ANISH RAJESH ADNANI
USC Email: adnani@usc.edu USCID: 4092610491
github: midnightbot

ISLR
3.7.4 I collect a set of data ($n=100$ observations) containing a single predictor and a quantitative response. I the fit a linear regression model to the data, as well as a seperate cubic regression. i.e $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3 + \varepsilon$

(a) Suppose that the true relationship between X and Y is linear, ie $Y = \beta_0 + \beta_1 X + \epsilon$. Consider the training residual sum of squares (RSS) for the linear regression and also the training RSS for the cubic regression. would we expect one to be lower than the other, would we expect them to be the same, or is there not enough information to tell? Justify your answer.

→ $$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Hence the train RSS will be lower if $\hat{y}_i$ is more closer to y_i.
As said above we need $\hat{y}_i$ to be more closer to y_i, that is we want to hammer the curve more to fit the data well. Hence cubic regression will have lower train RSS.

(b) Answer (a) using test rather than training RSS.

→ As we know that true model is linear

So while fitting the model, if we use more flexible models like cubic regression it will hammer itself to the data very well including noise/outliers which will definately give low training RSS but will increase the testing RSS

Hence linear model will have lower testing RSS

(c) Suppose that the true relationship between x and y is not linear, but we do not know how far is it from linear. Consider the training RSS for the linear regression, and also the training RSS for cubic regression. Would we expect them to be same, or is there not enough information to tell? Justify your answer.

→ As in the previous case while we are considering training RSS

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

we want the curve to fit the data well / hammer it more.

If we have a lookup table training RSS = 0, hence increasing the model flexibility will reduce the training RSS, as it fits / hammers the curve to fit the data well compared to less flexible models.

Hence cubic regression will have low training RSS

(d) Answer (c) using test rather than training RSS.

→ As in case (b) for test RSS to be low we want the model to immitate the true curve as closely as possible and not to hamper the curve more to include/fit noise/outliers.

Since we do not know the true curve.

If true ~~curve~~ curve is more closer towards linear and the ~~training~~ test RSS for linear model will be low.

But if true curve is more closer towards cubic and the ~~training~~ test RSS for cubic regression will be low.

Hence there is not enough evidence to correctly answer the question.