PA1

alanpoon

Friday, March 27, 2015

Investigation of the type of transmission best for cars

Synopsis

In this report, analysis is drawn to determine which tranmission between automatic and manual is better for Miles per Gallon of fuel consumed by the different types of car using the mtcar dataset in regression model.

Regression model Theory Applied

The research problem of whether automatic or manual tranmission is better for miles per Gallon(mpg) can either be tackled using single variable linear model or multivarable linear model. There are many variables like weights, number of cyclinder, horse power and many more can also affect the mpg.

Factorizing some variables

We use str(mtcars) to take a look at the data. Those variables are binary, are factorized.

```
data(mtcars)
##str(mtcars)
mtcars$cyl <- factor(mtcars$cyl)
mtcars$vs <- factor(mtcars$vs)
mtcars$gear <- factor(mtcars$gear)
mtcars$carb <- factor(mtcars$carb)
mtcars$am <- factor(mtcars$am,labels=c('Automatic','Manual'))
```

Basic Linear model unadjusted

We define the null hypothesis that there is strong correlation between the tranmission type with the mpg. We will test the hypothesis under 95% confidence and reject H(0) if p-value is less than alpha, which is 0.05.

```
lmfit<-lm(mpg~.,data=mtcars)
```

Using summary(lmfit), we see that the slope cofficient is 7.245 with 17.147 as intercept with p-value:0.000285. We can reject the H(0).

Next,we do a backward elimation to find a combination of variables that gives the least p-values. Consideration of these variables gives the highest p-values: cyl + hp + wt + am.

```
##lmfit2 <- step(lmfit, direction = "backward")
```

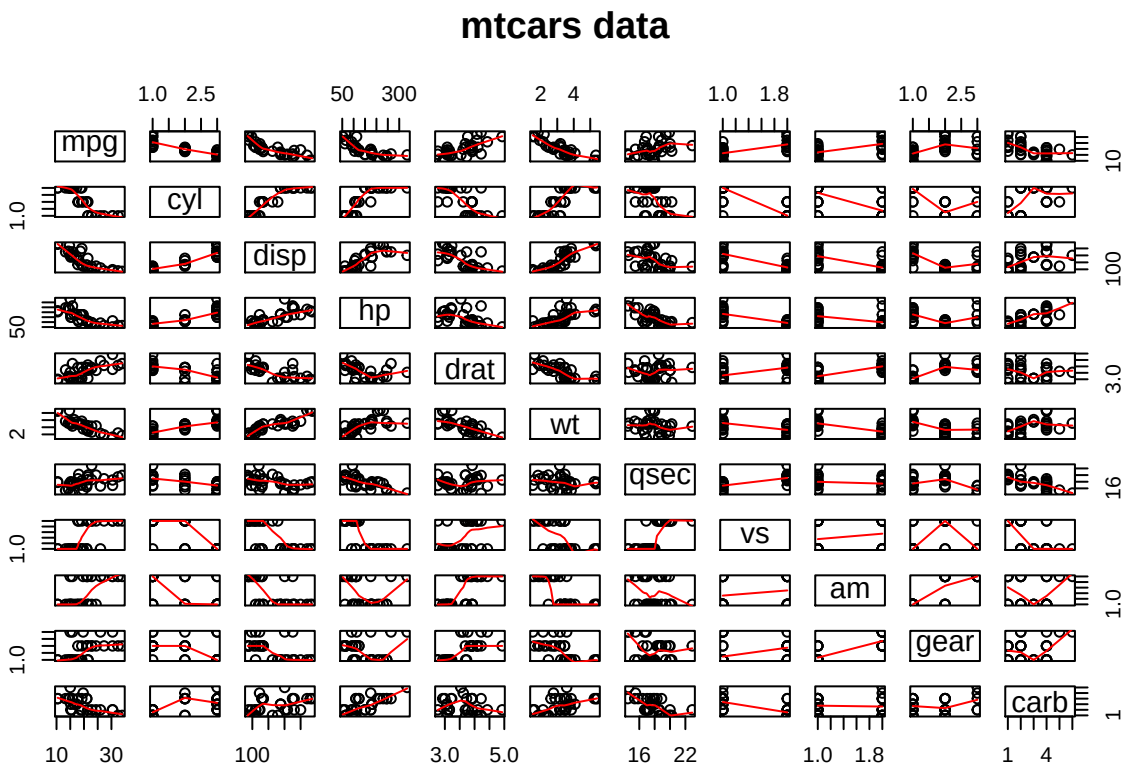
summary(lmfit2) shows these combination of variables yearn p-value: 1.506e-10. The linear model suggests a decrease in the mpg by Miles per gallon may be constituented by the a increase of (3.03 of cylinder16),(2.16 of cylinder),(0.03 of horsepower) and (2.49 of weight). On the other hand, an increase of miles per gallon can be constituented by having manual transmission (by 1.8)

Multivariable Linear model regression

Many other variables are required to be kept constant. In this case, we need to add more regressors to generalize the simple linear regressors in order to get the purpose of prediction. However, there are costs adding more regressors that may not really be related to mpg.

A preliminary graph can be plotted to have an overview of the inter-correlation of the variables.

```
require(stats)
require(graphics)
pairs(mtcars,panel=panel.smooth,main="mtcars data")
```



T-test Using t-test statistic, we will be able find out the confidence interval for the mpg using transmission type.

```
##
##  Welch Two Sample t-test
##
## data:  mpg by am
## t = -3.7671, df = 18.332, p-value = 0.001374
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.280194  -3.209684
## sample estimates:
## mean in group Automatic    mean in group Manual
##                17.14737                24.39231
```

The 95% confidence interval is -11.28 to -3.21.

Conclusion

In conclusion, according to our findings, cars with manual transmission are better in mpg than auto transmission. The rate of change of the mean miles per gallon with respect to auto/manual tranmission type is about 1.8.