STAT 440 Homework 11

Charlie Lu (Cxl5159)

October 2022

1 A

```
20
21  # model 1
22  set.seed(440)
23  cv.error <- rep(0, 2)
24  glm.fit <- glm(GPA ~ EntranceExam, data = samples)
25  cv.error[1] <- cv.glm(samples, glm.fit)$delta[1]
26
27  set.seed(440)
28  glm.fit <- glm(GPA ~ poly(EntranceExam, 2) , data = samples)
29  cv.error[2] <- cv.glm(samples, glm.fit)$delta[1]
30
31  cv.error
```

45:1 (Top Level)

Console Terminal Background Jobs

R 4.2.1 · C:/Users/Charlie Lu/Desktop/

```
> cv.error
[1] 0.02801622 0.03030053
```

The cross validation error for out first equation is lower than the the second equation that uses polynomial of degree 2

2 B

```
34  #3 fold cv
35  set.seed(440)
36  glm.fit <- glm(GPA ~ EntranceExam, data = samples)
37  cv.error[3] <- cv.glm(samples, glm.fit, K = 3)$delta[1]
38
39  set.seed(440)
40  glm.fit <- glm(GPA ~ poly(EntranceExam,2), data = samples)
41  cv.error[4] <- cv.glm(samples, glm.fit, K = 3)$delta[1]
42
43  cv.error
44
45
```

13:14 (Top Level)

Console Terminal Background Jobs

R 4.2.1 · C:/Users/Charlie Lu/Desktop/

```
> cv.error[4] <- cv.glm(samples, glm.fit, K = 3)$delta[1]
> cv.error
[1] 0.02801622 0.03030053 0.03437119 0.03452298
```

The cross validation error computed with the 3-fold CV for both equations is significantly higher than the leave one out cross validation method done in the previous section.