The Fuel Consumption of Automatic and Manual Vehicles (1973-74 models)

Jonathan Hill, MPA

Thursday, April 23, 2015

# Overview

The two questions addressed in this project are:

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

I will attempt to answer these questions using a multi-linear regression model with a dependent variable of miles per gallon. The model will include independent variables (control variables) for the characteristics of the engine, such as the number of cylinders and displacement volume, and the car itself, such as rear axle ratio and weight.

# The Data

The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973-74 models). A table with more detailed information about each variable can be found in the appendix.

# Exploratory Analysis

The result of the Shapiro-Wilk normality test suggests it is safe to assume the dependent variable is normal with a mean of 20.09 mpg (as illustrated by the dark black line through the histogram in the appendix).

The exploratory plots suggest that manual transmission vehicles may be more fuel efficient than automatic transmission vehicles because the top row is slightly higher than the bottom row, but they also illustrate that other factors are probably effecting fuel efficiency.

For example, there seems to be an indirect relationship between fuel efficiency and weight. This builds confidence in the quality of our data because it is reasonable to assume that heavier cars are less fuel efficient. It also necessitates the use of a multi-variate method such as multi-linear regression in order to accurately assess how a car's transmission type (automatic/manual) effects its fuel efficiency.

## First Iteration of the Regression Model

reg <- lm(mpg~.,data=mtcars)  
summary(reg)

The first iteration of my regression model included all of the independent variables from the mtcars dataset. Only *wt* seemed to significantly impact *mpg*, but many insignificant variables were causing a lot of noise.

## Final Regression Model

In order to determine which variables to include in the final model, I used an R stepwise regression function, step(), and set the direction=both. This automatic process adds and removes independent variables to the model until it finds the combination of independent variables that maximizes the R2 of the model.

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

# Conclusion

Many characteristics of cars are highly correlated with each other. For example, heavy cars tend to have larger engines and powerful engines tend to have more cylinders (see correlation matrix in the appendix). This results in multi-colinearity issues.

In order to deal with this problem, the 1/4 mile speed of a vehicle is extremely useful. If the car does not have the correct balance of attributes, this medium distance speed will decrease and so to will its fuel efficiency. Therefore, it is an ideal control variable in combination with the vehicle's weight in determining the effect that a car's transmission type has on its fuel efficiency.

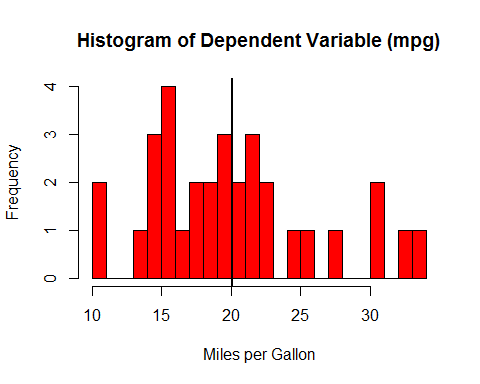
Manual transmission vehicles appeared to be more fuel efficient than automatic transmission vehicles in 1973-74. Based on the mtcars dataset, one could argue that, in 1974, a manual transmission vehicle's mpg was between 0.11 and 5.76 greater than a similar automatic transmission vehicle with a 95% confidence level.

## Appendix

#### Variable Descriptions

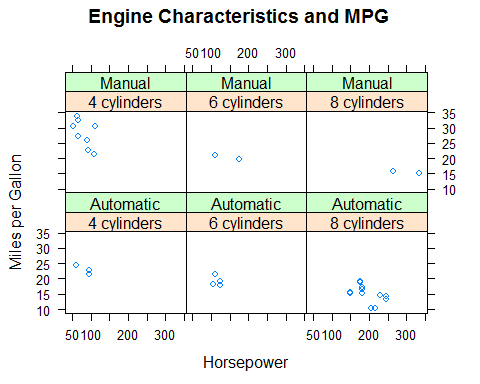
|  |  |
| --- | --- |
| Variable Name | Description |
| *mpg* | Miles/(US) gallon |
| **am** | Transmission (0 = automatic, 1 = manual) |
| cyl | Number of cylinders |
| disp | Displacement (cu.in.) |
| hp | Gross horsepower |
| drat | Rear axle ratio |
| wt | Weight (lb/1000) |
| qsec | 1/4 mile time |
| vs | V/S |
| gear | Number of forward gears |
| carb | Number of carburetors |

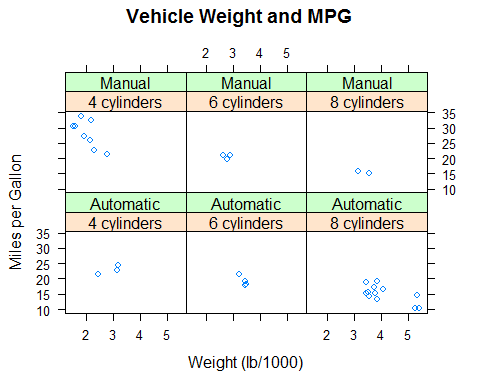
#### Dependent Variable *mpg*



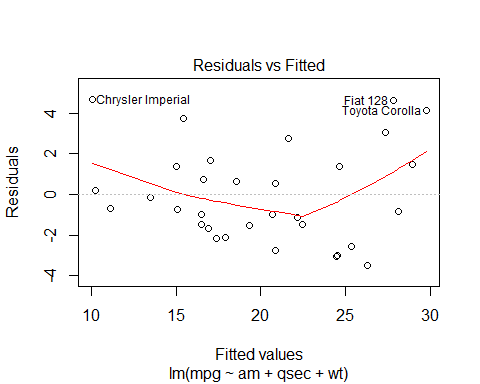
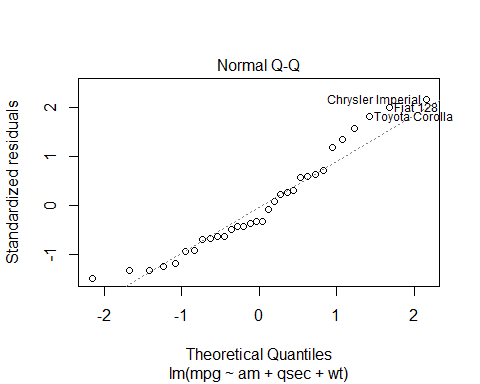
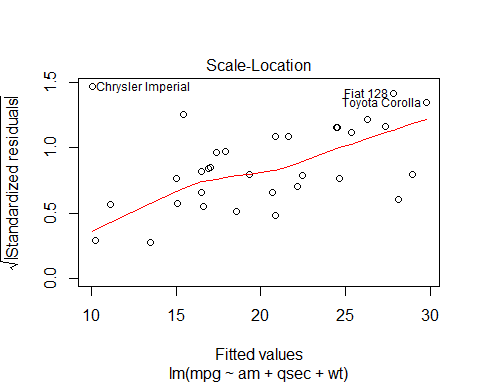
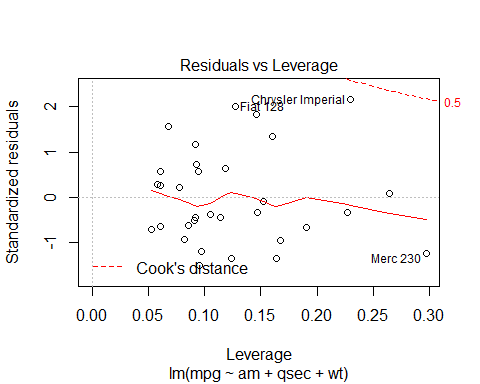
##   
## Shapiro-Wilk normality test  
##   
## data: mtcars$mpg  
## W = 0.9476, p-value = 0.1229

#### Exploratory Plots





#### Regression Diagnostics

#### Summary Statistics

## mpg cyl disp hp   
## Min. :10.40 4 cylinders:11 Min. : 71.1 Min. : 52.0   
## 1st Qu.:15.43 6 cylinders: 7 1st Qu.:120.8 1st Qu.: 96.5   
## Median :19.20 8 cylinders:14 Median :196.3 Median :123.0   
## Mean :20.09 Mean :230.7 Mean :146.7   
## 3rd Qu.:22.80 3rd Qu.:326.0 3rd Qu.:180.0   
## Max. :33.90 Max. :472.0 Max. :335.0   
## drat wt qsec vs   
## Min. :2.760 Min. :1.513 Min. :14.50 Min. :0.0000   
## 1st Qu.:3.080 1st Qu.:2.581 1st Qu.:16.89 1st Qu.:0.0000   
## Median :3.695 Median :3.325 Median :17.71 Median :0.0000   
## Mean :3.597 Mean :3.217 Mean :17.85 Mean :0.4375   
## 3rd Qu.:3.920 3rd Qu.:3.610 3rd Qu.:18.90 3rd Qu.:1.0000   
## Max. :4.930 Max. :5.424 Max. :22.90 Max. :1.0000   
## am gear carb   
## Automatic:19 Min. :3.000 Min. :1.000   
## Manual :13 1st Qu.:3.000 1st Qu.:2.000   
## Median :4.000 Median :2.000   
## Mean :3.688 Mean :2.812   
## 3rd Qu.:4.000 3rd Qu.:4.000   
## Max. :5.000 Max. :8.000

#### Correlation Matrix

## mpg cyl disp hp drat wt  
## mpg 1.0000000 -0.8521620 -0.8475514 -0.7761684 0.68117191 -0.8676594  
## cyl -0.8521620 1.0000000 0.9020329 0.8324475 -0.69993811 0.7824958  
## disp -0.8475514 0.9020329 1.0000000 0.7909486 -0.71021393 0.8879799  
## hp -0.7761684 0.8324475 0.7909486 1.0000000 -0.44875912 0.6587479  
## drat 0.6811719 -0.6999381 -0.7102139 -0.4487591 1.00000000 -0.7124406  
## wt -0.8676594 0.7824958 0.8879799 0.6587479 -0.71244065 1.0000000  
## qsec 0.4186840 -0.5912421 -0.4336979 -0.7082234 0.09120476 -0.1747159  
## vs 0.6640389 -0.8108118 -0.7104159 -0.7230967 0.44027846 -0.5549157  
## am 0.5998324 -0.5226070 -0.5912270 -0.2432043 0.71271113 -0.6924953  
## gear 0.4802848 -0.4926866 -0.5555692 -0.1257043 0.69961013 -0.5832870  
## carb -0.5509251 0.5269883 0.3949769 0.7498125 -0.09078980 0.4276059  
## qsec vs am gear carb  
## mpg 0.41868403 0.6640389 0.59983243 0.4802848 -0.55092507  
## cyl -0.59124207 -0.8108118 -0.52260705 -0.4926866 0.52698829  
## disp -0.43369788 -0.7104159 -0.59122704 -0.5555692 0.39497686  
## hp -0.70822339 -0.7230967 -0.24320426 -0.1257043 0.74981247  
## drat 0.09120476 0.4402785 0.71271113 0.6996101 -0.09078980  
## wt -0.17471588 -0.5549157 -0.69249526 -0.5832870 0.42760594  
## qsec 1.00000000 0.7445354 -0.22986086 -0.2126822 -0.65624923  
## vs 0.74453544 1.0000000 0.16834512 0.2060233 -0.56960714  
## am -0.22986086 0.1683451 1.00000000 0.7940588 0.05753435  
## gear -0.21268223 0.2060233 0.79405876 1.0000000 0.27407284  
## carb -0.65624923 -0.5696071 0.05753435 0.2740728 1.00000000