

LOOOCV.R

vijaykalmath

2022-01-04

```
# Leave One Out Cross Validation

# cv.glm does LOOCV and is part of boot library

library(boot)
library(ISLR)

glm.fit = glm(mpg~horsepower,data=Auto)

summary(glm.fit)

##
## Call:
## glm(formula = mpg ~ horsepower, data = Auto)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -13.5710  -3.2592  -0.3435   2.7630  16.9240
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 39.935861   0.717499   55.66  <2e-16 ***
## horsepower  -0.157845   0.006446  -24.49  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 24.06645)
##
##      Null deviance: 23819.0  on 391  degrees of freedom
## Residual deviance:  9385.9  on 390  degrees of freedom
## AIC: 2363.3
##
## Number of Fisher Scoring iterations: 2

# If K value is not given , cv.glm defaults to LOOCV
cv.err = cv.glm(Auto,glm.fit)

cv.err$delta

## [1] 24.23151 24.23114
```

```
cv_array = rep(0,5)

for (i in 1:10){
  glm.fit = glm(mpg~poly(horsepower,i),data = Auto)
  cv_array[i] = cv.glm(Auto,glm.fit)$delta[1]
}

cv_array
```

```
## [1] 24.23151 19.24821 19.33498 19.42443 19.03321 18.97864 18.83305 18.96115
## [9] 19.06863 19.49093
```

We can see that apart from a sharp drop of Error from linear to quadratic model, there are no signifi