Comparison of MPG for manual and automatic transmisson

Kiryl Batsiukov

02/18/2015

This is a regression analysis that tries to respond to the following 2 questions:

- Is an automatic or manual transmission better for MPG ?
- Quantify the MPG difference between automatic and manual transmissions

Executive Summary

On average, similar cars will have 2.94 more MPG with manual transmission. With 95% confidence, the increase in MPG for manual transmission will lie between 0.05 and 5.83 MPG. So, manual tranmission is better for MPG.

Exploratory Data Analysis

Please refer to the mtcars dataset help file in R for details on our data source.

```
require(datasets)
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
```

Box plot shows MPG for different transmission types in Appendix 1.

Regression Model

A simple modell for mpg with transmission as the only one predictor

```
simpleModel <- lm(mpg ~ factor(am), data = mtcars)
coef(summary(simpleModel))
```

```
##              Estimate Std. Error   t value     Pr(>|t|)
## (Intercept) 17.147368   1.124603 15.247492 1.133983e-15
## factor(am)1  7.244939   1.764422  4.106127 2.850207e-04
```

```r
summary(simpleModel)$r.squared
```

```
## [1] 0.3597989
```

We conclude that on average MPG of cars with manual transmission type is 7.24 higher than with automatic transmision type, with significant P-Value. However, our model explains only 34% of variance. What is the influece of transmission in conjunction with other factors? The answer will help quantify the mpg difference between automatic and manual transmission more exactly. Based on our domain knowledge in cars, we propose more predictors with low covariance: horse power, weight and transmission type. They should explain more variance in the model.

```r
cor(mtcars[,c(7,6,9)])
```

```
##            qsec         wt         am
## qsec  1.0000000 -0.1747159 -0.2298609
## wt   -0.1747159  1.0000000 -0.6924953
## am   -0.2298609 -0.6924953  1.0000000
```

```r
complexModel<-lm(mpg~factor(am)+qsec+wt,mtcars)
coef(summary(complexModel))
```

```
##              Estimate Std. Error   t value     Pr(>|t|)
## (Intercept)  9.617781  6.9595930  1.381946 1.779152e-01
## factor(am)1  2.935837  1.4109045  2.080819 4.671551e-02
## qsec         1.225886  0.2886696  4.246676 2.161737e-04
## wt          -3.916504  0.7112016 -5.506882 6.952711e-06
```

```r
summary(complexModel)$r.squared
```

```
## [1] 0.8496636
```

The multivariate model with 3 predictors explains 83% of variance, which is a better result. All predictors have significant T-Values. The residuals plot (in appendix) has no specific patterns. On average, similar cars will have 2.94 more MPG with manual transmission. This effect is much lower than for simple model, where qsec and wt are not considered. With 95% confidence, the increase in MPG for manual transmission will lie between 0.05 and 5.83 MPG:

```r
confint(complexModel)[2,]
```

```
##      2.5 %     97.5 %
## 0.04573031 5.82594408
```

To confirm that the last model should be selected we use anova method:

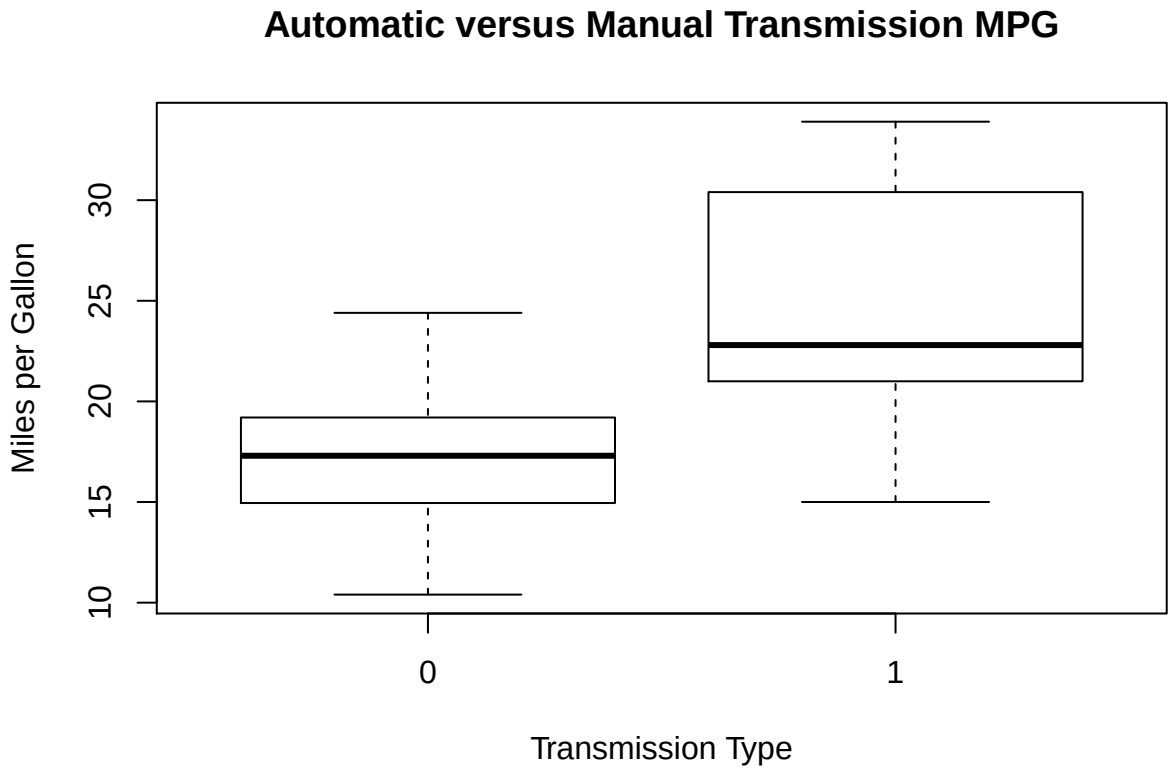
```r
anova(simpleModel, complexModel)[c(4,6)]
```

```
##   Sum of Sq   Pr(>F)
## 1
## 2    551.61 1.55e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The multivariate model is our final selection.

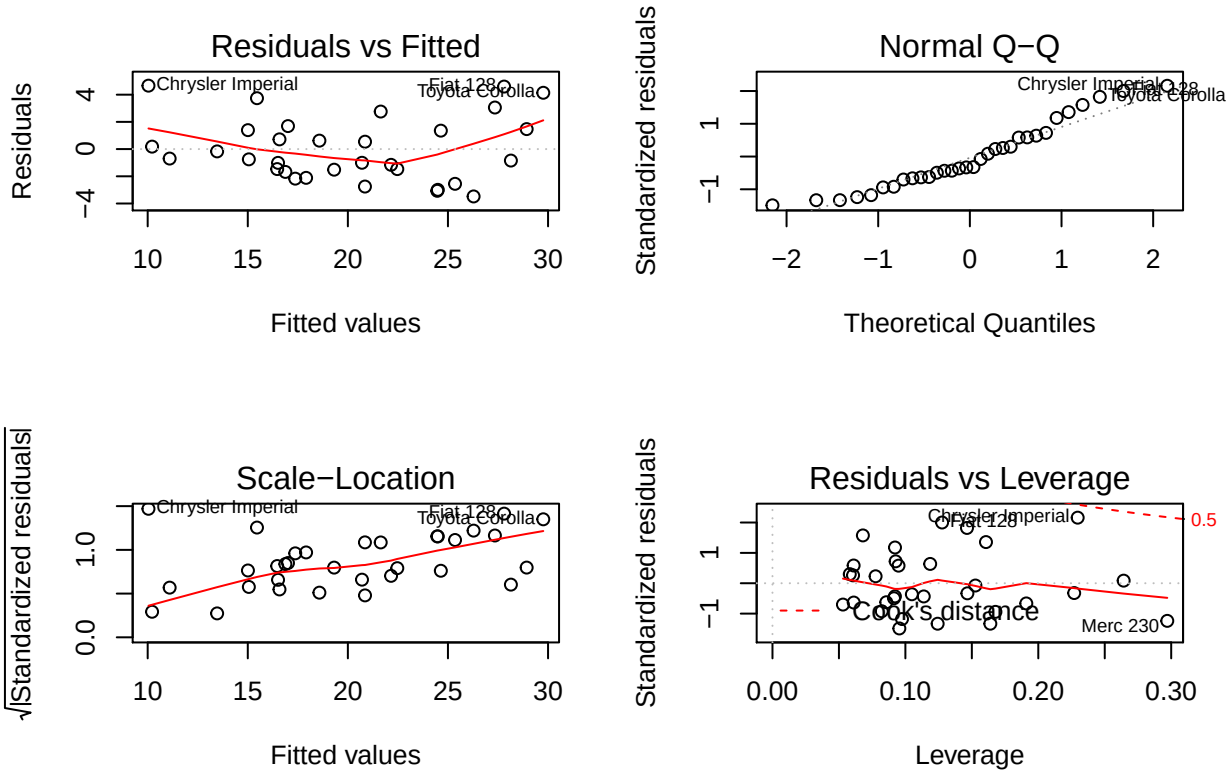
Appendix 1. Box Plot of MPG for different transmission types

```
boxplot(mpg ~ am, data=mtcars, xlab="Transmission Type", ylab="Miles per Gallon",
        main="Automatic versus Manual Transmission MPG")
```



Appendix 2. Residuals Plot for complex model

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


Residuals vs Fitted
Residuals
Fitted values
Chrysler Imperial
Fiat 128
Toyota Corolla
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Leverage
Cook's distance
0.5
Merc 230