Automatic versus Manual Transmission for Fuel Efficiency

Li Xu

Executive Summary

This report answers the following two questions: 1. Is an automatic or manual transmission better for MPG? 2. Quantify the MPG difference between automatic and manual transmissions. Here we study "mtcars" data set and show that manual tranmission is better for fuel efficiency than automatic transmission. Also we quantify the difference between these two ways of transmission in several models.

Data Set

The data set "mtcars" includes 32 observations on 11 variables, and was extracted from the 1974 Motor Trend US magazine. It has been bundled in RStudio, and its description can be found at http://stat.ethz.ch/R-manual/R-devel/library/datasets/html/mtcars.html. Another way to obtain its information is simply typing "?mtcars" in the RStudio console.

Exploratory Data Analysis

We conduct some exploratory analysis on "mtcars". First we draw the scatterplot figure with respect to "mpg" and "am" (See Figure 1). Then we fit the linear regression model of "mpg" onto "am", and get the boxplot graph (see Figure 2).

Single Variable Case

We fit the linear regression model of "mpg" onto single variable "am". This linear regression model can be obtained by "**lm()**" function. Typing "**SingleModel<-lm(mpg~am, data=mtcars); summary(SingleModel)$coefficients**" gives the summary of "model1" as follows:

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   17.147      1.125  15.247 1.13e-15
am             7.245      1.764   4.106 0.000285
```

Interpret Coefficients

From the above summary, we find that the manual transmission makes a 7.245 miles per gallon increase compared to the automatic way. Since the coefficient of "am" is 7.245, a positive number, this variable has a positive effect on fuel efficiency. Therefore, we conclude that manual transmission has better fuel efficiency than automatic way.

Quantify Uncertainty:

The coefficients of intercept and slope are very large relative to their estimation standard error. And the p values for two coefficients are both <0.001. Hence we could reject the two null hypotheses (slope==0 and

intercept==0), and conclude that both the intercept and slope are non-zero. Then this conclusion leads us to infer that the "am" and "mpg" has a significant relationship. And manual transmission has better fuel efficiency.

Two Variables Case

We consider one more variable "cyl" and examine the relationship between mpg and two variables "am", "cyl". The scatterplot can be found in Figure 3. Type "**PairModel<-lm(mpg~am+cyl,data=mtcars); summary(PairModel)$coefficients**", and then we obtain that

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  34.5224     2.6032  13.262 7.69e-14
am            2.5670     1.2914   1.988   0.0564
cyl          -2.5010     0.3608  -6.931 1.28e-07
```

Interpret Coefficients

This regression result shows that with a fixed number of cylinders, manual transmission has 2.57 higher fuel efficiency than automatic transmission.

Quantify Uncertainty

Here the significance of "am" is 0.0564. So the effect of manual transmission is not statistically significant in this model.

Multiple variables case

Here all variables are used to fit the linear regression for mpg. Type "**TotalModel<-lm(mpg~.,data=mtcars); summary(TotalModel)$coefficients** ", and we obtain the estimation of each variable (See Summary 1). It shows that with other nine variables fixed, the number of miles per gallon is increased by 2.52 when we conduct manual transmission, compared to the automatic transmission.

Model Selection

We start at a model of "mpg" onto all variables. Then to find an optimal model, we make backward selection of these variables using Akaike Information Criterion (AIC). Typing "**allvariablesmode<-lm(mpg~.,data=mtcars), backwardmodel<-step(allvariablesmodel, direction="backward")**", we obtain the optimal model "mpg ~ wt + qsec + am". Its summary can be found in Summary 2.

Residuals Analysis and Diagnostics

We examine the residuals plot as Figure 5. It can be shown that the residual plots are randomly around the x-axis. So a linear regression model is appropriate for the "mtcars" data set.

Answers to the Questions

After the analysis of multiple models, we can conclude that manual tranmission is better for fuel efficiency than automatic transmission. In the single variable model, the difference between these two ways of transmission in several models is 7.245 miles per gallon with a significance level 0.000285 (<0.001, very strongly significant).

Appendix

Figure 1

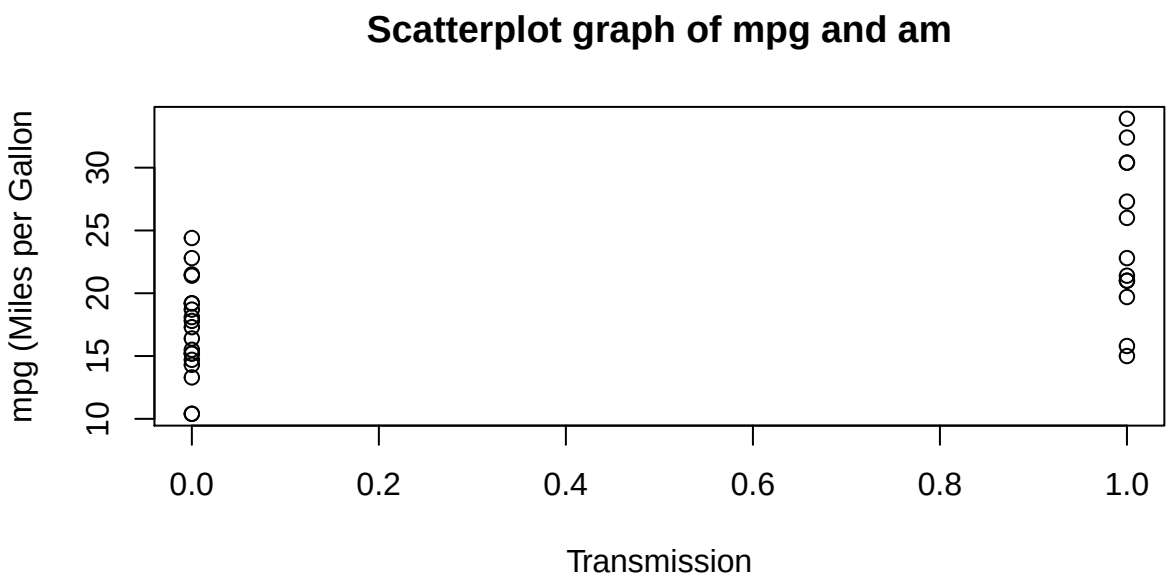


Figure 2

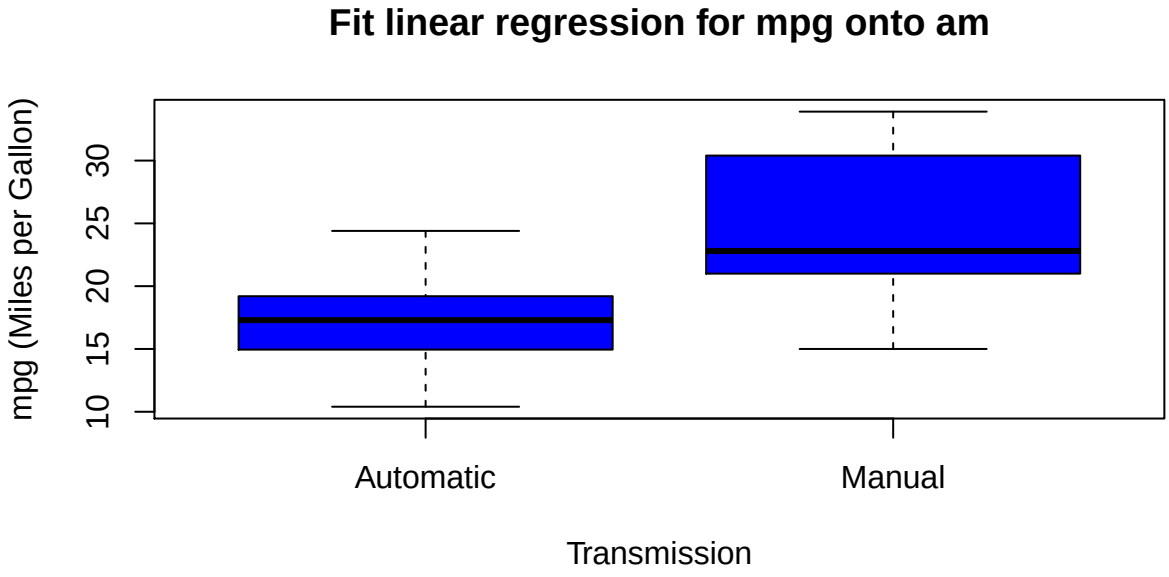
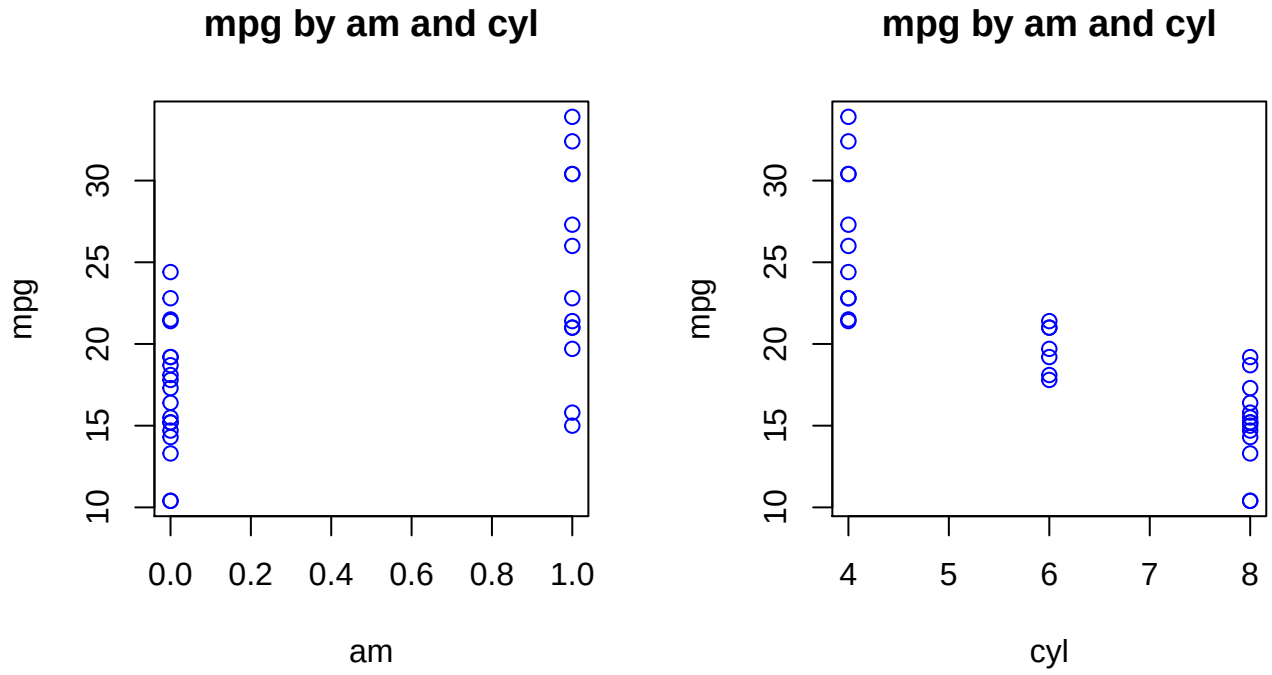


Figure 3



Summary 1

```
               Estimate  Std. Error    t value   Pr(>|t|)
(Intercept) 12.30337416 18.71788443  0.6573058 0.51812440
am           2.52022689  2.05665055  1.2254035 0.23398971
cyl         -0.11144048  1.04502336 -0.1066392 0.91608738
disp         0.01333524  0.01785750  0.7467585 0.46348865
hp          -0.02148212  0.02176858 -0.9868407 0.33495531
drat         0.78711097  1.63537307  0.4813036 0.63527790
wt          -3.71530393  1.89441430 -1.9611887 0.06325215
qsec         0.82104075  0.73084480  1.1234133 0.27394127
vs           0.31776281  2.10450861  0.1509915 0.88142347
gear         0.65541302  1.49325996  0.4389142 0.66520643
carb        -0.19941925  0.82875250 -0.2406258 0.81217871
```

Summary 2

```
Call:
lm(formula = mpg ~ wt + qsec + am, data = mtcars)

Residuals:
    Min      1Q  Median      3Q     Max
-3.4811 -1.5555 -0.7257  1.4110  4.6610
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   9.6178     6.9596   1.382 0.177915
wt           -3.9165     0.7112  -5.507 6.95e-06 ***
qsec          1.2259     0.2887   4.247 0.000216 ***
am            2.9358     1.4109   2.081 0.046716 *
---

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 2.459 on 28 degrees of freedom
Multiple R-squared:  0.8497,  Adjusted R-squared:  0.8336
F-statistic: 52.75 on 3 and 28 DF,  p-value: 1.21e-11
```

Figure 5