

# #4 Auto Loss Analysis - Shahin Shakeri

## 1- Autoloss

### a) Using forward stepwise method:

```
Autoloss <- read.csv(url("https://drive.google.com/uc?export=download&id=1-QuNWq7k4w3c8kBJ8BedIK1m-cpw0"))
Autoloss <- na.omit (Autoloss)

library(leaps)
regfit.fwd=regsubsets(Losses~., data = Autoloss,nvmax=15, method="forward")
summary=summary(regfit.fwd)
preds=summary$which[10,][-1]
names(preds)[preds==TRUE]

## [1] "NumDoorstwo"      "BodyStylehatchback" "BodyStylesedan"
## [4] "BodyStylewagon"   "DriveWheelsrwd"     "Height"
## [7] "EngineSize"       "Horsepower"         "PeakRPM"
## [10] "Citympg"
```

The best model with 10 predictors has the above columns.

### b) What number of predictors minimizes Cp?

```
min=which.min(summary$cp)
min

## [1] 5

preds=summary$which[min,][-1]
names(preds)[preds==TRUE]

## [1] "NumDoorstwo"      "BodyStylesedan" "DriveWheelsrwd" "Height"
## [5] "Citympg"
```

The best model which minimizes Cp has the above 5 predictors

## 2- Autoloss

### a) LASSO model with 5-fold cross validation

```
library(glmnet)

## Loading required package: Matrix
## Loading required package: foreach
## Loaded glmnet 2.0-16
```

```

set.seed(1)
x=model.matrix(Losses~.,data=Autoloss)[,-1]
y=Autoloss$Losses
cv.out=cv.glmnet(x,y,alpha=1,nfolds=5)
bestLam=cv.out$lambda.min
bestLam

```

```
## [1] 2.153395
```

## b) LASSO with best lamda

```

cv.out=glmnet(x,y,alpha=1,lambda = bestLam )
coefs=coef(cv.out)
coefs

```

```

## 19 x 1 sparse Matrix of class "dgCMatrix"
##                               s0
## (Intercept)          2.939268e+02
## FuelTypegas           .
## Aspirationturbo       .
## NumDoorstwo          1.400973e+01
## BodyStylehardtop      .
## BodyStylehatchback    .
## BodyStylesedan        1.555973e+00
## BodyStylewagon       -1.105739e+01
## DriveWheelsfwd        .
## DriveWheelsrwd        1.618569e+01
## Length               .
## Width                .
## Height               -3.581782e+00
## Weight               .
## EngineSize           .
## Horsepower           .
## PeakRPM              5.611809e-03
## Citympg              -7.270771e-01
## Price                4.096412e-05

```

Predictors that are not zero are in the model with best lambda.

## 3 Answers for a dataset with P predictors

i True

ii True

iii True

iv True i tested with hitters

v True