Regression Models Project

AI

July 24, 2015

Executive summary of Analysis

Data collected by Motor Trend in 1974 is used to see if

1. Is an automatic or manual transmission better for MPG
2. Quantify the MPG difference between automatic and manual transmissions

Data was visualized for various aspects and after several careful considerations a mathematical model was built to approximate mileage of the car. Based on the analysis among cars with **same weight and horsepower**,

1. We **can not** state that manual transmission or automatic transmission give better mileage than the other. With only **85%** certainity, it can be stated that manual transmission cars have better mielage but the general accepted certainity level to make such a claim is **95%**.
2. Manual transmission cars can give 0.736 *less* miles per gallon to 4.9 *more* miles per gallon compared to auto transmission cars

There seems to be a strong relationship between weight and horsepower of the car than with the type of transmission.

Details of analysis is provided in the following section. Any supporting graphs are provided in corresponding section in appendix.

Exploring *mtcars* data

After basic visualization, it appears that manaul trasmission cars have better mpg than automatic transmission. However, the first graph tells something interesting

1.Manual cars seem to start at lower weight and there are quite a few light weight cars among manual transmission cars that are lower weight than any auto car in data set

2.There a few automatic transmission cars that are heavier than any other manual transmision cars

3.Holding the weight constant, the relationship between transmission type and mpg is not apparent

Model Selection

There are multiple possiblities to mathematically model the mileage of a car. We consider a few in this section and look the merits and pitfalls of each and select one.

Fit 1: mpg vs am

This is the simplest and direct model between mpg and am. This is literally writing the question down as a linear regression problem. This model simply compares the average mileage of automatic transmission and manual tranmission cars.

Since cars in two categories donot necessarily have similar characterstics that can potentially affect the mileage, this can be error prone.

Fit 2: mpg vs am + wt

One of the observations from exploratory analysis is that auto transmission cars weight more. Do they have lesser mileage because of the extra weight? One way to know is by adjusting for weight. This model accounts for the difference in weights across the categories.

Fit 3: mpg vs am + wt + hp

More power a engine has, more fuel it tends to draw. Hence horse power may have an effect on mpg. It is reasonable to consider horse power in the model. This model does exactly that.

Now that we have three models, we performed various diagnositics to pick the best and also considered other models just to ensure none of the significant contributors to mileage is left out from the model.

Diagnositics and Selecting a model

Residual error has gradually decreased by adding the variables (see plots in corresponding section). Also, the predictors weight and horsepower seem to be significant using maximum likelihood analysis. We can reject the hypothesis that adding weight and horsepower is not required. Let us consider adding variables like cylinders and gears.

Other variables such as number of gears and cylinders may have an effect on mileage. For example, more cylinders there are, more fuel is drawn. It is reasonable to consider these models as well. See the code in diagnositcs section.

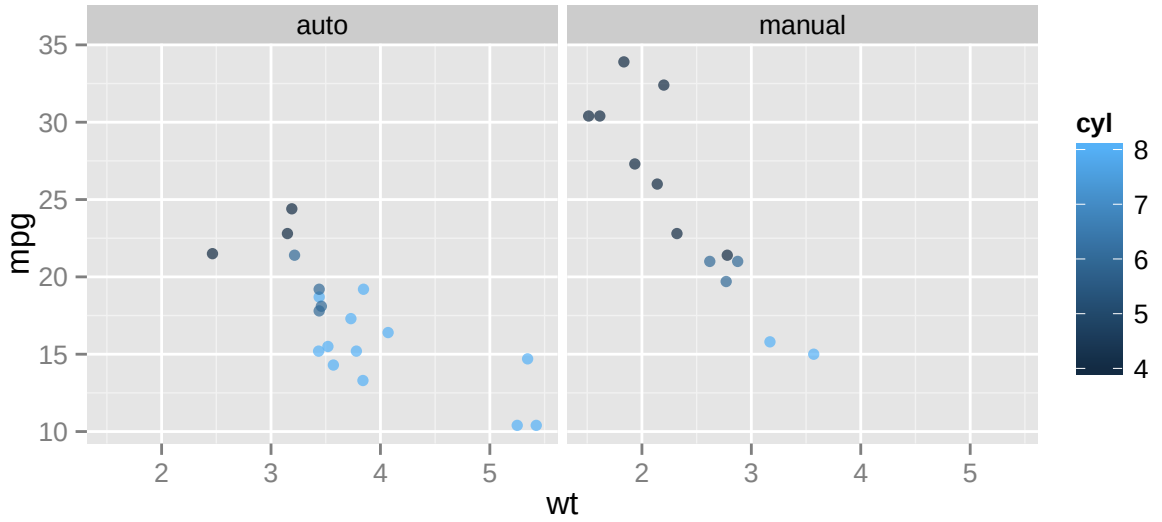
Based on the maximum likelihood analysis, we can reject the hypothesis that cylinders and gears contribute to outcome. Hence our final model does not include these. **We will stick with model 3**. May be the effect of cylinders is captured by the horsepower of the vehicle already in the model.

Conclusion with uncertainity quantified

For *given weight and horse power*, based on the coefficent of transmission type in the model, manual transmission car is expected to give 2.08 more miles per gallon compared to a auto transmission car for this particular data set. Since this is only a small sample of the cars, With 95% confidence we can generalize that for the population of cars a manual transmission car gives 0.736 *less* miles per gallon to 4.9 *more* miles per gallon compared to automatic transmission car.

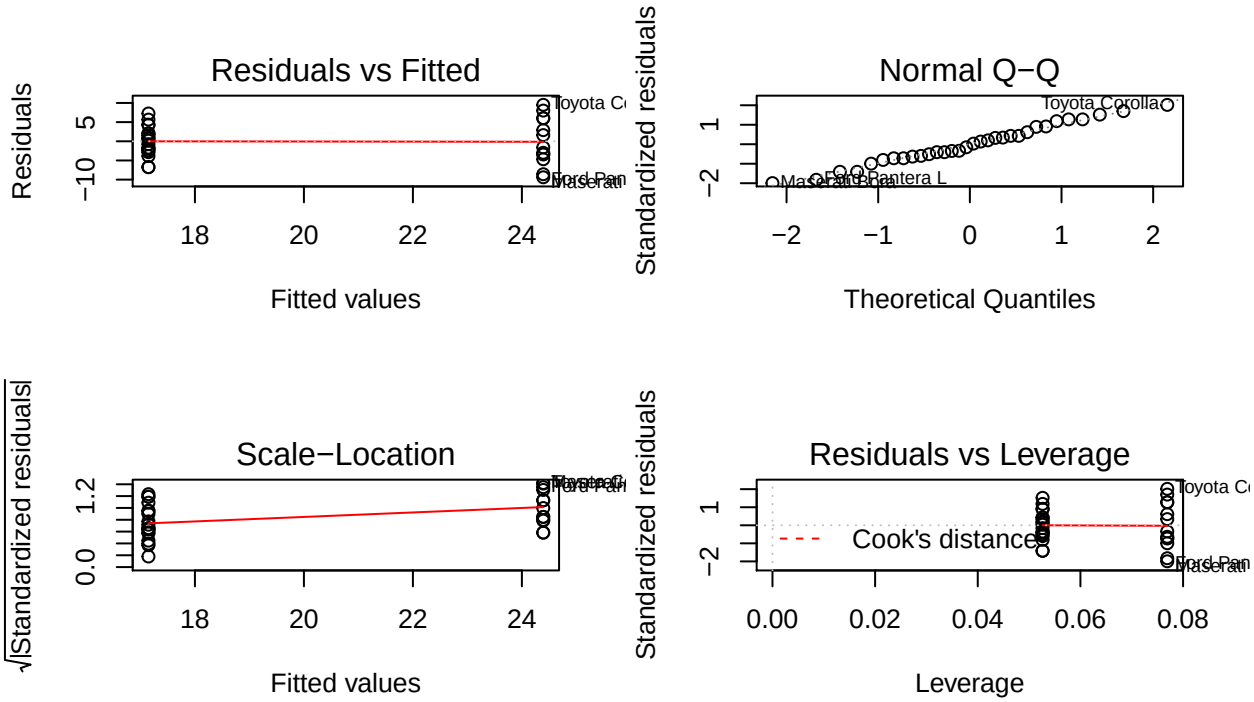
Appendix

Exploring data

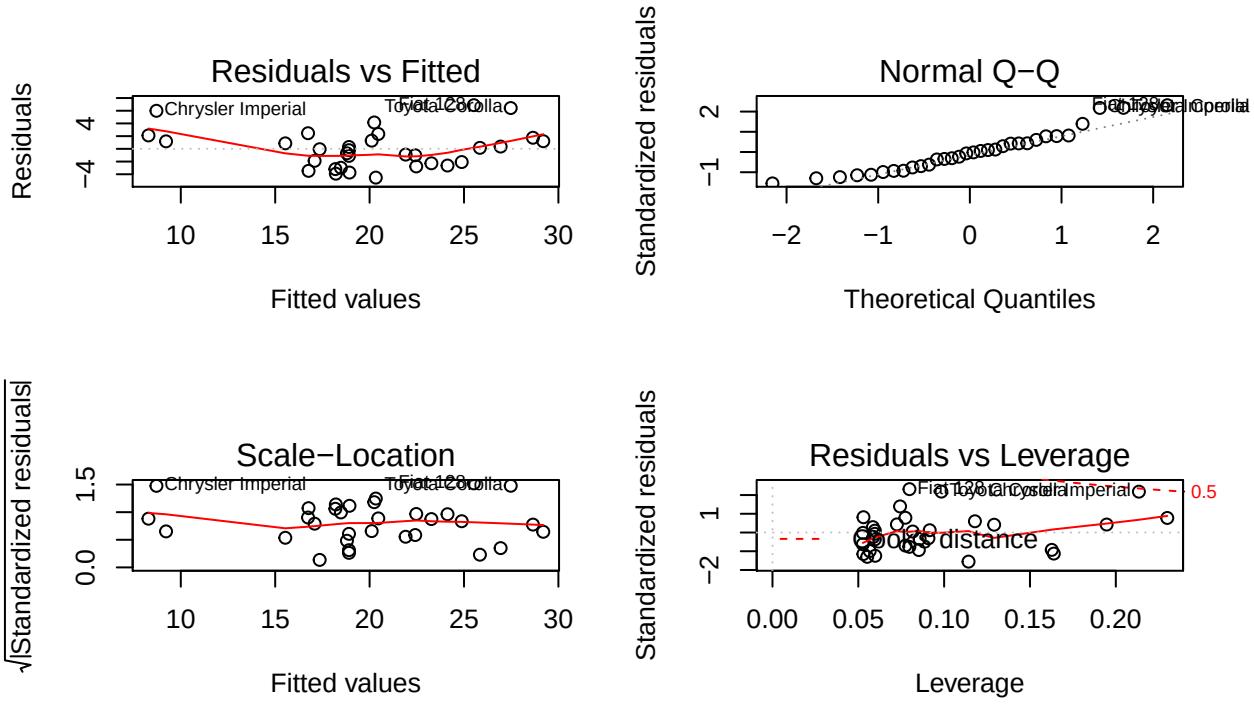


Model Selection

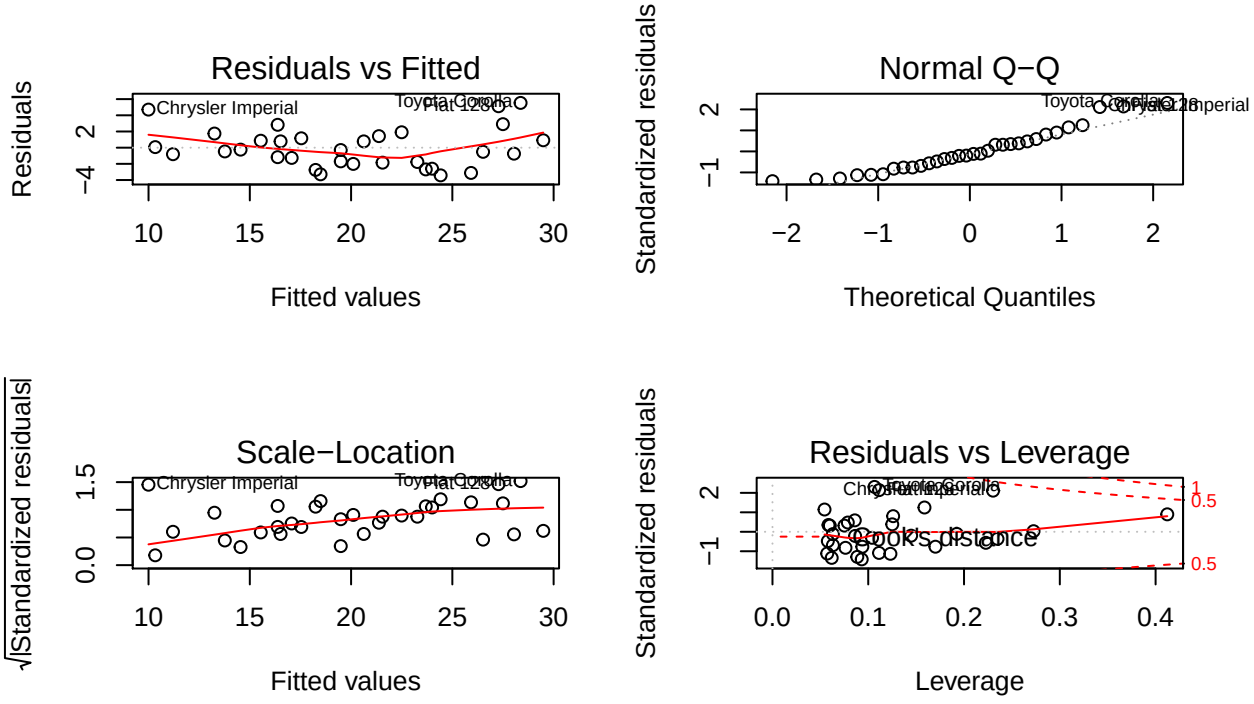
Fit 1



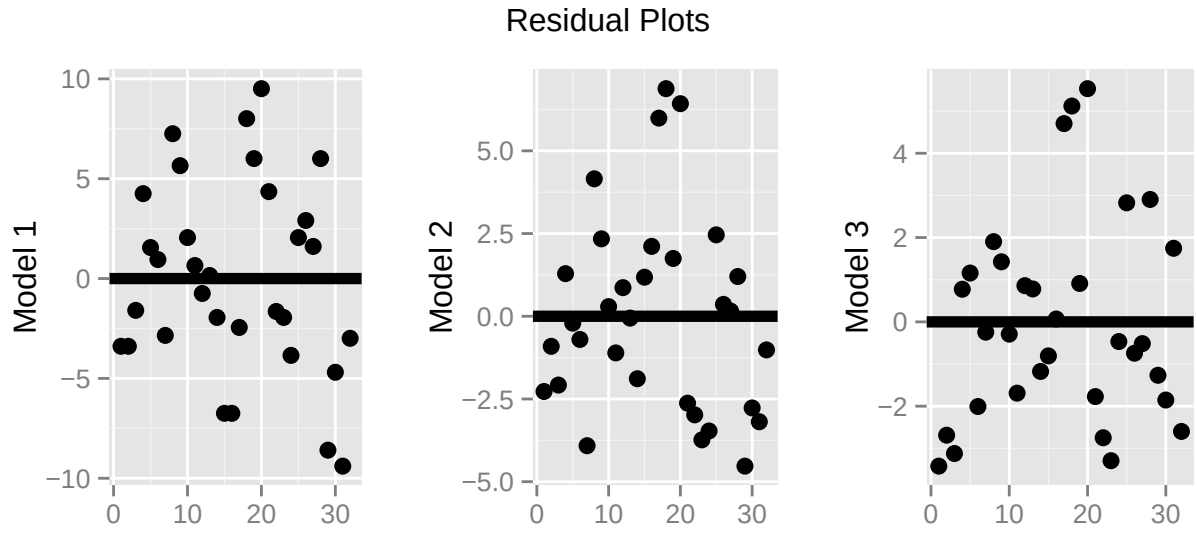
Fit 2


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

Fit 3


Residuals vs Fitted
Residuals
Fitted values
Chrysler Imperial
Toyota Corolla
Fiat 128
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Leverage
Cook's distance
1
0.5
0.5

Diagnositics and Selecting a model




```
## Analysis of Variance Table
##
## Model 1: mpg ~ am - 1
## Model 2: mpg ~ am + wt + hp
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     30 720.90
## 2     28 180.29  2    540.61 41.979 3.745e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am + wt + hp
## Model 2: mpg ~ am + wt + hp + cyl
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     28 180.29
## 2     27 170.00  1    10.293 1.6348 0.2119
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am + wt + hp
## Model 2: mpg ~ am + wt + hp + gear
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     28 180.29
## 2     27 179.34  1   0.95072 0.1431 0.7081
```

Conclusion with uncertainity quantified

```
## [1] -0.7357587  4.9031790
```

```
## [1] 0.04641094 4.12100932
```