Take home 2

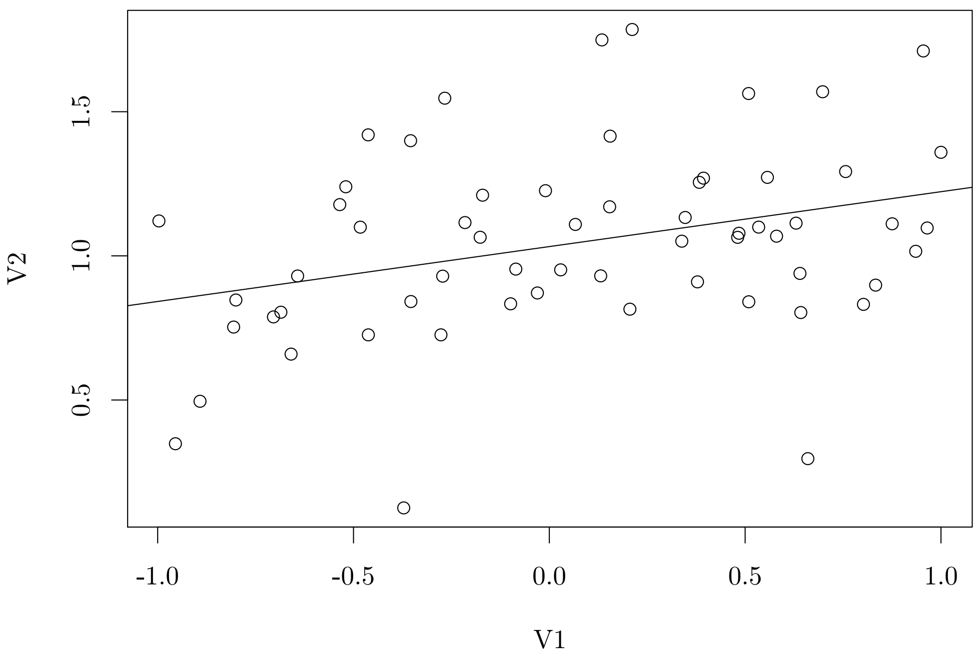
Pieter Luyten

11 december 2019

Question 1

(a)

The value for the intercept of the fit is 1.0321764 and for the rico of the fit is 0.1904323



Figuur 1: line fit throught the data in Ex1.txt

(b)

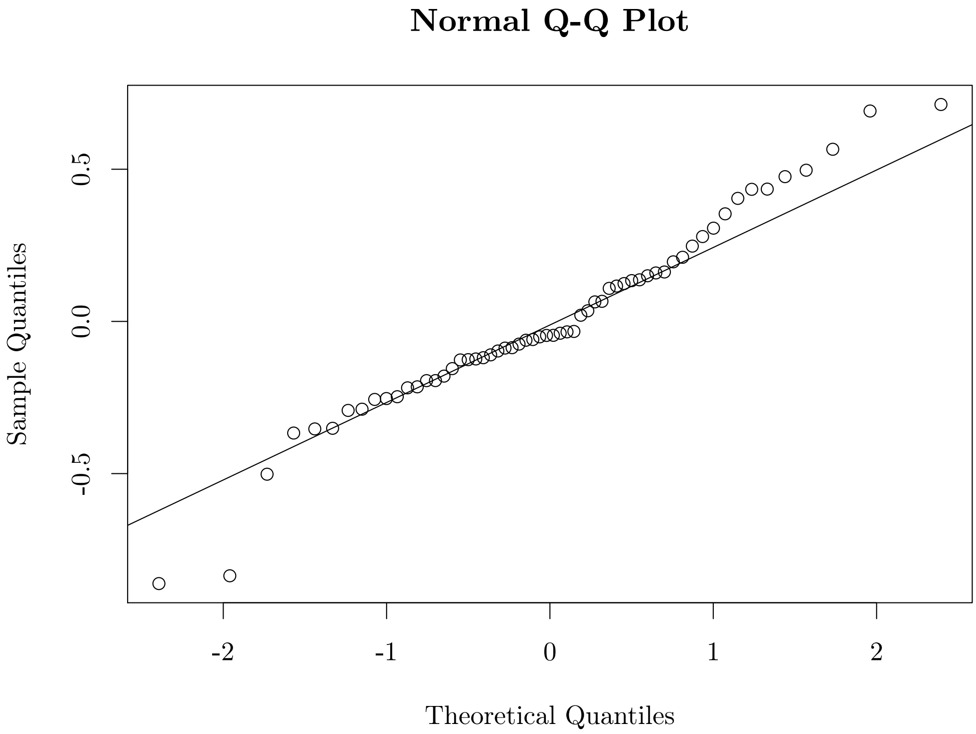
Using the result from section 7.4.1 in <REFERENCE cursus> we know that the random variable

$$T = \frac{\beta_1}{\sqrt{\frac{S^2}{\sum_{i=1}^{n}(x_i - \bar{x}^2)}}} \tag{1}$$

has a Student distribution with $n-2$ degrees of freedom. The test value is 2.603. Using a student-t distribution with $60-2=58$ degrees of freedom we find a p-value of 0.0117. At the confidence level $\alpha = 0.01$, the null hypothesis that $\beta_1 = 0$ holds. The 99% confidence region for the test value is $[-2.663, 2.663]$.

(c)

A q-q plot is a plot where the quantiles of the assumed distribution are plotted against the quantiles from the sample. So in a sample of n points where we label the observation $x_i, i \in \{1, 2, \ldots, n\}$ from lowest to highest, the ith point will be plotted at the $(Q(i/n), x_i)$. Here $Q(x)$ is the quantile function of the assumed distribution. This is a function such that $P(X < Q(p)) = p$. If the assumed distribution is a good model for the observed sample, the points will well fitted by a linear function.



Figuur 2: qq plot for the errors on the fit

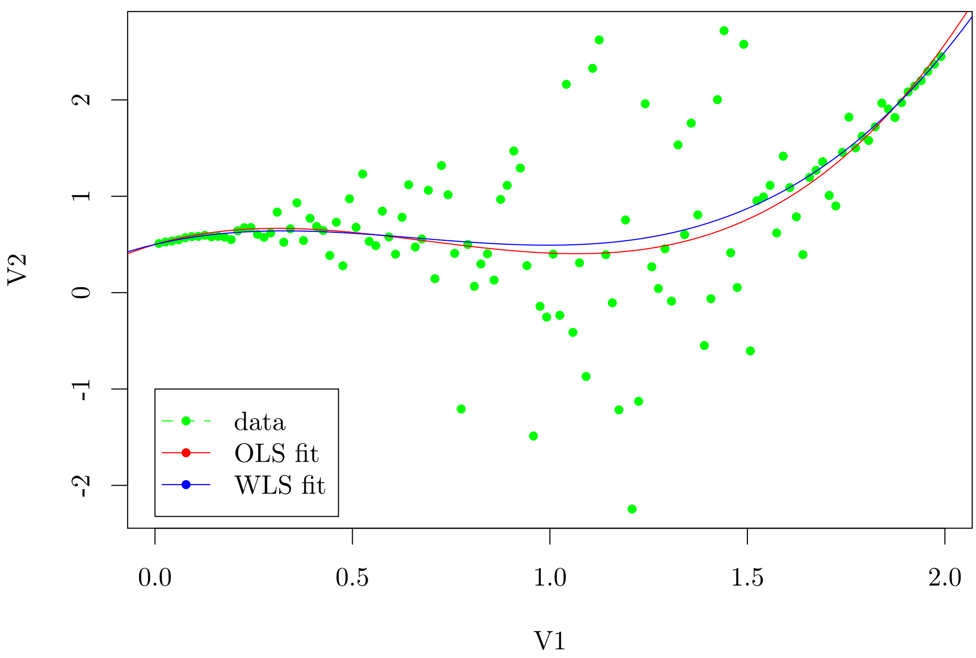
Question 3

(a)

The values for the fitted coefficients with ordinary least squares are: $[0.4994, 1.1985, -2.4959, 1.2082]$. In figure 3 the fit is plotted along with the data.

(b)

The coeffitients fitted with the weighted least square algorithm are $[0.5003, 0.9669, -1.9655, 0.99100]$.



Figuur 3: fit of cubic function throught the data in ex3.txt using ordinary least squares

(c)

The coefficients that are calculated using the least squares method are closer to the real values then the ones calculated using the ordinary least squares. To really say something about the difference and how well they match the given real values of the function we need more information about the distribution of the parameters.

(d)

I guess I will have to adjust the derivation for ordinary least squares from the book???

Acknowledgements

Rune Buckinx and Michaël Maex for helping to fix stupid mistakes in stupid R code Seppe for stupid discussions about how well the questions are asked Thibeau for organising a group crying session Robbe for genuinly good input about trying to solve stupid questions