

Automatic vs. Manual transmission for better MPG

Junho Song

22 July, 2014

Contents

1	Executive Summary	1
2	Data Processing	1
2.1	Load and Check the Data Set	1
2.2	Data Processing and Exploratory Data Analysis	2
3	Inference	2
3.1	Fist question	2
4	Model Selection	3
4.1	Second question	3
5	Appendix	4

1 Executive Summary

The MPG(miles per gallon) of vehicle is affected by many factors like cylinders, weight, gears, hourse power, automatic trasmission etc. In this study, the data set extracted from the 1974 Motor Trend US magazine is used, and I focused on the difference of MPG between automatic and manual transmission.

Main questions are below.

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

2 Data Processing

2.1 Load and Check the Data Set

Check the data set for processing. And for better understanding as a whole, we can see the whole relationship with each variables at fig1(see the appendix).

```
library(datasets)
data(mtcars)
#str(mtcars)
```

2.2 Data Processing and Exploratory Data Analysis

We can see the ‘cyl’, ‘vs’, ‘am’, ‘gear’ and ‘carb’ which are numerical variables. But we can use it as factors for analysis about relationships with mpg. And check the boxplot as relation between mpg and other variables at fig2(see appendix).

```
library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##     filter, lag
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

mpgMtcars <- select(mtcars, mpg, cyl, vs, am, gear, carb)
mpgMtcars$cyl <- as.factor(mpgMtcars$cyl)
mpgMtcars$vs <- as.factor(mpgMtcars$vs)
mpgMtcars$am <- as.factor(mpgMtcars$am)
mpgMtcars$gear <- as.factor(mpgMtcars$gear)
mpgMtcars$carb <- as.factor(mpgMtcars$carb)
#str(mpgMtcars)
```

3 Inference

We will focus on the variable ‘am’, automatic transmission vs. manual transmission (factor ‘0’ is automatic, ‘1’ is manual).

3.1 First question

“Is an automatic or manual transmission better for MPG”

```
mpgAmMtcars <- select(mtcars, mpg, am)
mpgAmMtcars$am <- factor(mpgAmMtcars$am, labels = c("Automatic", "Manual"))

# mpg vs. Automatic transmission
summary(mpgAutoMtcars <- subset(mpgAmMtcars, am == "Automatic"))
```

```
##           mpg           am
## Min.      :10.4   Automatic:19
## 1st Qu.:14.9   Manual    : 0
## Median :17.3
## Mean      :17.1
## 3rd Qu.:19.2
## Max.      :24.4
```

```
# mpg vs. Manual transmission
summary(mpgManuMtcars <- subset(mpgAmMtcars, am == "Manual"))
```

```
##           mpg           am
## Min.      :15.0   Automatic: 0
## 1st Qu.:21.0   Manual    :13
## Median :22.8
## Mean     :24.4
## 3rd Qu.:30.4
## Max.     :33.9
```

According to the above result, the mean of mpg in manual is better ; 17.15 (automatic) vs. 24.39 (manual). Therefore, we can say manual transmission for mpg is better, and check it at fig3 (see appendix).

Furthermore, we can check the difference as t-test. - H0 : no difference between automatic transmission in terms of mpg - H1 : difference between automatic transmission in terms of mpg

```
mpgAmTtest <- t.test(x = mpgAutoMtcars$mpg,
                    y = mpgManuMtcars$mpg, paired = FALSE)
mpgAmTtest$p.value
```

```
## [1] 0.001374
```

The p-value is 0.00137. Considering the critical value 0.05, the null hypothesis (H0) is should be rejected and we can say there is a statistical significant mpg difference between automatic and manual transmission.

4 Model Selection

4.1 Second question

“Quantify the MPG difference between automatic and manual transmissions”

```
fitMtcars <- lm(mpg ~ as.integer(am), data = mpgMtcars )
summary(fitMtcars)
```

```
##
## Call:
## lm(formula = mpg ~ as.integer(am), data = mpgMtcars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.392 -3.092 -0.297  3.244  9.508
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      9.90        2.63   3.77 0.00072 ***
## as.integer(am)    7.24        1.76   4.11 0.00029 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 4.9 on 30 degrees of freedom
## Multiple R-squared:  0.36,    Adjusted R-squared:  0.338
## F-statistic: 16.9 on 1 and 30 DF,  p-value: 0.000285
```

According to the summary, we can check the p-value 0.000285, suggests that the difference can be explained by this model statistically. And in terms of mpg, manual transmission has better efficiency, which increased by 7.245.

5 Appendix

fig1. Relationships among variables

```
require(stats);require(graphics)
pairs(mtcars, panel=panel.smooth, main="Relationships among variables", col = 3 + (mtcars$mpg >20))
```

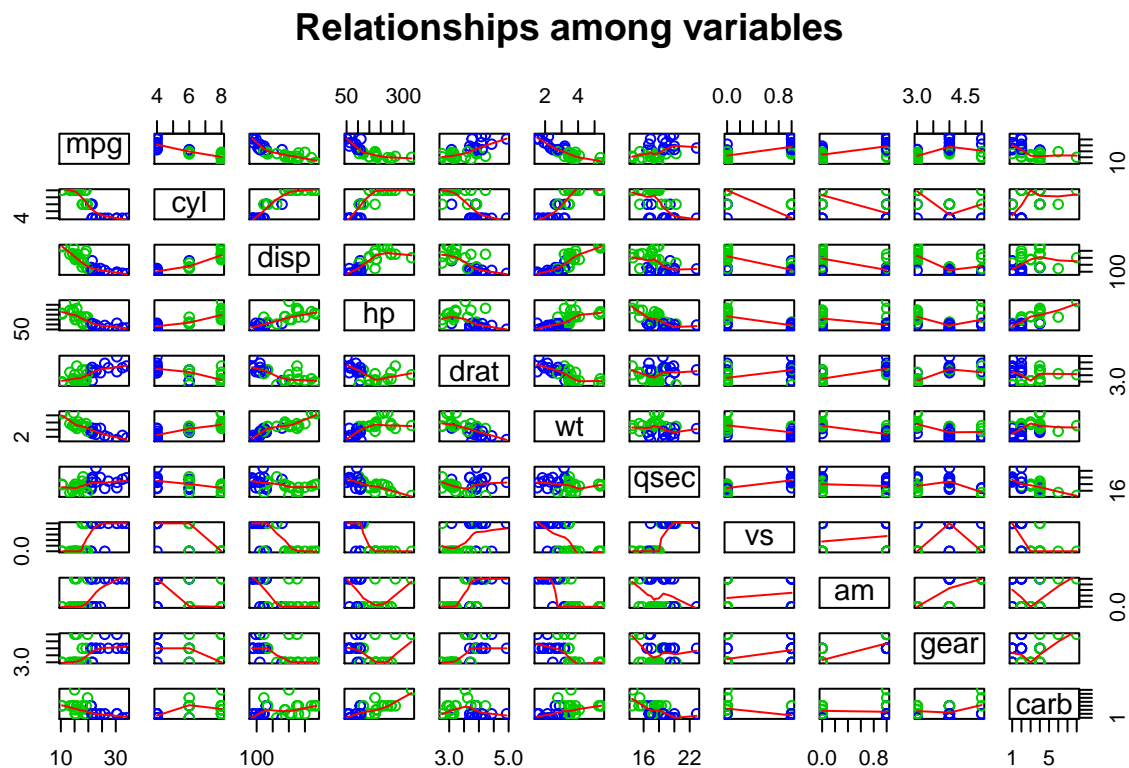


fig1. mpg and other variables

```
## boxplot for checking difference
par(mfrow = c(1,5))
boxplot(mpg ~ cyl, data = mpgMtcars, ylab = "mpg", xlab = "cyl")
boxplot(mpg ~ vs, data = mpgMtcars, ylab = "mpg", xlab = "vs")
boxplot(mpg ~ am, data = mpgMtcars, ylab = "mpg", xlab = "am")
boxplot(mpg ~ gear, data = mpgMtcars, ylab = "mpg", xlab = "gear")
boxplot(mpg ~ carb, data = mpgMtcars, ylab = "mpg", xlab = "carb")
```

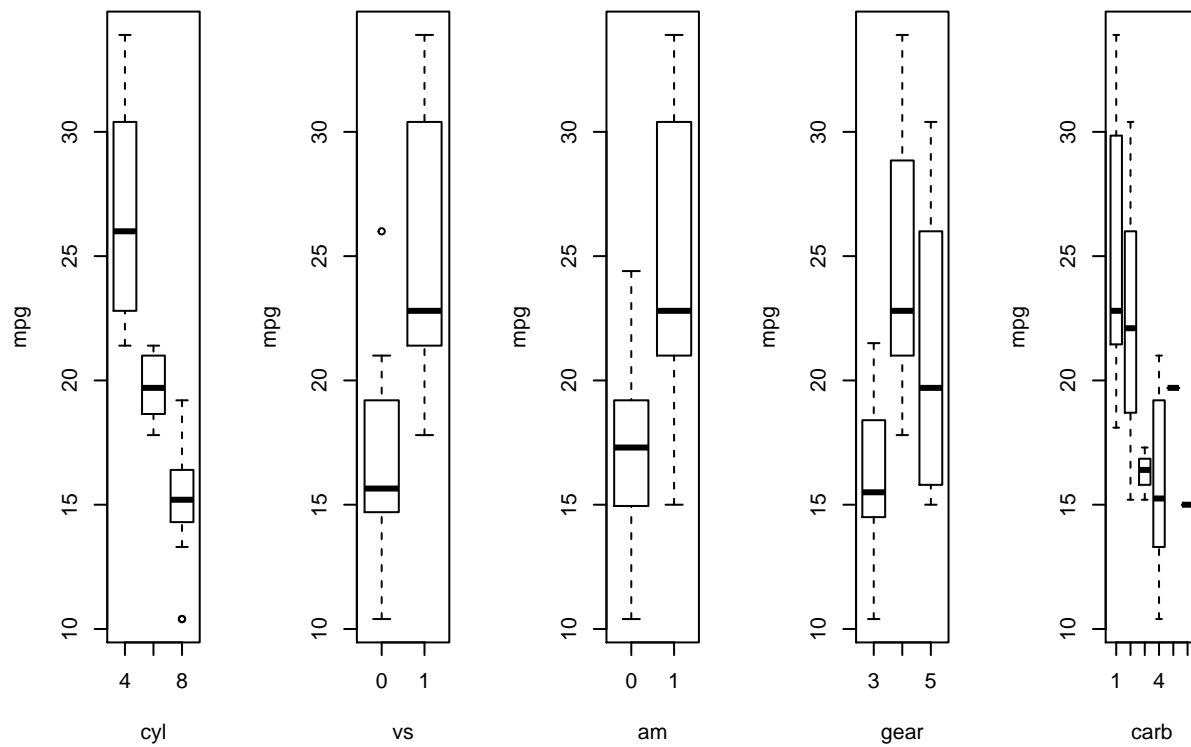


fig 2. mpg vs transmission (automatic vs. manual)

```
mpgAmMtcars <- select(mtcars, mpg, am)
mpgAmMtcars$am <- factor(mpgAmMtcars$am, labels = c("Automatic", "Manual"))
par(mfrow = c(1,1))
boxplot(mpg ~ am, data = mpgAmMtcars, ylab = "mpg", xlab = "am")
```

