Regression Models Class Project (JHU via Coursera) - July 2014

Leif Ulstrup

July 22, 2014

Executive Summary

The graphical analysis of the relationship of mpg to automobile transmission type using the 1974 Motor Trend data on 32 cars gives some hints that manual transmissions get better mileage for cars of equivalent horsepower; however the linear regression analysis shows that the combination of weight + horsepower + horsepower*wt explains 89% of the predicted mpg and the introduction of transmission type adds no significant additional factor to the prediction of mpg.

The only way to absolutely confirm whether automatic or manual transmissions are more fuel efficient is to compare the same vehicles outfitted with each option for a side by side comparison. Most auto makers provide this comparison information on the web for their cars.

Introduction

Our project assignment for this class (Regression Models at Johns Hopkins U via Coursera - July 2014) is to assume we work for Motor Trend Magazine and are asked to use data from a 1974 Motor Trend magazine survey of car performance for 32 automobiles to answer two questions. The information about this data source "mtcars" can be found here: http://stat.ethz.ch/R-manual/R-devel/library/datasets/html/mtcars.html

The questions are:

(1) Is an automatic or manual transmission better for MPG (miles per gallon)?

(2) Quantify the MPG difference between automatic and manual transmissions?

The challenge in answering these questions is that the data do NOT include automatic AND manual transmission versions of the same cars for comparison. Solely using the data in the mtcars data set to answer this question by trying to normalize the physical characteristics such as weight, horsepower, acceleration, cylinders, etc and then testing for differences in MPG for automatic vs. manual transmissions will likely have a low confidence factor.

The performance factors in the "mtcars" data set are ultimately governed by physics (Newton's Laws), car designer engineering trade-offs, and how the car was driven in testing. Given that, one can refer to the physics that drive the major factors in MPG performance to think through ways to normalize the physical characteristics in

answering the question. Per this reference: http://www.driverside.com/auto-library/top_10_factors_contributing_to_fuel_economy-317

Four of these major physical factors are:

(1) engine efficiency in converting the energy in a gallon of gasoline into power,
(2) mechanical heat loss due to inefficiencies in the power transmission to the tires,
(3) rolling friction losses (these are function of the weight of the car and coefficient of rolling friction of the tires), and, most importantly at high speeds (>50mph),
(4) drag on the car where the power (energy/unit time) required to overcome

drag = (1/2)*(fluid density of air)*(front surface "drag area" of the car)*(drag coefficient of the car body)* velocity^3

If one assumes that the test driving for the mtcars data approximated the same velocities, then the major difference in energy consumption is based on the front surface of the car * the car's drag coefficient. The drag area and drag coefficients can vary by multiples and create major differences in fuel economy per this reference and these tables: http://en.wikipedia.org/wiki/Automobile_drag_coefficient. We do NOT have this critical drag related data in the mtcars data set.

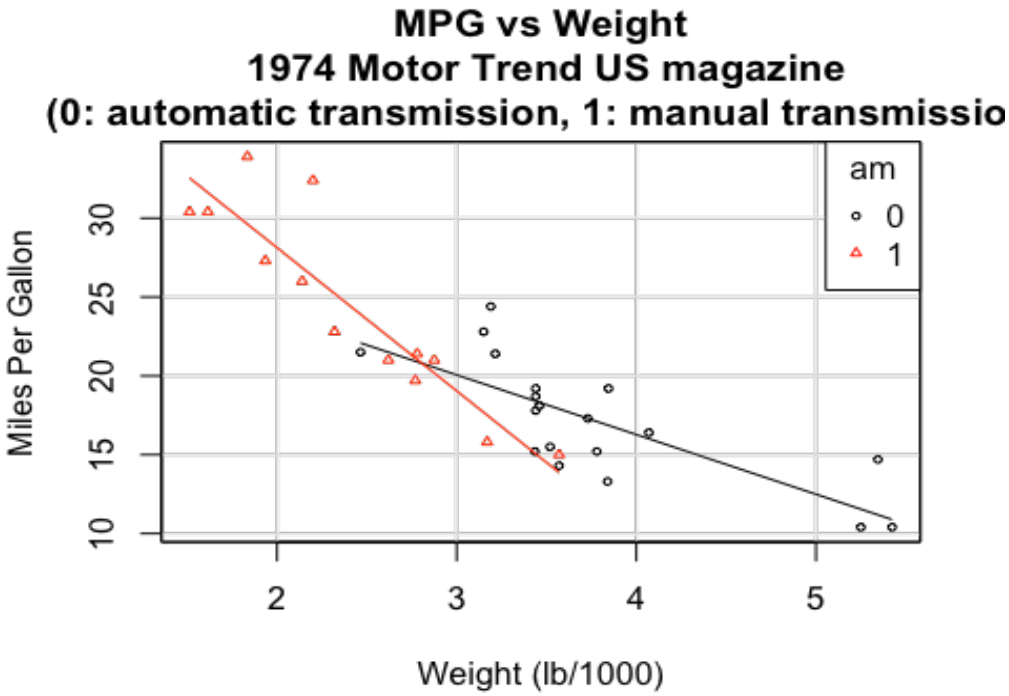
The exploratory data analysis below explores the relationship between MPG and transmission type by first exploring the relationships between MPG and Weight, MPG and Horsepower (HP), Quarter Mile time vs HP, and Quarter Mile time vs Power Ratio (Weight/HP). In each case the data are plotted noting whether the car has an automatic (am=0) or manual transmission (am=1). The analysis shows a number of differences between the automatic and manual car performance but also a number of overlaps in physical parameters where the data show little variation between the two types of transmissions.

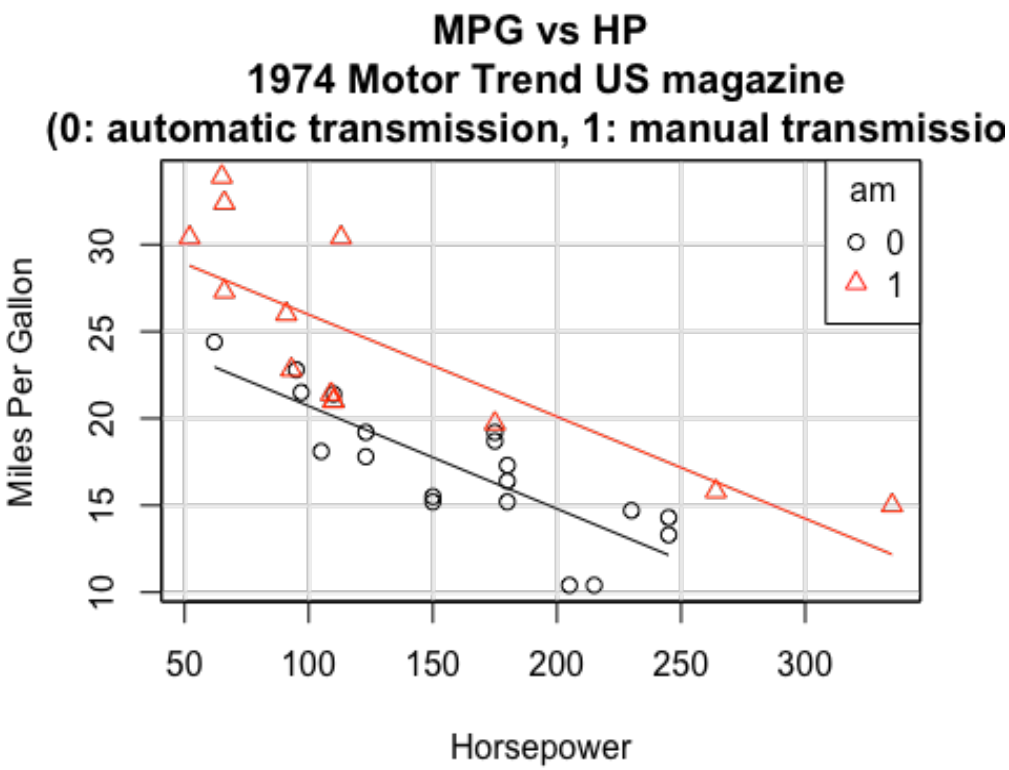
The most striking diagram "MPG vs HP" seems to show that manual transmission cars are on a trend line of ~4 MPG superior performance when horsepower is the assumed factor affecting MPG; however, the data is limited and high horsepower manual transmission cars are relatively low weight cars (sports cars) vs. the automatic transmission set.

A simple linear regression on transmission type (am) as a predictor only explains 36% of the MPG difference (R^2) and has a slope confidence interval that is very broad (from 3.6 to 10.9) while weight (wt) explains 75%, hp explains 60%, and weight + hp + wt * hp explains 89% of the MPG prediction. Adding the transmission type has no effect and has a p-value of .93.

The exporatory data analysis below explains the process used to evaluate mtcars.

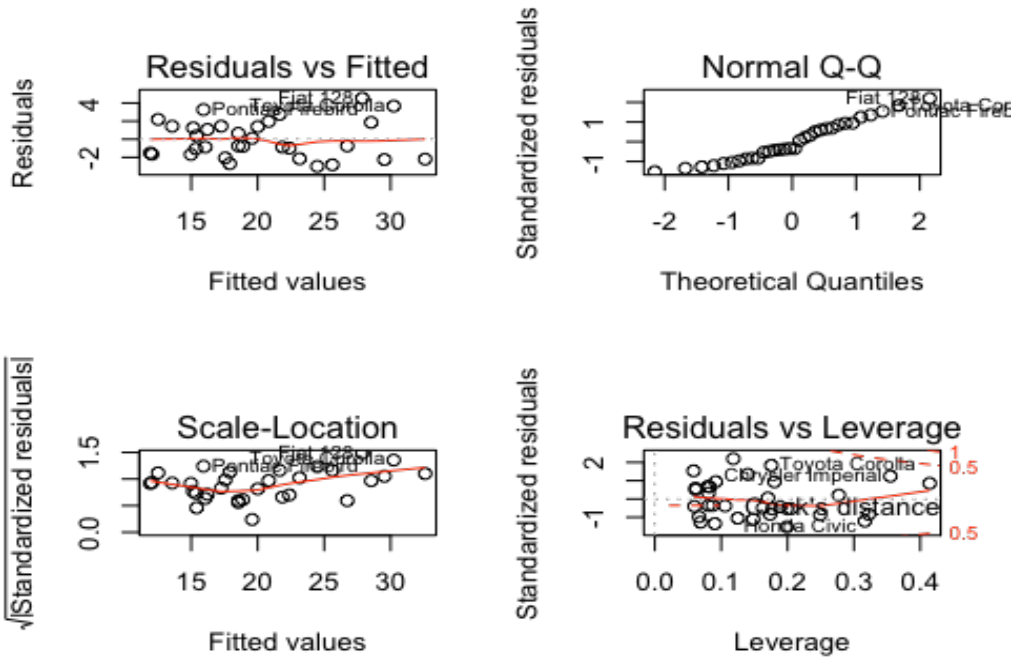
Exploratory Data Analysis (Including Appendix of Figures)

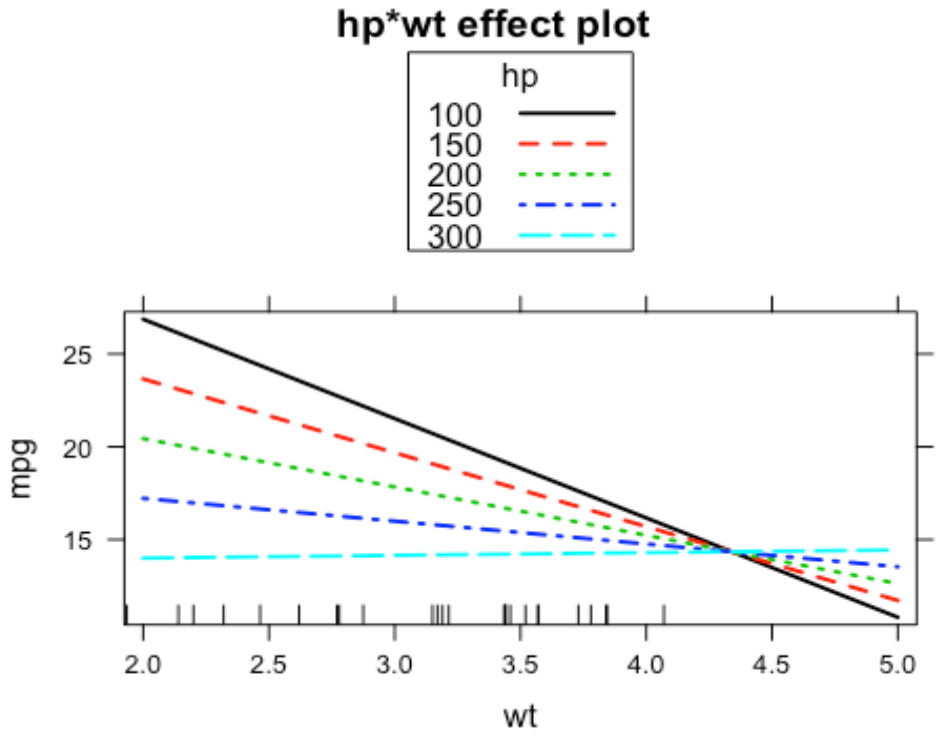




```
##
## Call:
## lm(formula = mpg ~ hp + wt + hp:wt + am, data = mtcars)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -2.985 -1.658 -0.741  1.436  4.527
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 49.45224    5.28073    9.36  5.7e-10 ***
## hp          -0.11930    0.02655   -4.49  0.00012 ***
## wt          -8.10056    1.78933   -4.53  0.00011 ***
## am           0.12511    1.33343    0.09  0.92594
## hp:wt        0.02749    0.00847    3.24  0.00313 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.19 on 27 degrees of freedom
## Multiple R-squared:  0.885,  Adjusted R-squared:  0.868
## F-statistic: 51.8 on 4 and 27 DF,  p-value: 2.76e-12
```

```
##                2.5 %   97.5 %
## (Intercept)  38.6171 60.28741
## hp           -0.1738 -0.06483
## wt          -11.7719 -4.42917
## am           -2.6109  2.86108
## hp:wt         0.0101  0.04487
```





Reference: "R in Action", Robert Kabacoff, pages 186-188