

Impact of Automatic and Manual Transmission on MPG

Liam Damewood

September 21, 2014

Summary

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

The answer to #1 is: Manual, in general, but it depends on the weight of the vehicle. For lighter vehicles (under 2800 lbs.) manual gives better mpg, while automatic gives better mpg for heavier vehicles.

The dataset

The `mtcars` dataset has the following variables:

- [, 1] mpg Miles/(US) gallon
- [, 2] cyl Number of cylinders
- [, 3] disp Displacement (cu.in.)
- [, 4] hp Gross horsepower
- [, 5] drat Rear axle ratio
- [, 6] wt Weight (lb/1000)
- [, 7] qsec 1/4 mile time
- [, 8] vs V/S
- [, 9] am Transmission (0 = automatic, 1 = manual)
- [,10] gear Number of forward gears
- [,11] carb Number of carburetors

```
data(mtcars)
mtcars$vs <- as.factor(mtcars$vs)
mtcars$am[mtcars$am == 0] <- "Automatic"
mtcars$am[mtcars$am == 1] <- "Manual"
mtcars$am <- as.factor(mtcars$am)
mtcars$gear <- as.factor(mtcars$gear)
mtcars$carb <- as.factor(mtcars$carb)
```

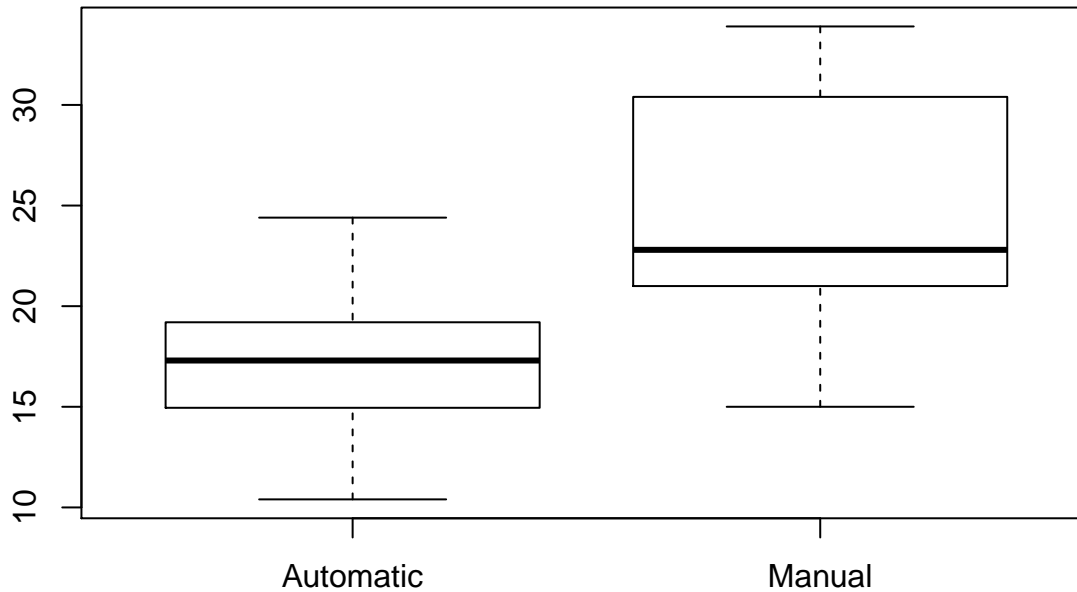
Comparing mpg for the two transmission types

By comparing the mean mpg for the two types of transmission, it is clear that manual transmission is better for mpg overall, so the answer for question #1 is “Manual”. Below, I plot the bar distribution of the automatic and manual transmissions showing that the mean mpg for manual transmission is 24.3923 and the mean mpg for automatic transmission is 17.1474. The model for the transmission type alone is

$$mpg_i = \beta_1 + \beta_2 am_i + \epsilon_i$$

where $am_i = 1$ for manual transmission and $am_i = 0$ for automatic. This model simply compares the average mpg for the automatic (β_1) to the average mpg for manual transmission ($\beta_1 + \beta_2$)

```
plot(mtcars$am, mtcars$mpg)
```



```
fit1 <- lm(mpg ~ am, mtcars)
summary(fit1)$coef
```

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

Weight

This picture does not account for factors such as weight, horsepower or number of cylinders. First, I consider the interaction of the weight with the transmission type. The linear model is

$$mpg_i = \beta_1 + \beta_2 am_i + \beta_3 wt_i + \beta_4 am_i wt_i + \epsilon_i.$$

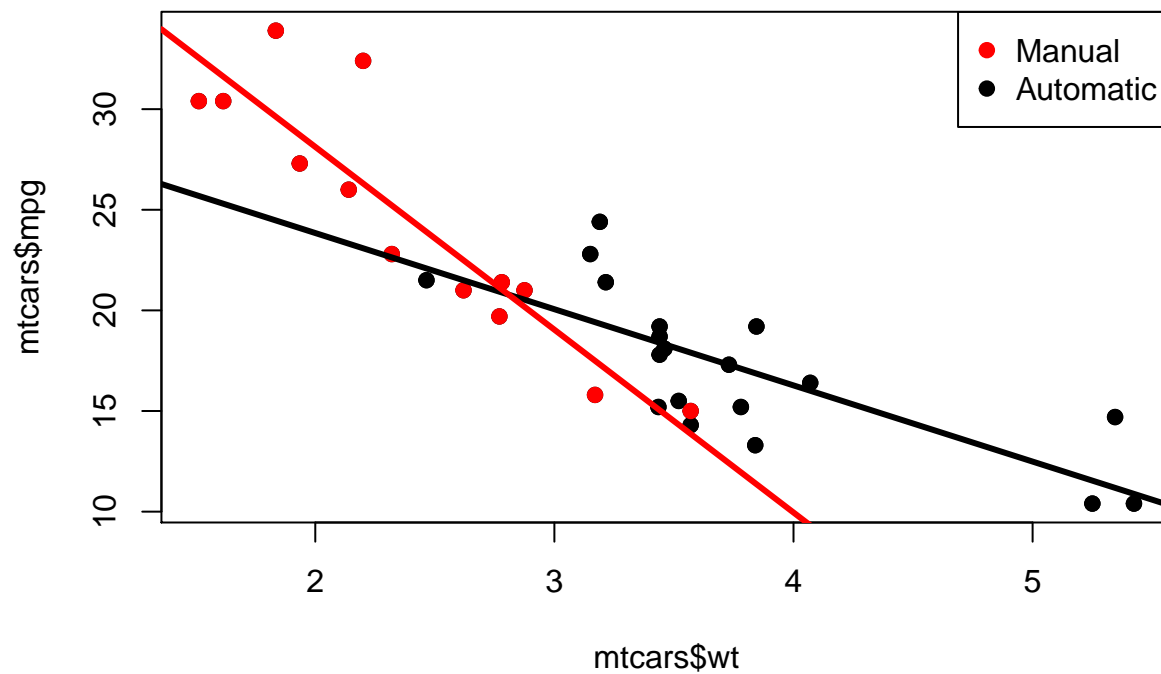
The best fit lines are

$$E[mpg] = \hat{\beta}_1 + \hat{\beta}_3 wt$$

$$E[mpg] = (\hat{\beta}_1 + \hat{\beta}_2) + (\hat{\beta}_3 + \hat{\beta}_4) wt$$

for automatic and manual transmissions, respectively. The data and best fit lines are shown below. Next the residuals

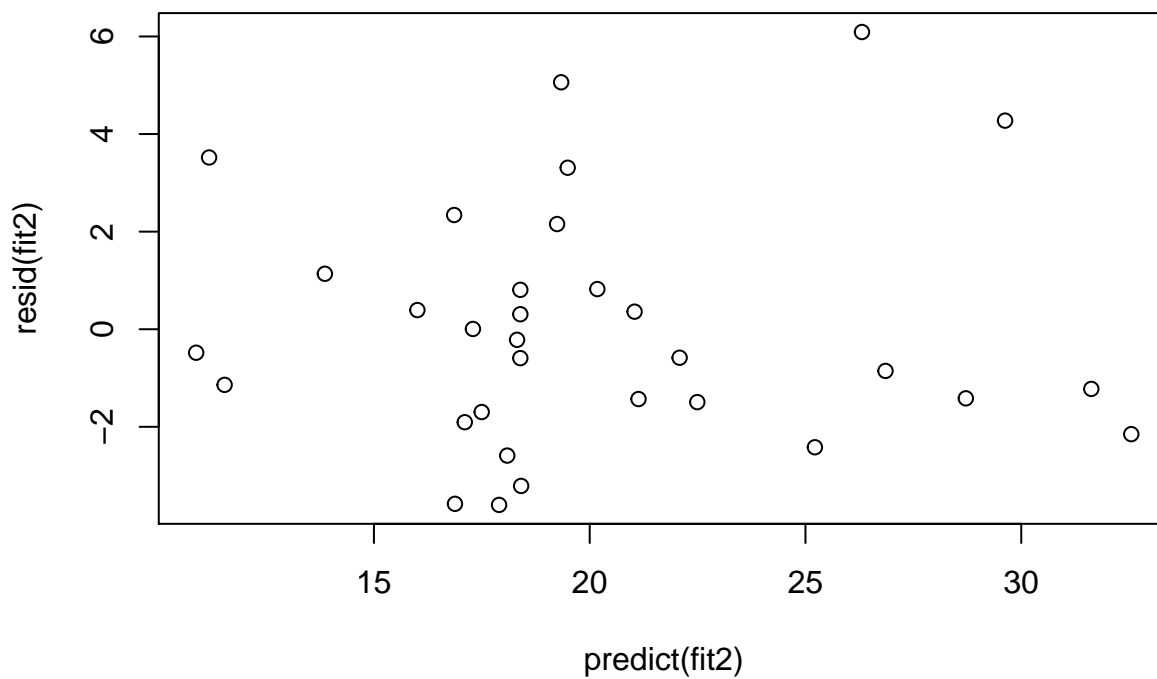
```
fit2 <- lm(mpg ~ am * wt, mtcars)
plot(mtcars$wt, mtcars$mpg, pch=19)
points(mtcars$wt, mtcars$mpg, pch=19, col=((mtcars$am=="Manual")*1+1))
abline(c(fit2$coeff[1], fit2$coeff[3]), col="black", lwd=3)
abline(c(fit2$coeff[1] + fit2$coeff[2], fit2$coeff[3] + fit2$coeff[4]), col="red", lwd=3)
legend("topright", c("Manual", "Automatic"), pch = 19, col = c(2,1))
```



```
summary(fit2)$coef
```

	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	31.416	3.0201	10.402	4.001e-11
## amManual	14.878	4.2640	3.489	1.621e-03
## wt	-3.786	0.7856	-4.819	4.551e-05
## amManual:wt	-5.298	1.4447	-3.667	1.017e-03

```
plot(predict(fit2), resid(fit2))
```



Clearly, the weight plays a significant factor in determining the mpg. Below around 3000 pounds, the manual transmission is better for mpg while it is better for automatic transmission above. The precise weight is the intersection of the best fit lines at $wt = -\hat{\beta}_2/\hat{\beta}_4 = 2.8081 \times 1000$ lbs.

Horsepower

Another factor that may effect the mpg is the horsepower. The new model is

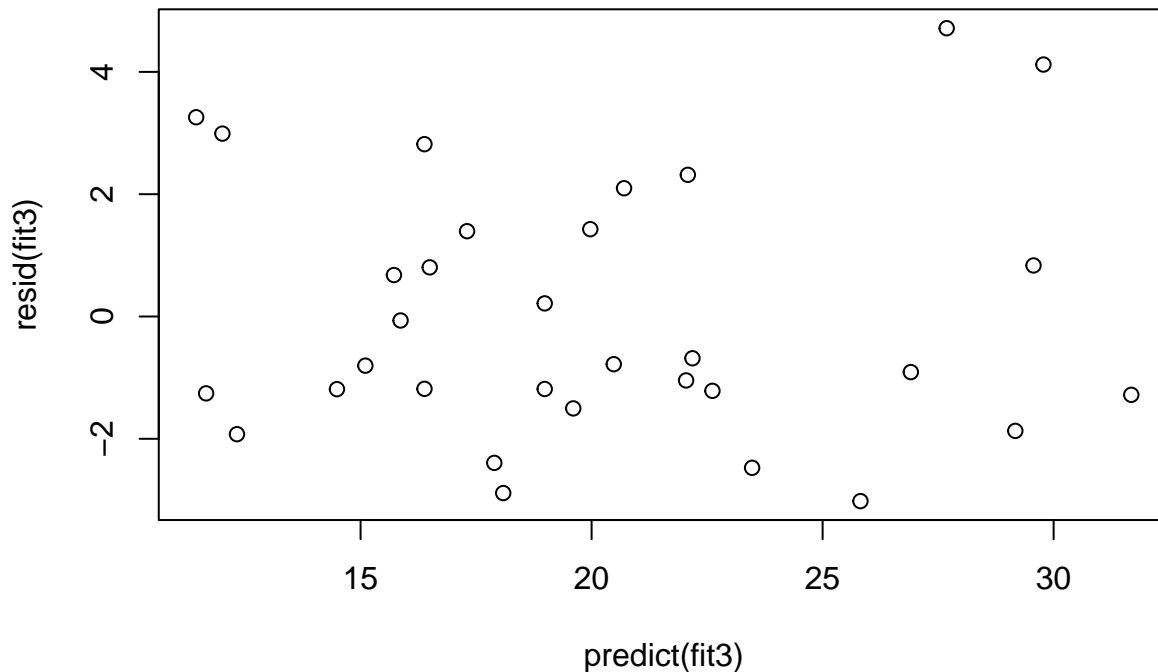
$$mpg_i = \beta_1 + \beta_2 am_i + \beta_3 wt_i + \beta_4 hp_i^\alpha + \beta_5 am_i wt_i + \epsilon_i$$

where I have used $\alpha = 1/2$.

```
alpha = 1./2
fit3 <- lm(mpg ~ am * wt + I(hp^alpha), mtcars)
summary(fit3)$coef
```

```
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)   35.541     2.8762  12.357 1.265e-12
## amManual      10.857     3.8568   2.815 8.998e-03
## wt           -2.279     0.8135  -2.801 9.302e-03
## I(hp^alpha)   -0.786     0.2371  -3.314 2.623e-03
## amManual:wt   -3.324     1.3759  -2.416 2.272e-02
```

```
plot(predict(fit3), resid(fit3))
```



To compare the models, I use the nested model testing, which shows that the horsepower is significant ($P < 0.05$).

```
anova(fit1, fit2, fit3)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am * wt
## Model 3: mpg ~ am * wt + I(hp^alpha)
##   Res.Df RSS Df Sum of Sq    F Pr(>F)
## 1      30 721
## 2      28 188  2      533 53.8 3.8e-10 ***
## 3      27 134  1      54 11.0  0.0026 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```