Mtcars mile per gallone regression study

WangYong

2022/1/29

- two pages

1. Executive Summary

Based on both data explorary plot mode and linear model coefficients check, it can be confidence to answer below to questions: ## Q1. Is an automatic or manual transmission better for MPG?
Yes. Compared to automatic transmission, manual is better for MPG. This answer will be showed in both boxplot EDA stage and linear mode coefficients interpret stage.

Q2. Quantify the MPG difference between automatic and manual transmissions

With stepAIC strategy to find the best model, we get quantified model : ** mpg = 9.6178 + 2.9358 * am + (-3.9165)* wt + 1.2259 * qsec ** it means: 1. Model can explain for: 84.9663556% variability of mpg 2. Compared to reference mpg(9.6) automatic transmission type, every gallon fuel could run extra 2.9358, that is 2.938+9.6178=12.558. 3. For every 1000lb weight added, the car will be run less -3.9165 miles. 4. the fast the miles will be added 1.2259 mileage per qsec.

2. Exploratory Data Analysis

The mtcars data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption(mpg) and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

2.1 quick look

First, let's quick look the data.

The mtcars dataset has 32 rows. The response variable is 'mpg'. and the independent variables are cyl, disp, hp, drat, wt, qsec, vs, am, gear, carb

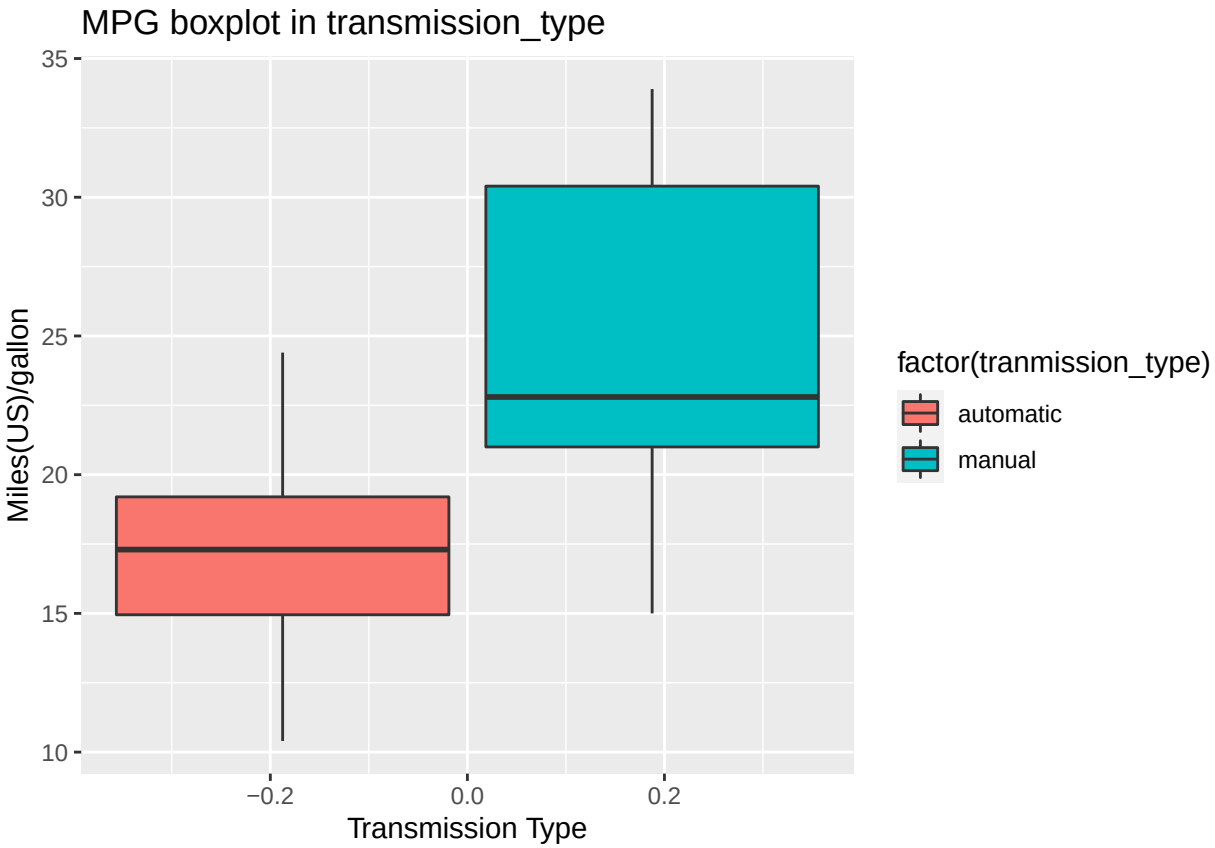
```
##                    mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 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
##                   tranmission_type
## Mazda RX4                   manual
## Mazda RX4 Wag               manual
```

```
## Datsun 710                 manual
## Hornet 4 Drive          automatic
## Hornet Sportabout       automatic
## Valiant                 automatic
```

Then, let use box plot have a overview of mpg compare am transmission.

3. Q1. Is an automatic or manual transmission better for MPG?

This question will be studied by boxplot and two sample t_test to compare. # 3.1 Box plot



For above figures, it can be clear find the different transmission type lead to different mile per gallon. The automatic transmission mean is about 17 miles per gallon. The manual type is a bit higher than automatic, 24 mile per gallon.

3.2 two sample t test

```
library(dplyr)
mpg_auto   <- mtcars %>% filter(am==0) %>% select(mpg)
mpg_manual <- mtcars %>% filter(am==1) %>% select(mpg)

tt_result <-t.test(mpg_manual, mpg_auto, alternative='greater')
```

According to above two sample test, we **reject null hypothesis** with p_value **0.07%** .
It showed that the difference between manual and automatic transmission is significant and manual transmission car MPG is greater than the automatic type.

After compared in by box plot, and student test, It can be confidences that that fuel consumption viewpoint, manual transmission type is better that automatic.

4. Q2. Quantify the MPG difference between automatic and manual transmissions

In order to quantify the mpg difference between two different transmission type. we need look inside,

4.1 Using stepAIC to final appropriate mode

There are 10 independent variable. They are: cyl, disp, hp, drat, wt, qsec, vs, am, gear, carb. The stepAIC will be used to final appropriate model.

```
#aic_result<-  stepAIC(lm(data=mtcars, mpg~.),direction = 'both')
#aic_result$anova
final_model <- lm(data=mtcars, mpg~factor(am)+wt+qsec)
r2 <-summary(final_model)$r.squared
adj_r2<- summary(final_model)$adj.r.squared

residual_normality <-shapiro.test(final_model$residuals)
coef <-summary(final_model)$coef
summary(final_model)
```

```
##
## Call:
## lm(formula = mpg ~ factor(am) + wt + qsec, 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
## factor(am)1   2.9358     1.4109   2.081 0.046716 *
## wt           -3.9165     0.7112  -5.507 6.95e-06 ***
## qsec          1.2259     0.2887   4.247 0.000216 ***
## ---
## 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
```

With stepAIC strategy to find the best model, we get quantified model : ** mpg = 9.6178 + 2.9358 * am + (-3.9165)* wt + 1.2259 * qsec ** it means: 1. Model can explain for: 84.9663556% variability of mpg 2. Compared to reference mpg(9.6) automatic transmission type, every gallon fuel could run extra 2.9358, that is 2.938+9.6178=12.558. 3. For every 1000lb weight added, the car will be run less -3.9165 miles. 4. the fast the miles will be added 1.2259 mileage per qsec.

4.2 Vif & residual normality test

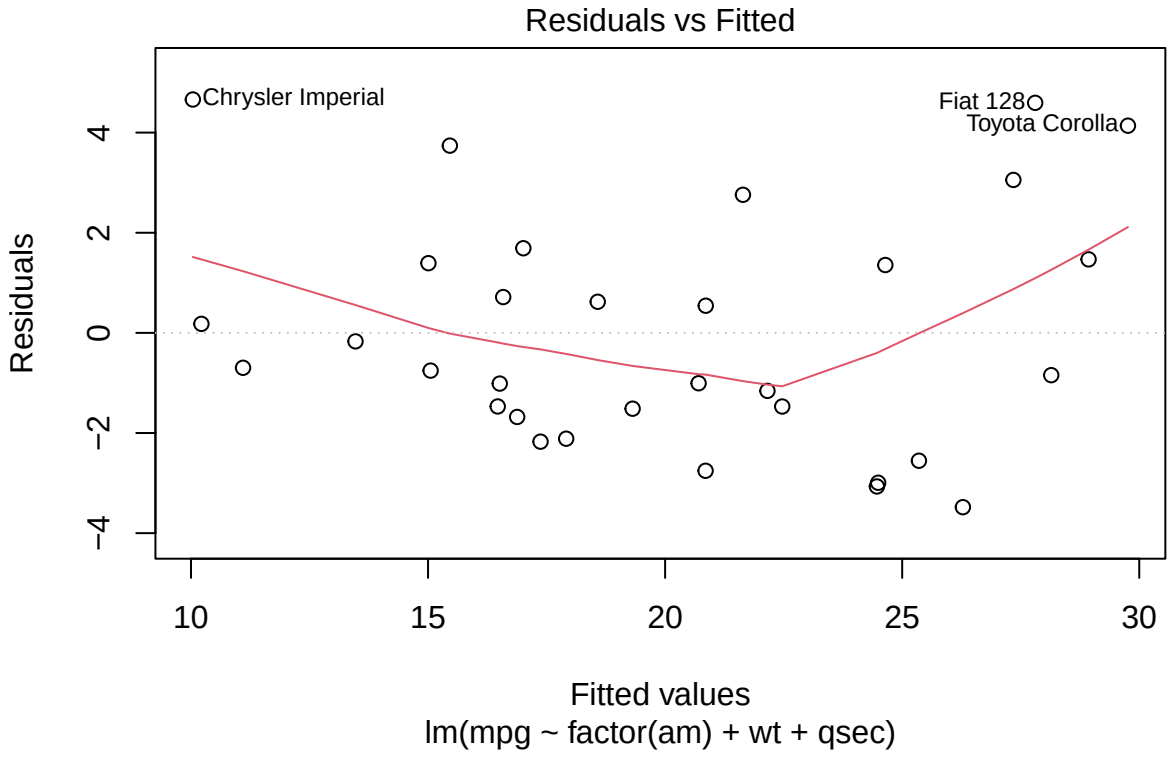
```
## factor(am)         wt       qsec
##   2.541437   2.482952   1.364339
```

1. Multilinear: According to vif result, it showed possibility of multicollinear. If the VIF value less than 4, in general, there is no serious multicollinear issue.

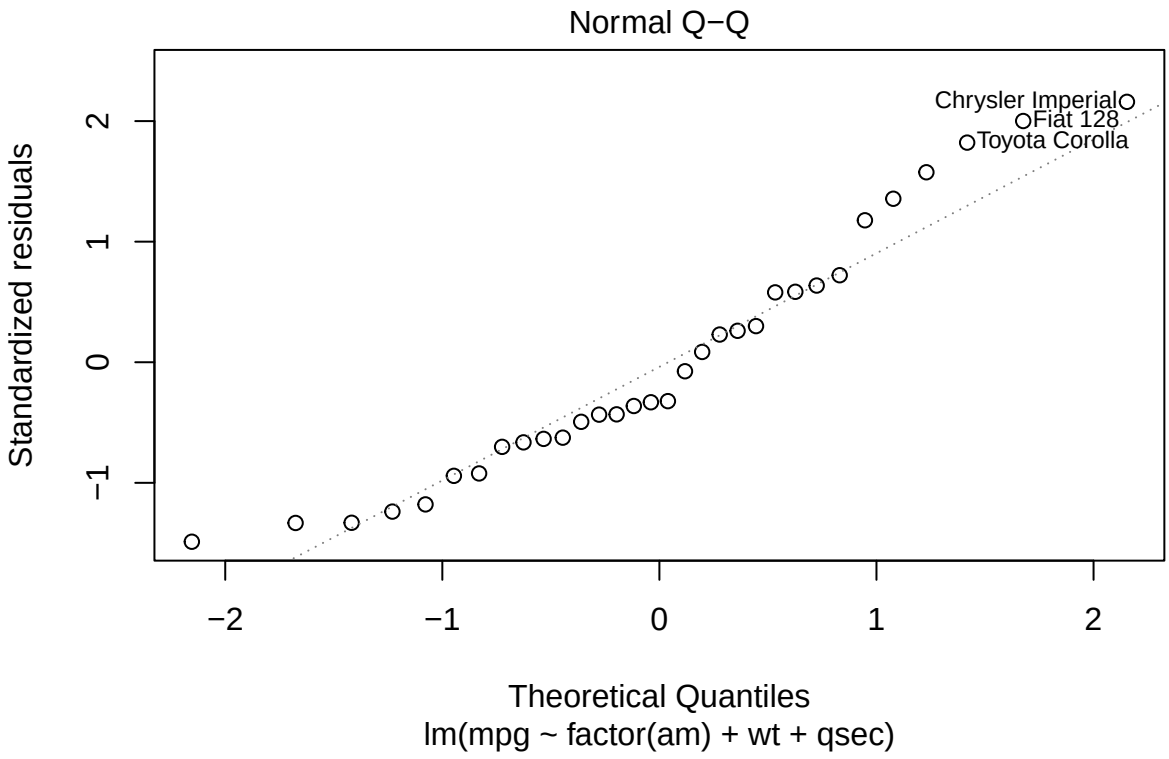
2. residual normality: According the shapiro normaility test, the p_value is 0.0804277 which large that 0.05. So it can keep the null hypothesis, residual normality.

5. Appendix figures

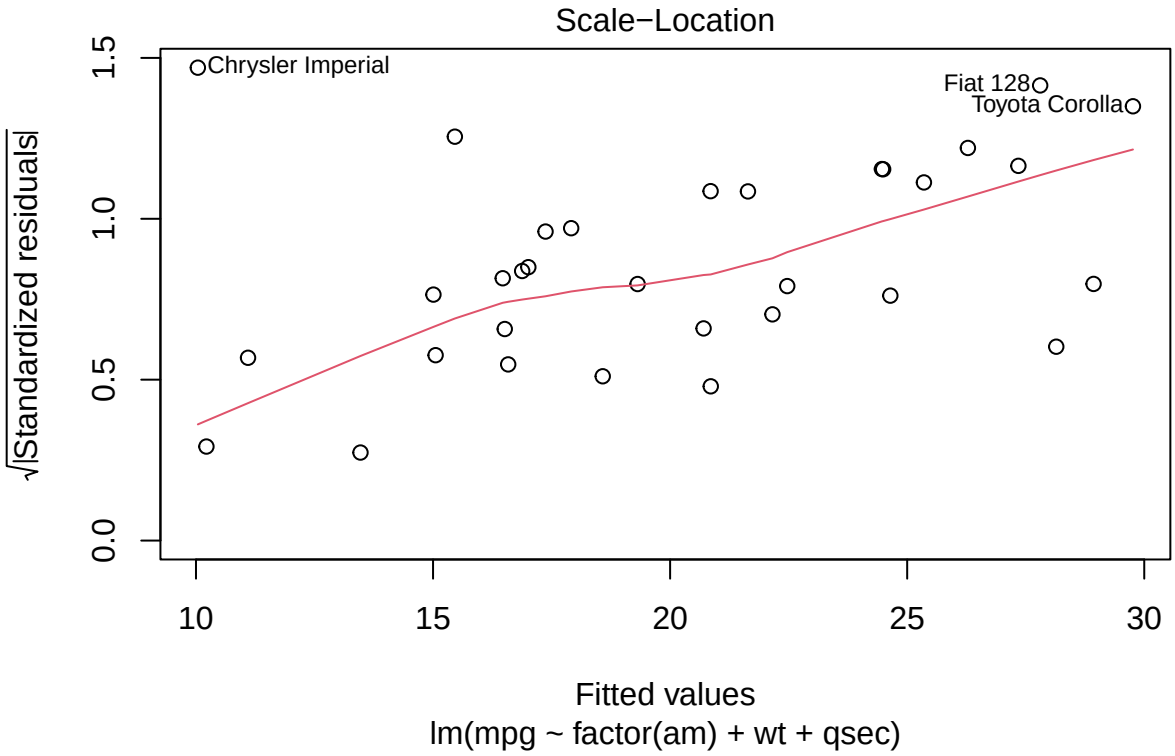
5.1 residual vs fitted plot



5.2 Q-Q plot


Normal Q−Q
Chrysler Imperial
Fiat 128
Toyota Corolla
Standardized residuals
−2
−1
0
1
2
Theoretical Quantiles
lm(mpg ~ factor(am) + wt + qsec)

5.3 scale-location



5.4 cook's distance