Q19

Rahul Atre

2023-10-19

Q19 [Regularization]

Suppose we estimate the regression coefficients in a linear regression model by minimizing $\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})$ subject to $\sum_{j=1}^{p} |\beta_j| \leq s$

For a particular value of s. For parts 1. through 5., indicate which of i. through v. is correct. Justify your answer.

1. As we increase s from 0, the training RSS will:

- i. Increase initially, and then eventually start decreasing in an inverted U shape.

- ii. Decrease initially, and then eventually start increasing in a U shape.

- iii. Steadily increase.

- iv. Steadily decrease.

- v. Remain constant.

Ans: The training RSS will **steadily decrease**. As more beta parameters are added, the model becomes more flexible to better fit the training data, which minimizes training RSS.

2. Repeat 1. for test RSS.

Ans: The test RSS will **decrease initially** but start **increasing** in a **U shape**. Initially, the regression coefficients are less, so the model is not too biased or overfitting. But as we increase the number of beta parameters and minimize to s, it overfits random noise in the data, which can cause an increase in error, thus increasing the test RSS.

3. Repeat 1. for variance.

Ans: The variance will **steadily increase**. When we increase s from 0, more beta parameters allow the model to replicate fluctuations in data. Less constrains are placed on the least squares model, and as such, the amount of variance will increase.

4. Repeat 1. for (squared) bias.

Ans: The bias will **steadily decrease**. As we are aware, bias and variance are inversely proportional to each other due to the bias-variance trade-off. As we increase s from 0, a larger volume of beta parameters will make less assumptions about the data, which will decrease the overall bias of our model.

5. Repeat 1. for the irreducible error.

Ans: The irreducible error will **remain constant**. The irreducible error has no relation to minimizing the function above, since it is always present due to unknowable variables that cannot be reduced further. Minimizing the beta parameters would have no impact on the irreducible error and is independent of s.