

Is an automatic or manual transmission better for MPG

Manual transmission in cars yeild about 3.33 more miles per gallon than their automatic counterparts.

A plot(Figure.1 in appendix) between mileage and transmission mode, shows that manual transmission in geneeral yeilds better mileage.

But to quantify the actual difference requires more than just a simple plot. Figure 2 in appendix shows a clear linear relationship between MPG & cyl, disp, hp, wt. Variance inflation factor(VIF) can give a picture regarding multicollinearity.

```
## Warning: package 'car' was built under R version 3.2.2

##      cyl      disp      hp      drat      wt      qsec      vs
## 15.373833 21.620241  9.832037  3.374620 15.164887  7.527958  4.965873
##      am      gear      carb
##  4.648487  5.357452  7.908747
```

The infuence of including 'cyl', 'disp', 'hp', 'wt' variables can be tested through nested likelyhood ratio test.

```
fit1 <- lm(mpg ~ factor(cyl) + factor(am), data = mtcars)
fit2 <- lm(mpg ~ factor(cyl) + factor(am) + disp, data = mtcars)
fit3 <- lm(mpg ~ factor(cyl) + factor(am) + disp + hp, data = mtcars)
fit4 <- lm(mpg ~ factor(cyl) + factor(am) + disp + hp + wt, data = mtcars)
anova(fit1, fit2, fit3, fit4)

## Analysis of Variance Table
##
## Model 1: mpg ~ factor(cyl) + factor(am)
## Model 2: mpg ~ factor(cyl) + factor(am) + disp
## Model 3: mpg ~ factor(cyl) + factor(am) + disp + hp
## Model 4: mpg ~ factor(cyl) + factor(am) + disp + hp + wt
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      28 264.50
## 2      27 230.46  1    34.036 5.6572 0.025339 *
## 3      26 183.04  1    47.421 7.8820 0.009541 **
## 4      25 150.41  1    32.630 5.4236 0.028246 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p values from anova indicate that fit4 is better amongst the above proposed ones.

cyl and disp have high correlation

Although no. of cylinders and displacement have high correlation, anova suggests that including both the variables gives a better model.

```
cor(mtcars$cyl, mtcars$disp)
```

```
## [1] 0.9020329

fit4 <- lm(mpg ~ factor(cyl) + factor(am) + disp + hp + wt, data = mtcars)
fit6 <- lm(mpg ~ factor(cyl) + factor(am) + hp + wt, data = mtcars)
fit7 <- lm(mpg ~ factor(am) + disp + hp + wt, data = mtcars)

anova(fit4, fit6)$Pr
## [1] NA 0.751489

anova(fit4, fit7)$Pr
## [1] NA 0.1066097
```

disp and weight also have high correlation

But, in this case omitting the 'wt' instead of 'disp' variable gives a better model.

```
cor(mtcars$wt, mtcars$disp)
## [1] 0.8879799

fit4 <- lm(mpg ~ factor(cyl) + factor(am) + disp + hp + wt, data = mtcars)
fit5 <- lm(mpg ~ factor(cyl) + factor(am) + disp + hp, data = mtcars)
fit8 <- lm(mpg ~ factor(cyl) + factor(am) + hp + wt, data = mtcars)

anova(fit4, fit5)$Pr
## [1] NA 0.02824553

anova(fit4, fit8)$Pr
## [1] NA 0.751489
```

p value from Analysis of variance(Anova) suggests that 'fit5' is better amongst the above models.

Difference between Manual & Automatic transmission

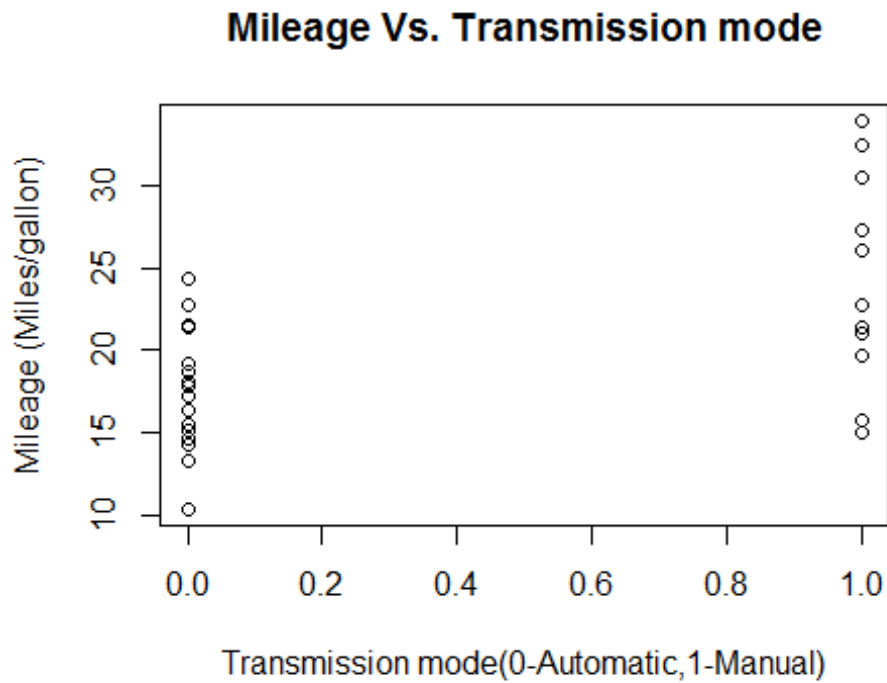
##	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	29.00423340	1.84536783	15.7173182	8.594568e-15
## factor(cyl)6	-3.22228303	1.58865398	-2.0283102	5.288842e-02
## factor(cyl)8	-1.01140831	3.03316951	-0.3334493	7.414694e-01
## factor(am)1	3.33382838	1.36363806	2.4448044	2.158019e-02
## disp	-0.01501139	0.01058469	-1.4182170	1.680029e-01
## hp	-0.03856603	0.01485958	-2.5953650	1.533180e-02

The coefficients of fit5 suggests that, keeping all other parameters constant, cars with 'am(1)' i.e manual transmission on an average yeild 3.33 miles more per gallon of fuel compared to automatic transmission.

*In actuality, it may vary from 1.97 miles/gallon to 4.69 miles/gallon. (Since standard error = 1.36miles/gallon)

Appendix

Figure 1



The plot below shows a clear linear relationship between MPG and cyl, disp, hp, wt

Figure 2

