Motor Trend

HTKeng

3/29/2020

Introduction

In this project, we aimed to explore the relationship between a set of variables and miles per gallon (MPG) (outcome) in the **mtcars** dataset and answer two questions: 1. Is an automatic or manual transmission better for MPG? and 2. Quantify the MPG difference between automatic and manual transmissions.

Method

We performed exploratory analyses to see the dataset and the correlation of each variables. Univariate analysis of transmission was performed using student t test. Multivariate linear regression with stepwise feature selection was used for model derivation. Residual plots were visulized for indentify potential bias.

Results

1. Univariate analysis

The mean mpg of the manual transmission is significantly higher to that of the automatic transmission.

```
t.test(mpg~am, mtcars)
```

```
##
##  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 0 mean in group 1
##        17.14737        24.39231
```

2. Multivariate liner regression with stepwise feature selection.

The model included wt, am, qsec as predicting features. The Adjusted R-squared was 0.8336 To opitimize the model, we added the interaction of am with wt and qsec to our model and the Adjusted R-squared was 0.8767.

```r
mtcars$am <- as.factor(mtcars$am)
levels(mtcars$am) <-c("Automatic", "Manual")
fitstep <- step(lm(mpg~., mtcars), direction = "both", trace = 0)
coef(fitstep)
```

```
## (Intercept)          wt        qsec    amManual
##    9.617781   -3.916504    1.225886    2.935837
```

```r
fit_interaction <- lm(mpg~ + am*wt + am*qsec, mtcars)
summary(fit_interaction)
```

```
##
## Call:
## lm(formula = mpg ~ +am * wt + am * qsec, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6832 -1.3222 -0.3747  1.0687  4.0907
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    11.2489     6.9922   1.609 0.119738
## amManual        8.9265    12.6662   0.705 0.487232
## wt             -2.9963     0.6910  -4.336 0.000194 ***
## qsec            0.9454     0.3067   3.082 0.004813 **
## amManual:wt    -3.7581     1.5158  -2.479 0.019969 *
## amManual:qsec   0.2355     0.5566   0.423 0.675630
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.117 on 26 degrees of freedom
## Multiple R-squared:  0.8966, Adjusted R-squared:  0.8767
## F-statistic: 45.07 on 5 and 26 DF,  p-value: 5.394e-12
```

Discussion and Conclusion

In this project, we found the mean mpg of the manual transmission is significantly higher which suggests the manul transmission is better for MPG when other variables remain constant. Furthermore, we developed a model with adjusted R-squared equaled 0.88. By adding the coefficients, we had the following conclusions:
1. Holding qsec constant, when the weight increase by 1000 lbs, the mpg decrease by 2.9963(beta2 = -2.9963) for automatic transmission cars, and 6.7544 (beta2+beta4 = -2.9963-3.7581) for manual transmission cars
2. Holding weight constant, when the 1/4 mile time increase by 1 sec , the mpg increase by 0.9454(beta3 = 0.9454) for automatic transmission cars, and 1.1809 (beta3+beta5 = 0.9454+0.2355) miles for manual transmission cars.
3. The mean mpg of manual transmission is higher than that of automatic transmisson by 8.93 miles.

Appendix.1 Correlation plot

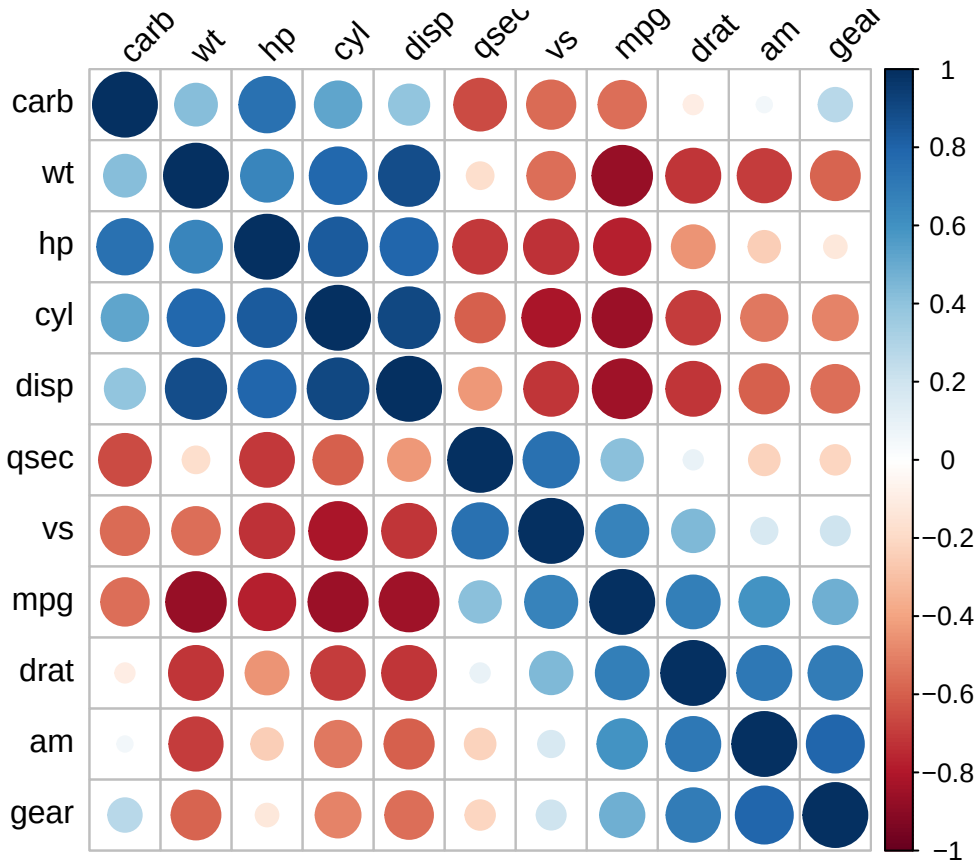
```r
data("mtcars")
head(mtcars)
```

```
##                    mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

```r
library(corrplot)
```

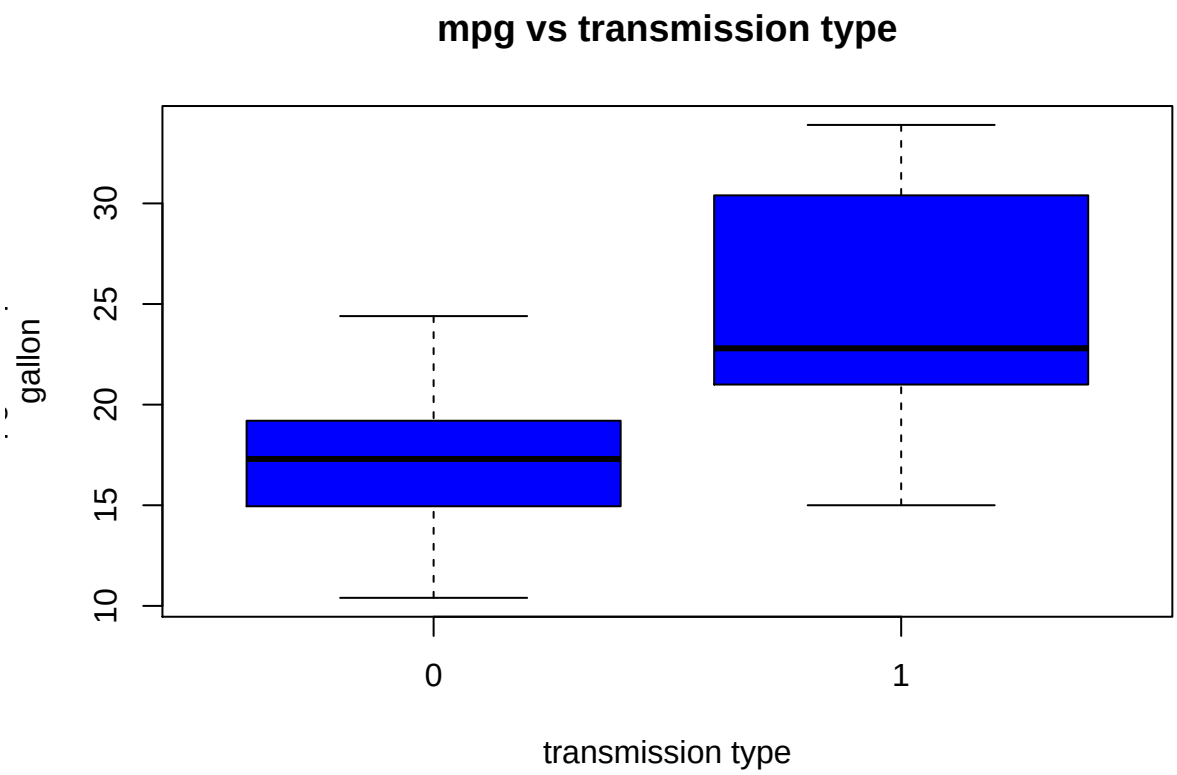
```
## corrplot 0.84 loaded
```

```r
res <- cor(mtcars)
corrplot(res, order = "hclust",tl.col = "black", tl.srt = 45)
```



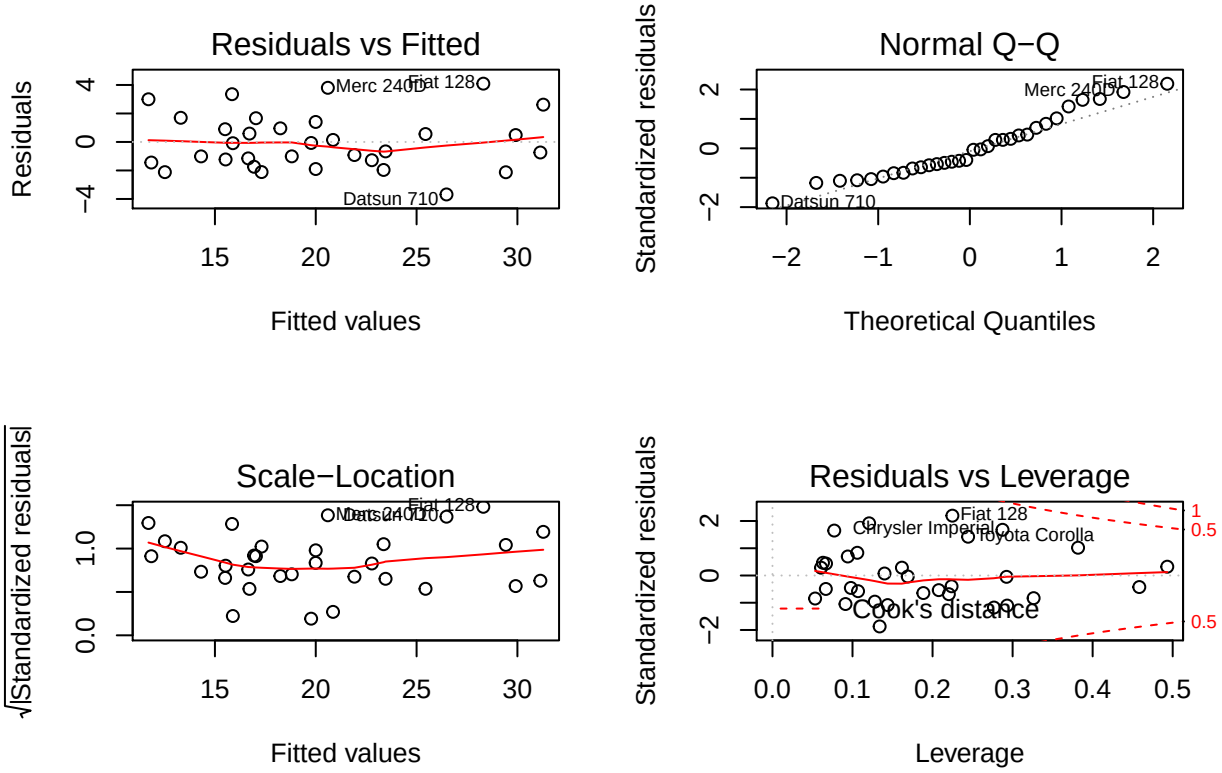
Appendix.2 Barplox of transsmisson

```r
boxplot(mtcars$mpg ~ mtcars$am, data = mtcars, outpch = 19, ylab="mpg:miles per
gallon",xlab="transmission type",main="mpg vs transmission type", col="blue")
```



Appendix.3 Residual diagnostic plots

```r
par(mfrow=c(2,2))
plot(fit_interaction)
```


Residuals vs Fitted
Residuals
Fitted values
Merc 240D
Fiat 128
Datsun 710
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Residuals vs Leverage
Cook's distance
Chrysler Imperial
Toyota Corolla
Leverage