STAT 4610

Michael Ghattas

16/11/2022

Chapter 6

```
library(ISLR)
library(glmnet)
```

```
## Warning: package 'glmnet' was built under R version 4.1.2
```

```
## Loading required package: Matrix
```

```
## Warning: package 'Matrix' was built under R version 4.1.2
```

```
## Loaded glmnet 4.1-4
```

```
library(pls)
```

```
## Warning: package 'pls' was built under R version 4.1.2
```

```
##
## Attaching package: 'pls'
```

```
## The following object is masked from 'package:stats':
##
##     loadings
```

Problem - 9

```
attach(College)
```

Part (a)

```
set.seed(1)
train = sample(c(TRUE, FALSE), nrow(College), rep = TRUE)
test = (!train)

College.train = College[train, ]
College.test = College[test, ]
```

Part (b)

```
lm.fit = lm(Apps ~ ., data = College.train)
lm.pred = predict(lm.fit, College.test, type = "response")
mean((lm.pred - College.test$Apps)^2)
```

```
## [1] 984743.1
```

-> Linear model fit test-error = 984743.1

Part (c)

```
set.seed(1)

train.mat = model.matrix(Apps~., data = College.train)
test.mat = model.matrix(Apps~., data = College.test)

cv.out = cv.glmnet(train.mat, College.train$Apps, alpha = 0)
bestlam = cv.out$lambda.min
bestlam
```

```
## [1] 394.2365
```

```
ridge.mod = glmnet(train.mat, College.train$Apps, alpha = 0)
ridge.pred = predict(ridge.mod, s = bestlam, newx = test.mat)
mean((ridge.pred - College.test$Apps)^2)
```

```
## [1] 940970.9
```

-> Ridge regression fit test error with a cross-validation based lambda = 940970.9 -> Lower than linear model test error

Part (d)

```
set.seed(1)

cv.out2 = cv.glmnet(train.mat, College.train$Apps, alpha = 1)
bestlam2 = cv.out2$lambda.min
bestlam2
```

```
## [1] 59.92044
```

```
lasso.mod = glmnet(train.mat, College.train$Apps, alpha = 1)
lasso.pred = predict(lasso.mod,s = bestlam2, newx = test.mat)
mean((lasso.pred - College.test$Apps)^2)
```

```
## [1] 993741.7
```

-> Lasso model fit test error with a cross-validation based lambda = 993741.7 -> Higher than linear model and ridge regression test error