Motor Trend - Does Transmission Affect Your MPG

Mohan

February 22, 2015

Introduction

Planning to buy a new car and you are undecided between automatic and manual transmission? Read the report below. We will look at how MPG affects between automatic and manual transmission and if its a myth. Looking at a data set of a collection of cars, we are interested in exploring the relationship between a set of variables and miles per gallon (MPG) (outcome). * "Is an automatic or manual transmission better for MPG" * "Quantify the MPG difference between automatic and manual transmissions"

Data

We have obtained the data from 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973-74 models).

We load the data as such.

```
library(datasets)
data(mtcars)
```

Converting variables to factor variables.

```
# Converting to factors
mtcars$cyl <- factor(mtcars$cyl)
mtcars$vs <- factor(mtcars$vs)
mtcars$am <- factor(mtcars$am, labels = c("Automatic", "Manual"))
mtcars$gear <- factor(mtcars$gear)
mtcars$carb <- factor(mtcars$carb)
```

Analysis

A boxplot was produced to show the difference between automatic and manual in terms of MPG. In figure 1, it is clear that manual transmission produces more MPG. Next, a pairwise graph (figure 2) was created in order to get a greater intuition of what other variables may be of interest. There is a linear relationship between MPG and each of cyl, disp, hp, drat, wt, qsec, vs, am. The covariance was also computed (figure 3) between every variable and the positive values were noted (qsec = 0.419, vs = 0.664, am = 0.600, gear = 0.480). Then a linear model was fit on all the variables to determine which variables should be used in the final models. In figure 4 the summary from this model is shown. The lowest p values were taken (i.e. wt = 0.063, am = 0.234, qsec = 0.274) due to their high significance in predicting MPG.

Model

From the initial model, covariance test and visually inspecting the pairwise graph the following variables stood out in particular: qsec, vs, am, wt and gear. Next a stepwise model process was used in order to obtain the most significant predictors to be used. As shown in figure 5, the most significant predictors in determining the MPG are cyl, hp, wt and am. The summary for this model is also shown. Next, the new model was compared with a basic model that only uses transmission type as its predictor. A p-value of 1.688e-08 was obtained (figure 6). This value is negligible indicating that the predictors are good for the model's accuracy.

```
# step wise selection process
modelStep <- step(modelFitAll, trace = 0)
```

Diagnostics

The residuals from the final model are plotted below (figure 8).

- The Residuals vs Fitted plot shows no pattern between the residuals and fitted values indicating that this regression model is well fit.

- The QQ plot shows that the points line up as expected meaning that the distribtion is normal and our model predictions are accurate.

- In both the Scale-Location plot and the Residuals vs Leverage plots, the points are in a group with none too far from the center indicating no point had too much leverage.

Statistical Inference

A Two Sample t-test was conducted between the different transmission types. The null hypothesis that transmission types don't have an effect on the MPG is discarded for a p-value greater than 0.05. The results are shown in figure 8. The p-value of 0.001374 and difference in means show that manual transmission has significantly more MPG than automatic.

```
test <- t.test(mpg ~ am, data = mtcars)
```

Conclusions

The transmission type of a car has a significant effect on its fuel efficiency. According to the model, manual transmission, on average, has 1.81 MPG more than automatics. According to the boxplot, manual transmission has ~ 6 MPG more than automatics.

Appendix

```
boxplot(mpg ~ am, data = mtcars,
        xlab = "Transmission", ylab = "MPG",
        main = "MPG vs Transmission", col = c("red", "green"),
        names = c("Automatic", "Manual"))
```

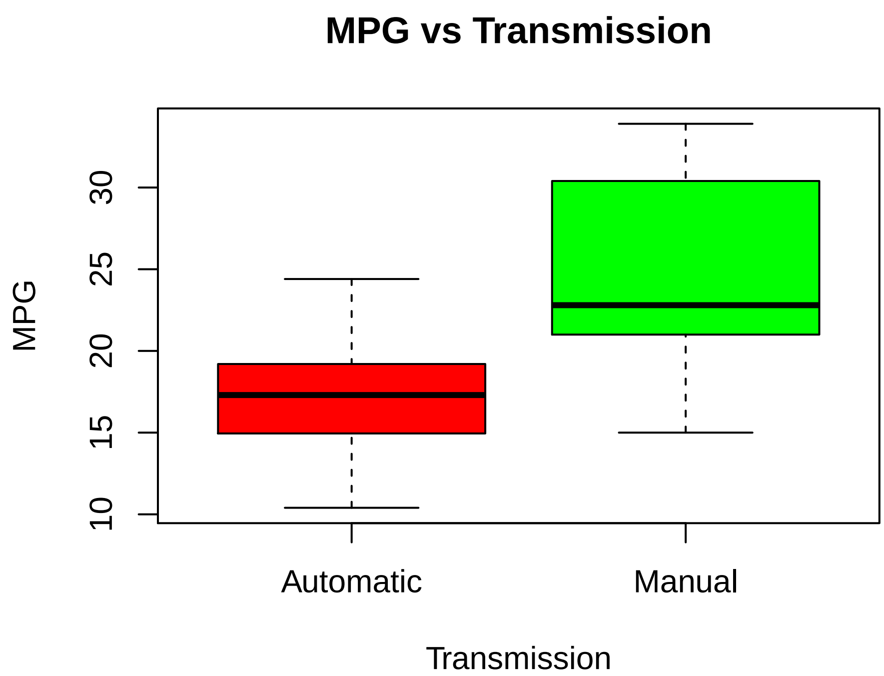

Figure 1.

```r
p1 = pairs(mtcars, panel = panel.smooth, main = "Pairwise plot of mtcars data")
```

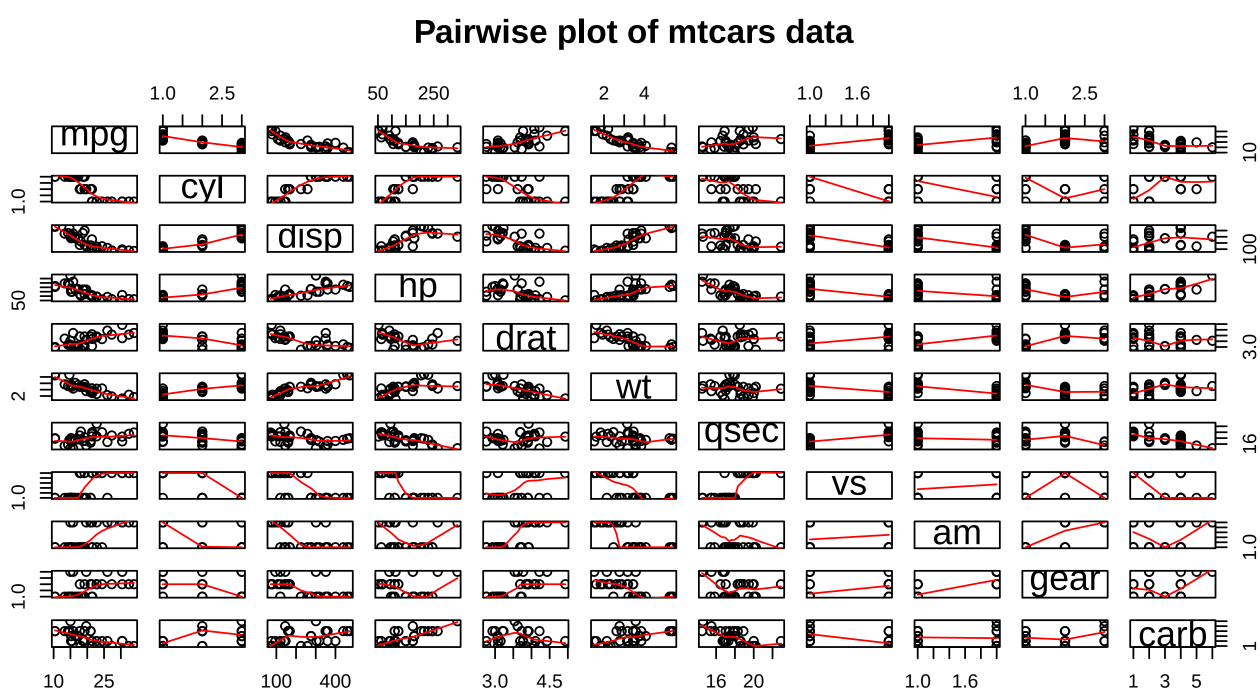

Figure 2.

```r
head(cov2cor(cov(sapply(mtcars, as.numeric))), 1)
```

Figure 3.

```
##     mpg       cyl       disp         hp      drat         wt     qsec
## mpg   1 -0.852162 -0.8475514 -0.7761684 0.6811719 -0.8676594 0.418684
##            vs        am      gear       carb
## mpg 0.6640389 0.5998324 0.4802848 -0.6067431
```

```
modelFitAll$coeff
```

Figure 4.

```
## (Intercept)        cyl6        cyl8        disp          hp        drat
## 23.87913244 -2.64869528 -0.33616298  0.03554632 -0.07050683  1.18283018
##          wt        qsec         vs1    amManual       gear4       gear5
## -4.52977584  0.36784482  1.93085054  1.21211570  1.11435494  2.52839599
##       carb2       carb3       carb4       carb6       carb8
## -0.97935432  2.99963875  1.09142288  4.47756921  7.25041126
```

```
summary(modelStep)$coef
```

Figure 5.

```
##                Estimate Std. Error   t value     Pr(>|t|)
## (Intercept) 33.70832390 2.60488618 12.940421 7.733392e-13
## cyl6        -3.03134449 1.40728351 -2.154040 4.068272e-02
## cyl8        -2.16367532 2.28425172 -0.947214 3.522509e-01
## hp          -0.03210943 0.01369257 -2.345025 2.693461e-02
## wt          -2.49682942 0.88558779 -2.819404 9.081408e-03
## amManual     1.80921138 1.39630450  1.295714 2.064597e-01
```

```
modelStep$coeff
```

```
## (Intercept)        cyl6        cyl8          hp          wt    amManual
## 33.70832390 -3.03134449 -2.16367532 -0.03210943 -2.49682942  1.80921138
```

```
compare <- anova(modelFitMileage, modelStep)
compare$Pr
```

Figure 6.

```
## [1]           NA 1.688435e-08
```

```
modelFitMileage
```

Figure 7.

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## Coefficients:
## (Intercept)     amManual
##      17.147        7.245
```

```
par(mfrow=c(2, 2))
plot(modelStep)
```

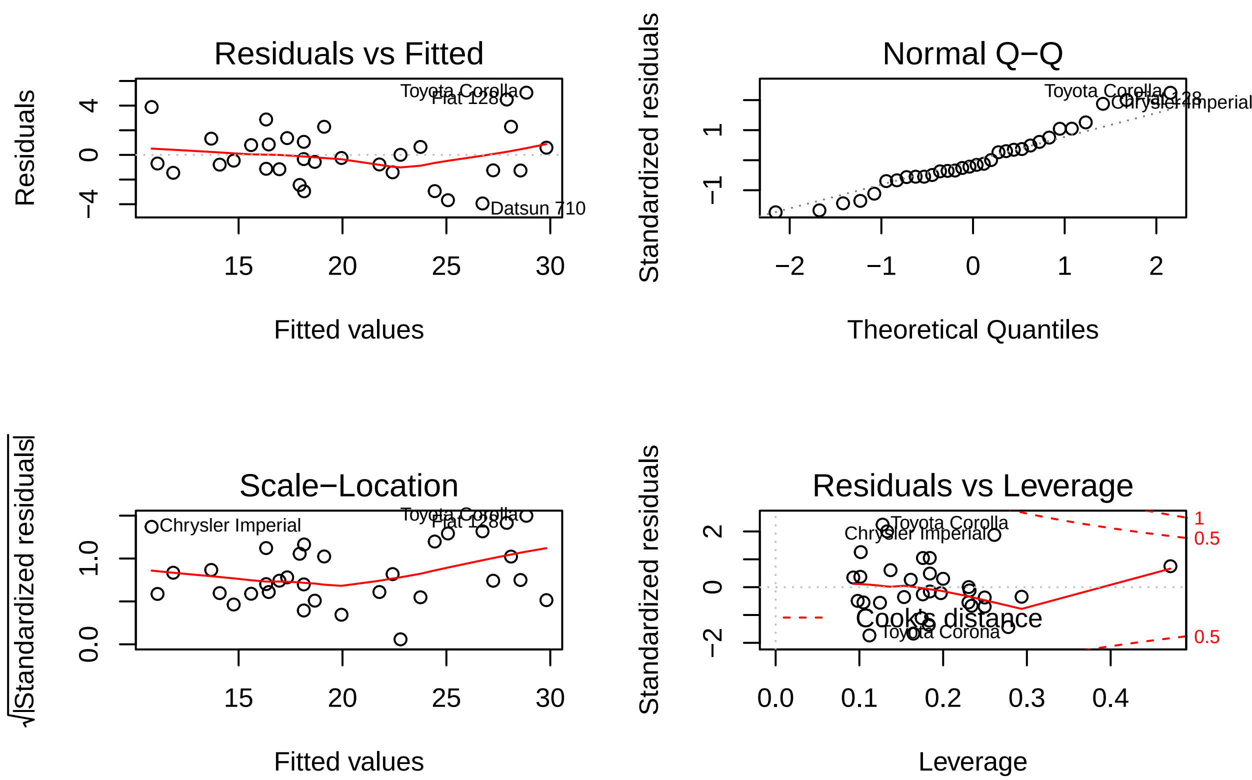

Figure 8.