Transmission Type and its Effect on Fuel Efficiency

Russell Allen

Tuesday, September 16, 2014

Executive Summary

This analysis sought to determine if an automatic or manual transmission was better, as measured by MPG. While a cursory examination of the data indicated a manual transmission as favorable to MPG, a deeper analysis revealed that the transmission type alone is insufficient to predict a change in MPG. This analysis concluded that weight is the primary explanatory variable for a vehicles fuel efficiency, and the initial naive observation of transmission type is the result of vehicle weight confounding the initial naive model.

Source Data

Data for this analysis is drawn from the `mtcars` data set, which is available within the R distribution. This data composes 32 observations of 11 variables. Of the 11 variables, we will focus on the transmission type (am), automatic or manual, as an explanatory variable to the vehicle's fuel efficiency expressed in the miles per gallon (mpg) response variable.

Analysis

A simple (naive) analysis of the transmission types relationship to MPG indicates a strong association between the two variables.

```
coef <- summary(lm(mpg ~ am , data = mtcars))$coefficients; coef
```

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.134e-15
## am             7.245      1.764   4.106 2.850e-04
```

This simple regression ignores all other variables and assumes that the transmission type is the only variable affecting the fuel efficiency of the vehicle. Under this naive model, we would clearly reject the null hypothesis (that transmission type has no effect on fuel efficiency) as the P value of 0.03% is well below the traditional 5% threshold.

However, a more detailed analysis which takes all the provided variables into consideration, reveals that the previously observed effect of transmission type is likely the result of confounding from other variables. The following regression model shows the effects of each variable on the fuel efficiency of a vehicle after isolating the effects of the other variables.

```
coef <- summary(lm(mpg ~ . , data = mtcars))$coefficients; coef
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.30337   18.71788  0.6573  0.51812
## cyl         -0.11144    1.04502 -0.1066  0.91609
## disp         0.01334    0.01786  0.7468  0.46349
## hp          -0.02148    0.02177 -0.9868  0.33496
## drat         0.78711    1.63537  0.4813  0.63528
## wt          -3.71530    1.89441 -1.9612  0.06325
```

```
## qsec          0.82104    0.73084  1.1234  0.27394
## vs            0.31776    2.10451  0.1510  0.88142
## am            2.52023    2.05665  1.2254  0.23399
## gear          0.65541    1.49326  0.4389  0.66521
## carb         -0.19942    0.82875 -0.2406  0.81218
```

With the effect of transmission type adjusted to exclude the effects of other variables, it's P value rises to 23.4%, well above 5% which means we fail to reject the null hypothesis. In other words, even though the model indicates that a manual transmission results in a 2.5 increase in MPG, there is a 23.4% chance that the observed affect was random yet normal variance in the underlying data.

In contrast, the weight (wt) of a vehicle appears to have a negative effect on fuel efficiency. Even after factoring out the effects of other variables, a 1000 lb increase in car weight results in a loss of 3.72 MPG … if we are a little loose in our alpha level (P value of weight, 6.3%, is greater than 5%.) This seems like a reasonable assessment, as common logic (and physics) dictate that a greater mass requires a greater fuel expenditure to propel it down the road.

Finally, a regression of MPG with weight and transmission type alone should be considered.

```r
coef <- summary(lm(mpg ~ wt + am , data = mtcars))$coefficients; coef
```

```
##             Estimate Std. Error  t value  Pr(>|t|)
## (Intercept) 37.32155     3.0546 12.21799 5.843e-13
## wt          -5.35281     0.7882 -6.79081 1.867e-07
## am          -0.02362     1.5456 -0.01528 9.879e-01
```

In this model the transmission type fails to affect fuel efficiency in any meaningful way (P value of 98.8%), but the weight of the vehicle is now clearly associated to the fuel efficiency. Under this restricted model, a 1000 lb increase in vehicle weight will result in a loss of 5.4 MPG (P value 1.87×10-5%).

Model Diagnostics

The above analysis assumes a linear relationship between the variables and the MPG. This is not always the case, as can be seen in Appendix A. All variables except the number of cylinders (cyl), weight (wt), cylinder orientation (vs), and transmission type (am) have non-linear relationships.

With regard to transmission type and vehicle weight, an examination of the residuals (See Appendix B) reveals quite a bit of noise in the observed weight's relationship to MPG. However it appears fairly uniformly random, indicating weight's relation is indeed linear to MPG. The residual noise in transmission type though is too large to even consider in the last model analyzed.

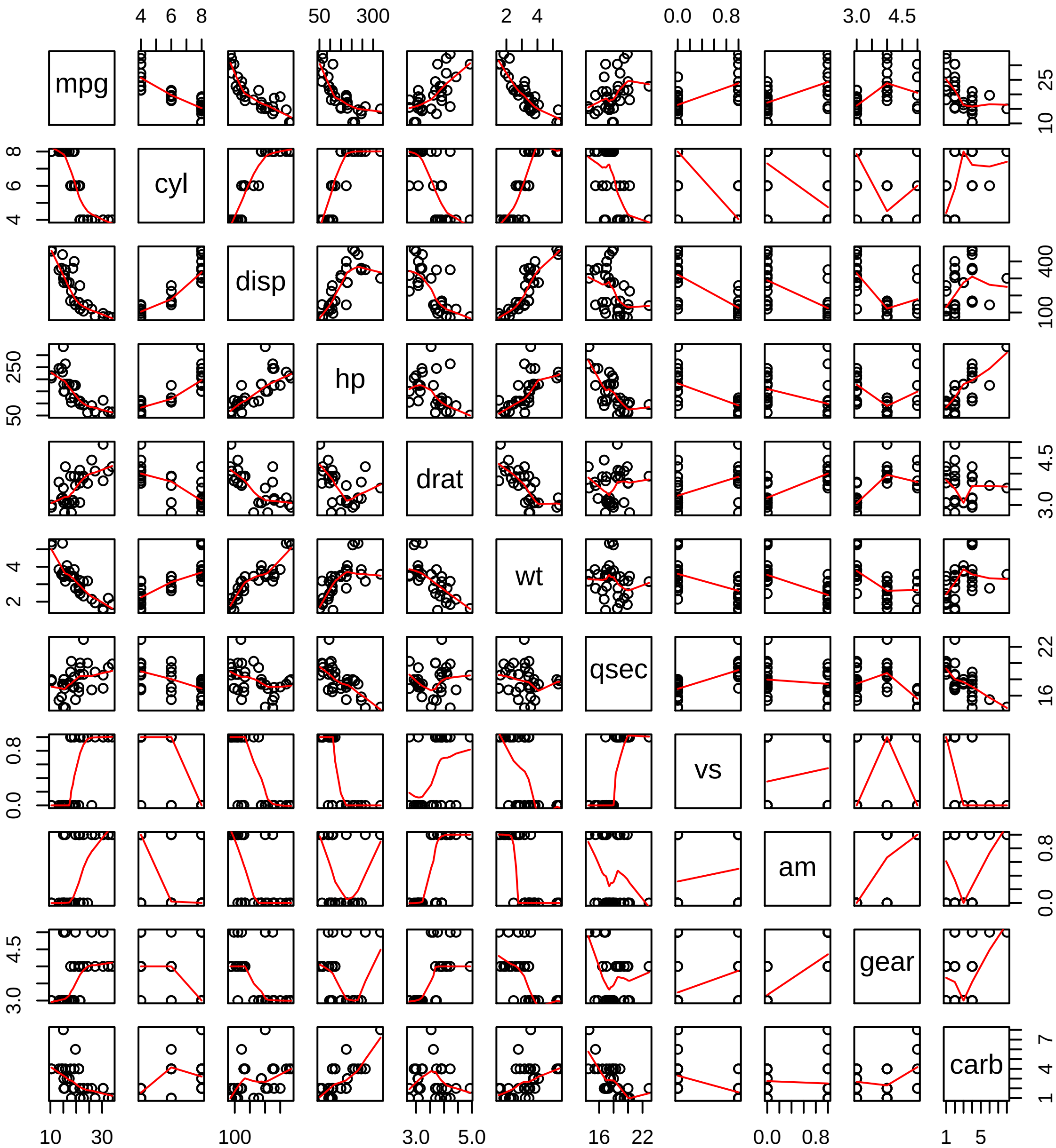
Conclusion

Of the variables measured, weight appears to be the strongest predictor of a vehicles MPG such that each 1000lb increase in weight costs 5.4 in miles per gallon. With a p value of 1.87×10-5%, we rejected the null hypothesis that weight was not related to MPG. We were unable to reject the similar null hypothesis for transmission type, which had a P value of 98.8%. Thus, we can not quantify a difference in MPG for an automatic or manual transmission.

Apendix

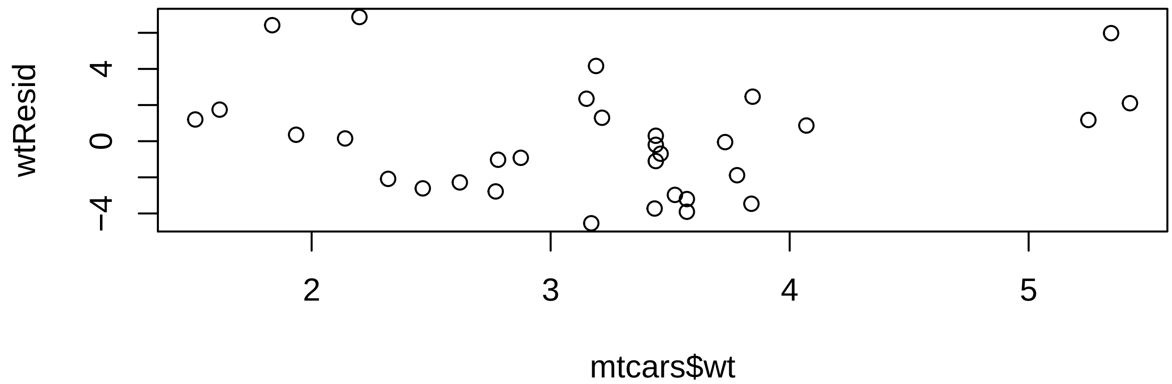
Apendix A

```
pairs(mtcars, panel = panel.smooth)
```



Apendix B - Residuals

```r
plot(wtResid ~ mtcars$wt)
```




```r
plot(amResid ~ mtcars$am)
```