Course Project

Daniel Fabian

June 20, 2015

Summary

The `mtcars` data set shows was analysed using a linear regression model, fitted with two variables: weight and automatic vs. manual transmission. However, no significant difference in `mpg` can be linked to automatic vs. manual transmission. Once weight of the car has been accounted for, the transmission has near-zero influence in the `mtcars` dataset.

Analysis

We start off with the `mtcars` data set, but we make two factor variables actually factors, first:

```
carsData <- mtcars
carsData$am <- as.factor(carsData$am)
levels(carsData$am) <- c("Automatic", "Manual")

carsData$vs <- as.factor(carsData$vs)
levels(carsData$vs) <- c("V", "S")
```

We start off by looking at the different pairwise plots, see fig 1.

Since automatic vs manual gear is a factor variable and not just a continuous one, we cannot directly use it for predicting the `mpg`. On the other hand, a very significant part of `mpg` can be explained with its weight, see fig 2.

we can start off by fitting a first model where we take care of the weight of the car:

```
fit1 <- lm(mpg ~ wt, data = carsData)
```

now when we took care of the weight as the main predictor, we can try and fit a linear model, where the gear is incorporated:

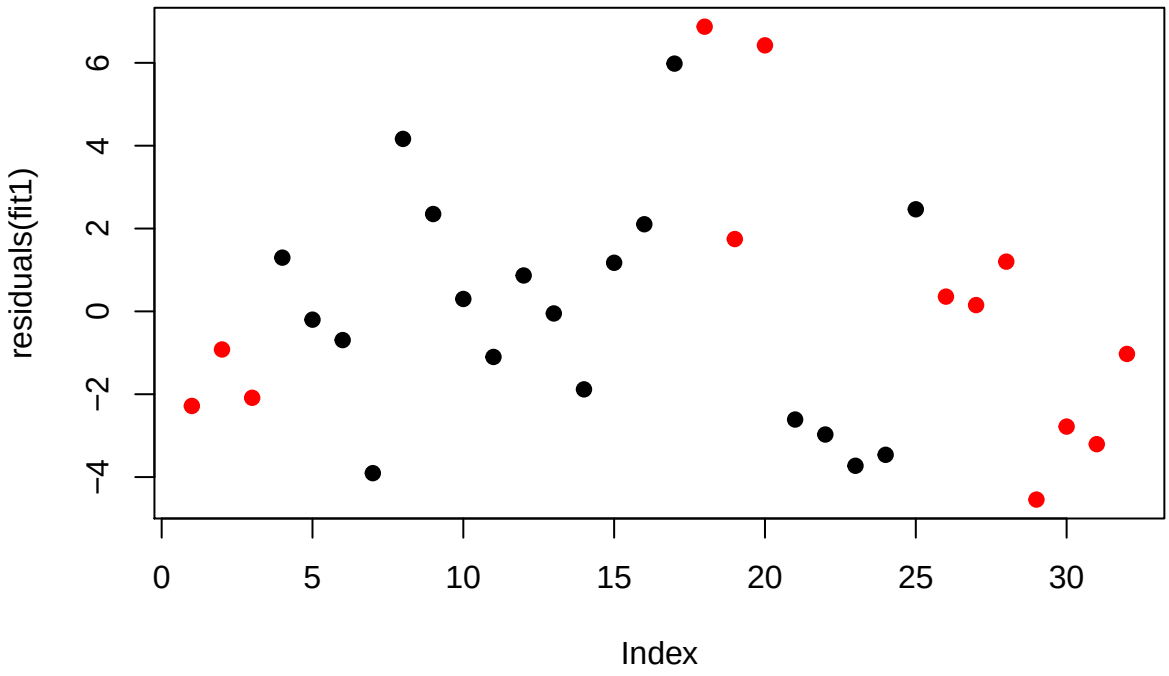
```
summary(update(fit1, mpg ~ wt + am))
```

```
##
## Call:
## lm(formula = mpg ~ wt + am, data = carsData)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.5295 -2.3619 -0.1317  1.4025  6.8782
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 37.32155    3.05464  12.218 5.84e-13 ***
```

```
## wt          -5.35281    0.78824  -6.791 1.87e-07 ***
## amManual    -0.02362    1.54565  -0.015    0.988
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.098 on 29 degrees of freedom
## Multiple R-squared:  0.7528, Adjusted R-squared:  0.7358
## F-statistic: 44.17 on 2 and 29 DF,  p-value: 1.579e-09
```

We see, that the t-value for the automatic vs. manual gear is very low, so it seems to be very much not significant.

```
plot(residuals(fit1), col=carsData$am, pch=19)
```



we see the same in the residuals plot.

So this data, does not support the evidence that any of automatic vs manual transmission is better for mpg. As such the difference cannot be quantified.

Appendix

fig 1

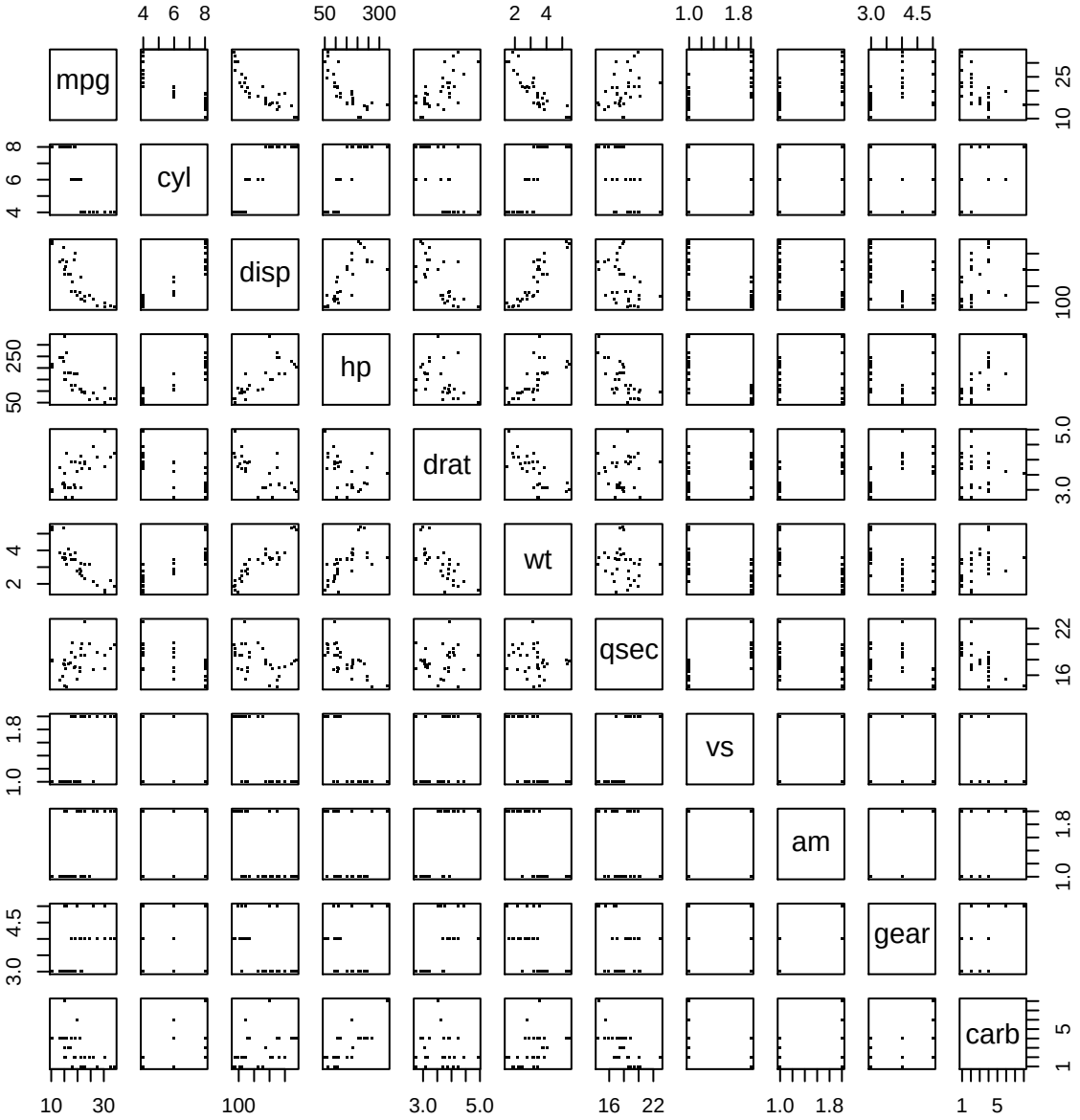


fig 2

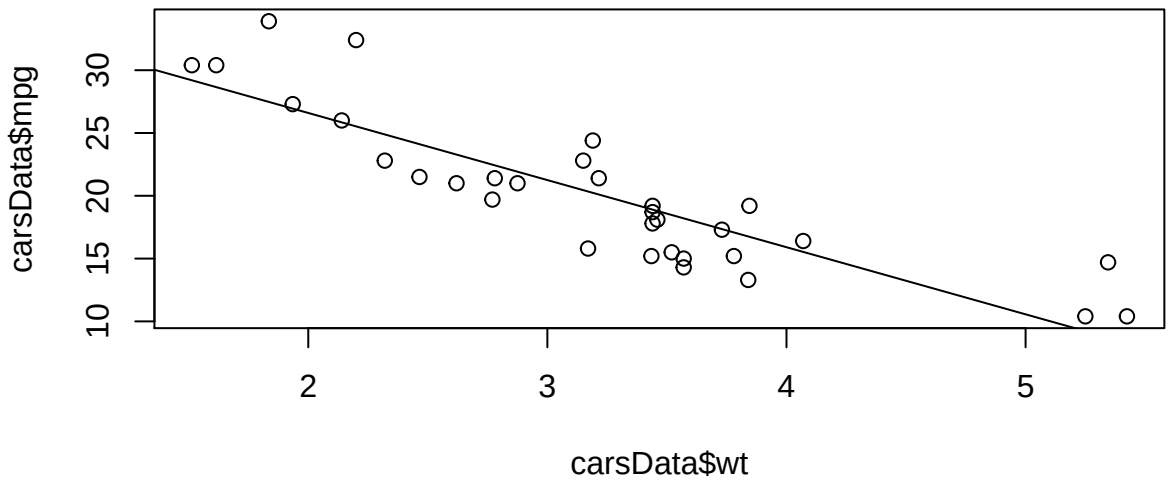


fig 3

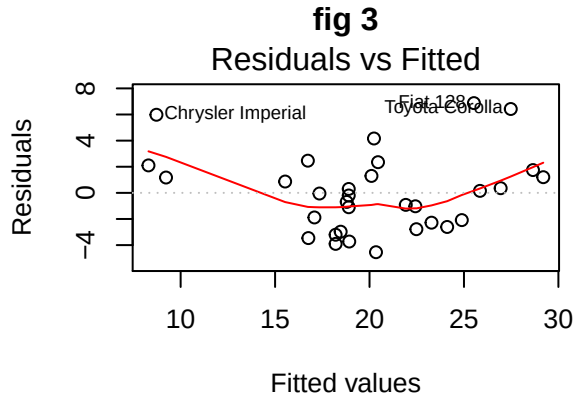


fig 3

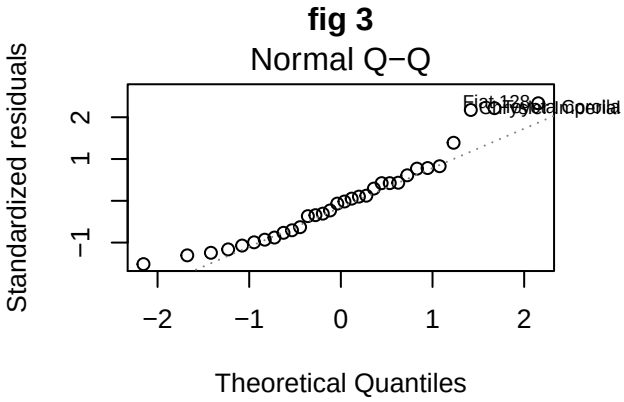


fig 3

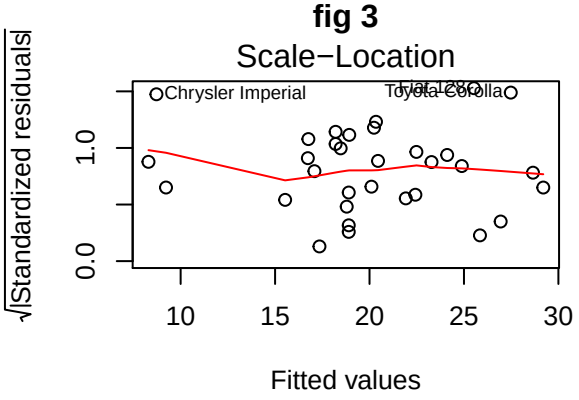


fig 3