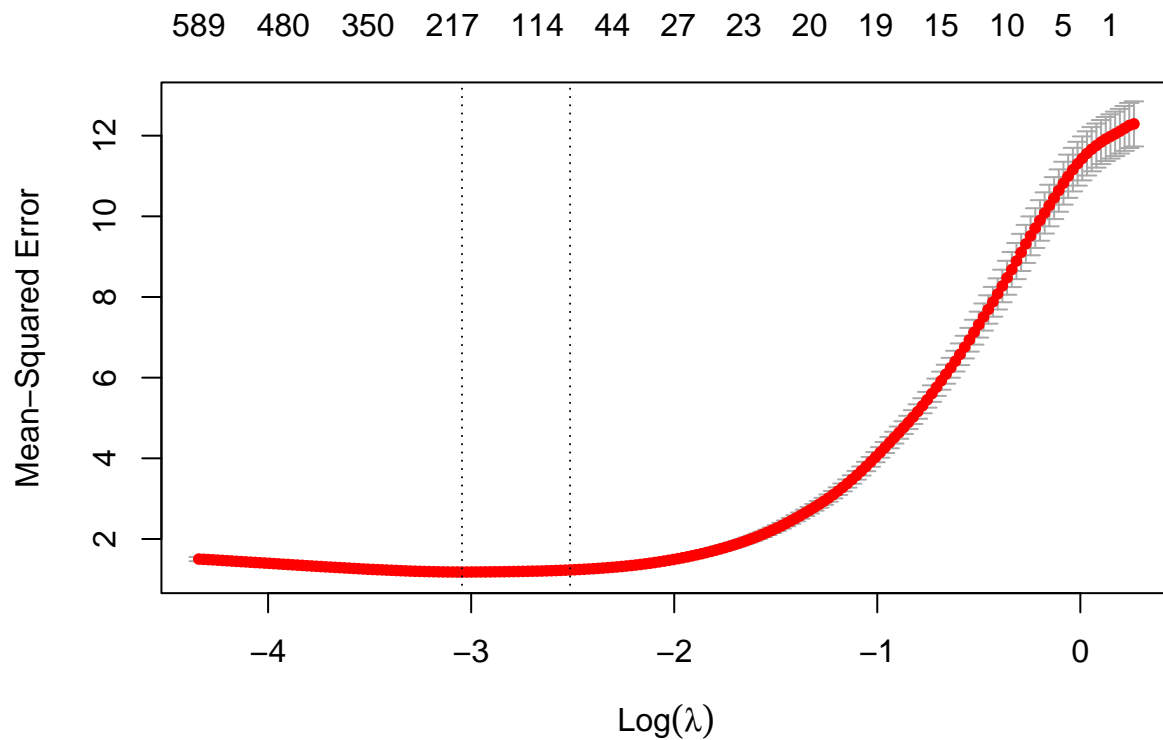
```
ridge.predicted <- predict(ridge, ridge$lambda.1se, new=pred.test)
mean((dv.test - ridge.predicted)^2)
```

```
## [1] 12.33611
```

1b

```
LASSO <- cv.glmnet(x=pred.train, y=dv.train, type.measure="mse",
alpha=1, family="gaussian", nlambdas=200)
```

```
plot(LASSO)
```



```
lasso.predicted <- predict(LASSO, LASSO$lambda.1se, new=pred.test)
mean((dv.test - lasso.predicted)^2)
```

```
## [1] 1.39962
```

1c

```
models_list <- list()
for (i in 0:20) {
  names <- paste0("alpha", i/20)
  models_list[[names]] <- cv.glmnet(pred.train,
                                   dv.train, type.measure="mse",
                                   alpha=i/20, family="gaussian")
}
```

```
results <- data.frame()
for (i in 0:20) {
  name <- paste0("alpha", i/20)
  predicted <- predict(models_list[[name]],
                      s=models_list[[name]]$lambda.1se, newx=pred.test)
  mse <- mean((dv.test - predicted)^2)
  temp <- data.frame(alpha=i/20, mse=mse, name=name)
  results <- rbind(results, temp)
}
```

```
results
```

```
##      alpha      mse      name
## 1  0.00 12.031937  alpha0
## 2  0.05  2.977836 alpha0.05
## 3  0.10  2.030369 alpha0.1
## 4  0.15  1.764349 alpha0.15
## 5  0.20  1.669560 alpha0.2
## 6  0.25  1.581807 alpha0.25
## 7  0.30  1.551845 alpha0.3
## 8  0.35  1.492186 alpha0.35
## 9  0.40  1.483776 alpha0.4
## 10 0.45  1.472959 alpha0.45
## 11 0.50  1.481883 alpha0.5
## 12 0.55  1.425983 alpha0.55
## 13 0.60  1.441359 alpha0.6
## 14 0.65  1.464152 alpha0.65
## 15 0.70  1.439209 alpha0.7
## 16 0.75  1.432360 alpha0.75
## 17 0.80  1.401791 alpha0.8
## 18 0.85  1.436598 alpha0.85
## 19 0.90  1.431786 alpha0.9
## 20 0.95  1.399239 alpha0.95
## 21 1.00  1.440575  alpha1
```

```
plot(results$alpha,results$mse)
```

