

DATA MINING - HMW2 - Julen Ferro

1.1) Ex 6:

If we minimize RSS we get the simple linear regression parameters.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

The regression will be $y_i = \hat{\beta}_0 + \hat{\beta}_1 \cdot x$

Therefore, by operating we get that the regression equation passes through the \bar{x} and \bar{y} point coordinates.

$$y = \hat{\beta}_0 + \hat{\beta}_1 x = (\bar{y} - \hat{\beta}_1 \bar{x}) + \hat{\beta}_1 \bar{x} = \bar{y}$$

4) When comparing the linear regression model vs the cubic one, I think that the ~~cubic~~ cubic one will have a lower training RSS due to the fact that fits better the prediction. It takes into account more terms in the regression equation, so it should fit fine the data prediction with less training residuals. But the final results will be more or less the same because there is not enough information in the statement.

AUTHOR: Julen Ferro Bañales