KEEGAN GYOERY
SID: 470413467

TUTORIAL: Thursday 11am Carslaw 610/11
TUTOR: Prof. E. Senata

21 / 8 / 17

MATH 1905 Assignment 1

Q1. $y_i = aw_i + bx_i + c + \varepsilon_i$

$$S_1(a,b,c) = \sum_{i=1}^{n} \left[y_i - (aw_i + bx_i + c)\right]^2$$

The minimisation will be performed in two steps:

$$\min_{a,b,c} S_1(a,b,c) = \min_{a,b}\left[\min_{c} S_1(a,b,c)\right]$$

Firstly, we must find the "best" c when a and b are held fixed. That is, find:

$$\hat{c}(a,b) = \arg\min_{c} S_1(a,b,c)$$

a) To perform the inner minimisation, it suffices to solve the equation:

$$\frac{\partial S_1(a,b,c)}{\partial c} = 0.$$

We are required to show that the solution to this equation is:

$$\hat{c}(a,b) = \bar{y} - a\bar{w} - b\bar{x}$$

and thus that:

$$S_2(a,b) = \sum_{i=1}^{n} \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x})\right]^2$$

The proof is then as follows.

$$S_1(a,b,c) = \sum_{i=1}^{n} \left[y_i - (aw_i + bx_i + c)\right]^2$$

$$\therefore \frac{\partial S_1(a,b,c)}{\partial c} = \frac{\partial}{\partial c}\left[\sum_{i=1}^{n}\left[y_i - (aw_i + bx_i + c)\right]^2\right]$$

$$= \sum_{i=1}^{n}\left[-2\left[y_i - (aw_i + bx_i + c)\right]\right]$$

$$= -2\sum_{i=1}^{n}\left[y_i - (aw_i + bx_i + c)\right]$$

$$\therefore \frac{\partial S_1(a,b,c)}{\partial c} = 0$$

$$\therefore -2\sum_{i=1}^{n}\left[y_i - (aw_i + bx_i + c)\right] = 0$$

$$\therefore \sum_{i=1}^{n}\left[y_i - aw_i - bx_i - c\right] = 0$$

$$\therefore \sum_{i=1}^{n}\left[y_i - aw_i - bx_i\right] = \sum_{i=1}^{n} c$$

$$\therefore \sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} w_i - b\sum_{i=1}^{n} x_i = nc$$

$$\therefore \frac{\sum_{i=1}^{n} y_i}{n} - \frac{a\sum_{i=1}^{n} w_i}{n} - \frac{b\sum_{i=1}^{n} x_i}{n} = c$$

$$\therefore c = \bar{y} - a\bar{w} - b\bar{x}$$

Now, to complete the proof, we use the following fact.

$$S_2(a,b) = S_1(a,b,\hat{c}(a,b))$$

$$\therefore S_2(a,b) = \sum_{i=1}^{n}\left[y_i - (aw_i + bx_i + \bar{y} - a\bar{w} - b\bar{x})\right]^2$$

$$= \sum_{i=1}^{n}\left[y_i - aw_i - bx_i - \bar{y} + a\bar{w} + b\bar{x}\right]^2$$

$$= \sum_{i=1}^{n}\left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x})\right]^2 \qquad (1)$$

b)i To perform the outer minimisations, it suffices to solve the equations:

$$\frac{\partial S_2(a,b)}{\partial a} = 0 \qquad (2)$$

$$\frac{\partial S_2(a,b)}{\partial b} = 0 \qquad (3)$$

The proof for the first equation is as follows.

$$S_2(a,b) = \sum_{i=1}^{n} \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right]^2$$

$$\therefore \frac{\partial S_2(a,b)}{\partial a} = \frac{\partial}{\partial a} \left[\sum_{i=1}^{n} \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right]^2 \right]$$

$$= \sum_{i=1}^{n} \left[-2(w_i - \bar{w}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right] \right]$$

$$= -2 \sum_{i=1}^{n} (w_i - \bar{w}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right]$$

$$\therefore \frac{\partial S_2(a,b)}{\partial a} = 0$$

$$\therefore -2 \sum_{i=1}^{n} (w_i - \bar{w}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right] = 0$$

$$\therefore \sum_{i=1}^{n} (w_i - \bar{w}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right] = 0$$

$$\therefore \sum_{i=1}^{n} (w_i - \bar{w}) \left[y_i - \bar{y} - b(x_i - \bar{x}) \right] = a \sum_{i=1}^{n} (w_i - \bar{w})^2$$

$$\therefore \sum_{i=1}^{n} (w_i - \bar{w})(y_i - \bar{y}) = a \sum_{i=1}^{n} (w_i - \bar{w})^2 + b \sum_{i=1}^{n} (x_i - \bar{x})(w_i - \bar{w})$$

$$\therefore S_{wy} = a S_{ww} + b S_{wx}$$

The proof for the second equation is as follows.

$$S_2(a,b) = \sum_{i=1}^{n} \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right]^2$$

$$\therefore \frac{\partial S_2(a,b)}{\partial b} = \frac{\partial}{\partial b} \left[\sum_{i=1}^{n} \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right]^2 \right]$$

$$= \sum_{i=1}^{n} \left[-2(x_i - \bar{x}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right] \right]$$

$$= -2 \sum_{i=1}^{n} (x_i - \bar{x}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right]$$

$$\therefore \frac{\partial S_2(a,b)}{\partial b} = 0$$

$$\therefore -2 \sum_{i=1}^{n} (x_i - \bar{x}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right] = 0$$

$$\therefore \sum_{i=1}^{n} (x_i - \bar{x}) \left[y_i - \bar{y} - a(w_i - \bar{w}) - b(x_i - \bar{x}) \right] = 0$$

$$\therefore \sum_{i=1}^{n} (x_i - \bar{x}) \left[y_i - \bar{y} - a(w_i - \bar{w}) \right] = b \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\therefore \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = b \sum_{i=1}^{n} (x_i - \bar{x})^2 + a \sum_{i=1}^{n} (w_i - \bar{w})(x_i - \bar{x})$$

$$\therefore S_{xy} = b S_{xx} + a S_{wx}$$

Now, we are required to put the solutions to equations (2) and (3) into the following format:

$$M \begin{bmatrix} a \\ b \end{bmatrix} = \underset{\sim}{v}$$

Thus we arrive at the following results.

$$a S_{ww} + b S_{wx} = S_{wy}$$

$$a S_{wx} + b S_{xx} = S_{xy}$$

$$\begin{bmatrix} S_{ww} & S_{wx} \\ S_{wx} & S_{xx} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} S_{wy} \\ S_{xy} \end{bmatrix}$$

ii If there is no unique solution to the system of equations:

$$a S_{ww} + b S_{wx} = S_{wy}$$

$$a S_{wx} + b S_{xx} = S_{xy}$$

then the determinant of M must be zero. This implies that M is not invertible, and there is no unique solution to the system of equations.

$$\therefore \det M = S_{ww} S_{xx} - S_{wx}^2$$

$$\therefore \det M = 0$$

$$\therefore S_{ww} S_{xx} - S_{wx}^2 = 0$$

$$\therefore \frac{S_{wx}^2}{S_{ww} S_{xx}} = 1$$

$$\therefore r^2 = 1$$

As it can be seen, if the determinant is equal to zero, then the value of r^2 is 1 for the x_i's and w_i's. Thus the x_i's and w_i's are perfectly correlated and lie on a straight (linear) line.

iii Assuming that $\det M \neq 0$, we can find M^{-1}. Thus we are able to solve for a and b, which is done in the following proof.

$$\det M = S_{ww}S_{xx} - S_{wx}^2$$

$$M = \begin{bmatrix} S_{ww} & S_{wx} \\ S_{wx} & S_{xx} \end{bmatrix}$$

$$\therefore M^{-1} = \frac{1}{S_{ww}S_{xx} - S_{wx}^2}\begin{bmatrix} S_{xx} & -S_{wx} \\ -S_{wx} & S_{ww} \end{bmatrix}$$

$$\therefore \begin{bmatrix} S_{ww} & S_{wx} \\ S_{wx} & S_{xx} \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} S_{wy} \\ S_{xy} \end{bmatrix}$$

$$\therefore \frac{1}{S_{ww}S_{xx} - S_{wx}^2}\begin{bmatrix} S_{xx} & -S_{wx} \\ -S_{wx} & S_{ww} \end{bmatrix}\begin{bmatrix} S_{ww} & S_{wx} \\ S_{wx} & S_{xx} \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} =$$

$$\frac{1}{S_{ww}S_{xx} - S_{wx}^2}\begin{bmatrix} S_{xx} & -S_{wx} \\ -S_{wx} & S_{ww} \end{bmatrix}\begin{bmatrix} S_{wy} \\ S_{xy} \end{bmatrix}$$

$$\therefore \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{S_{ww}S_{xx} - S_{wx}^2}\begin{bmatrix} S_{xx} & -S_{wx} \\ -S_{wx} & S_{ww} \end{bmatrix}\begin{bmatrix} S_{wy} \\ S_{xy} \end{bmatrix}$$

$$= \frac{1}{S_{ww}S_{xx} - S_{wx}^2}\begin{bmatrix} S_{xx}S_{wy} - S_{wx}S_{xy} \\ S_{ww}S_{xy} - S_{wx}S_{wy} \end{bmatrix}$$

$$\therefore a = \frac{S_{xx}S_{wy} - S_{wx}S_{xy}}{S_{ww}S_{xx} - S_{wx}^2}$$

$$\therefore b = \frac{S_{ww}S_{xy} - S_{wx}S_{wy}}{S_{ww}S_{xx} - S_{wx}^2}$$

Q2.a) In order to examine and analyse the linear relationship between flow and depth, we first look at the correlation coefficient.

$$r^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}}$$

$$= \frac{(3.74006)^2}{(0.27036)(54.65201)}$$

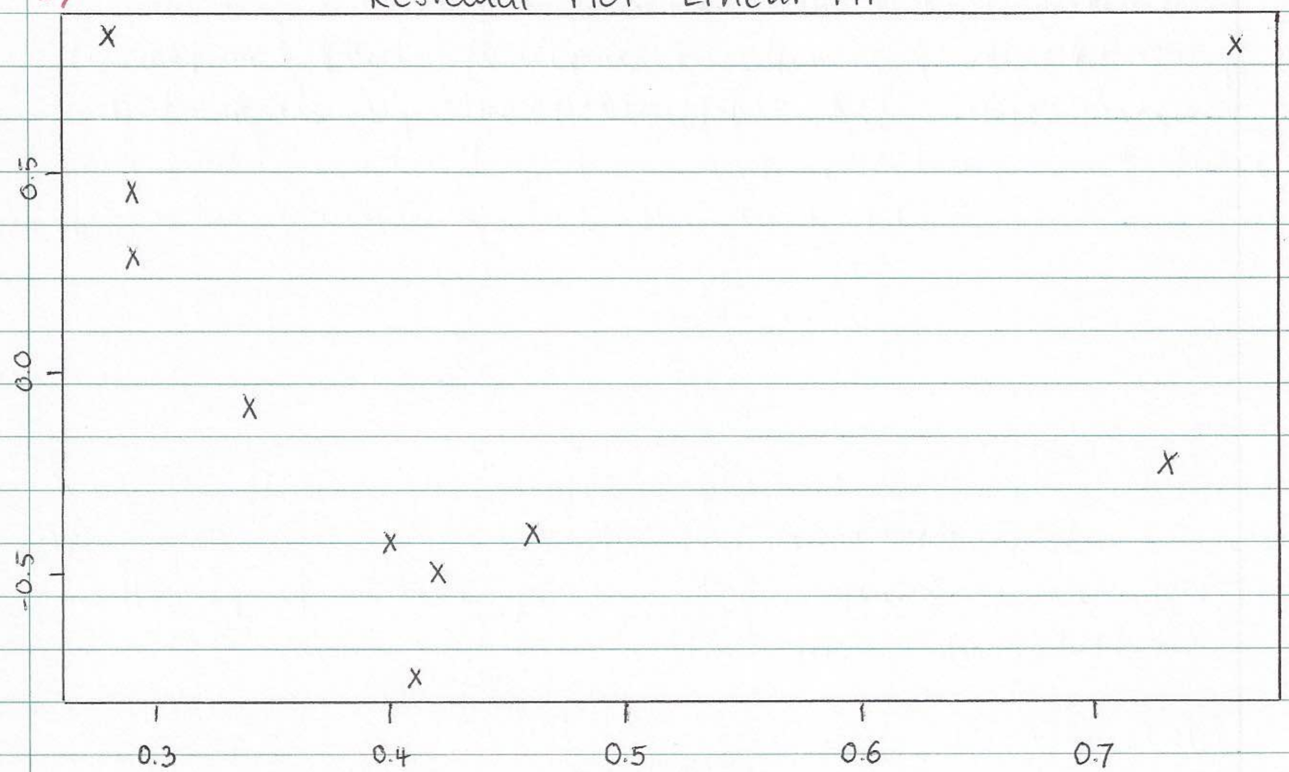
$$= 0.946691683\ldots$$

$$= 0.95$$

As the value for r^2 is close to 1 for this data set, we can interpret that the points lie close to a straight line. However, this does not mean that the linear relationship is a good fit. That is, the linear regression model may not be the most appropriate line of fit. Thus, other models may provide a better fit.

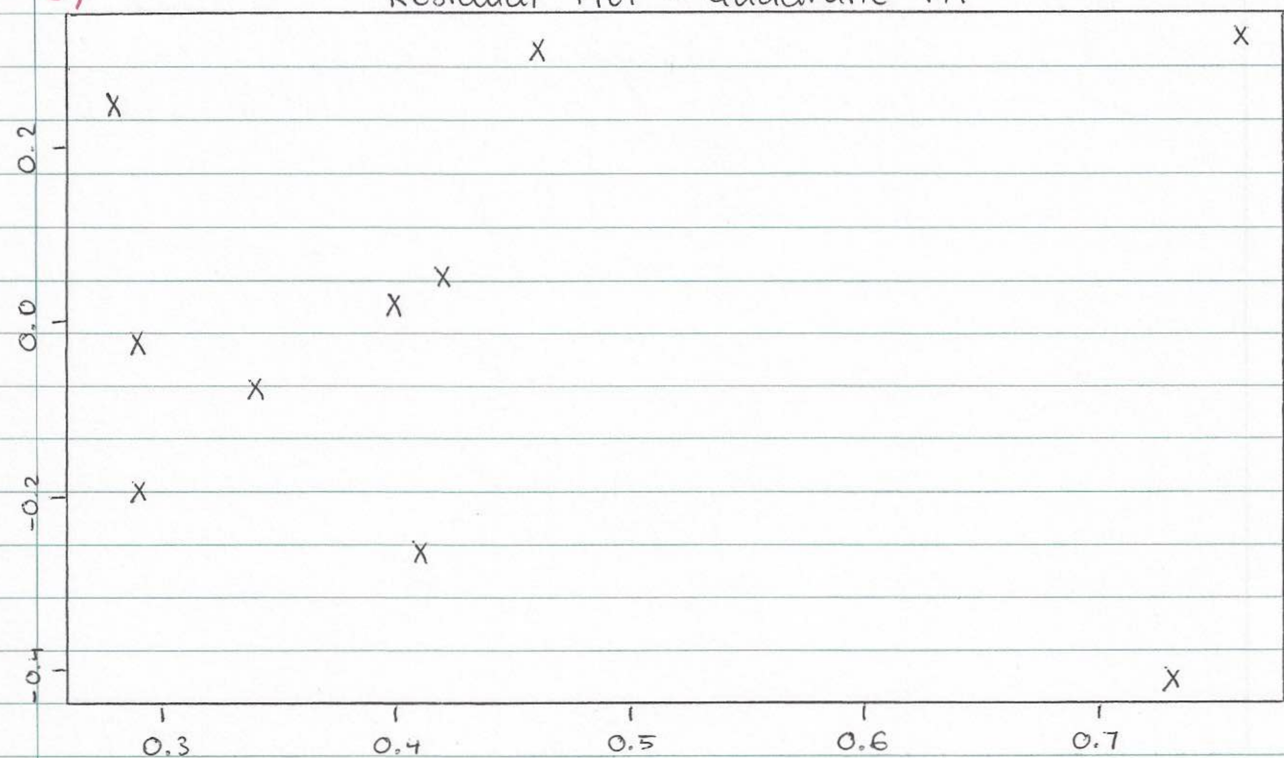
We do not have enough information to make a definite conclusion on the most appropriate model, and must instead use the residuals plots to give clarity on the model that will provide the best fit.

b) Residual Plot - Linear Fit



The flow values are not well explained as a linear function of depth plus "random errors", as the residual plot has a curvature, and is not randomly dispersed. This indicates systematic variation that is not captured by the linear model fit. Thus a linear regression model is not the most appropriate fit for the data, and non-linear models would be more appropriate.

c) Residual Plot - Quadratic Fit



The flow values are well explained as a quadratic function of depth plus "random errors", as the residual plot has no discernable curvature, and is indeed random. Thus there is no systematic variation, and as such, a quadratic regression model is a far more appropriate fit for the data.