

4. Suppose we estimate the regression coefficients in a linear regression model by minimizing

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2$$

for a particular value of λ . For parts (a) through (e), indicate which of i. through v. is correct. Justify your answer.

- (a) For training RSS :

Answer: iii. Steadily increase.

Explanation:

When $\lambda = 0$, the model is equivalent to standard linear regression, which minimizes training RSS without any penalty on the coefficients. As λ increases, the penalty on coefficients increases, making them shrink, which generally increases training RSS because the model is less able to fit the training data.

- (b) For test RSS :

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

Explanation:

As λ increases, the model moves from overfitting to underfitting. At first, reducing overfitting improves test performance, but too much regularization leads to underfitting and increased test error.

- (c) For variance :

Answer: iv. Steadily decrease.

Explanation:

As λ increases, the model becomes simpler and less sensitive to the training data, leading to a steady decrease in variance.

(d) For (squared) bias :

Answer: iii. Steadily increase.

Explanation:

Increasing λ makes the model simpler and less capable of capturing the true underlying patterns in the data, which increases the bias.

(e) For irreducible error :

Answer: v. Remain constant.

Explanation:

The irreducible error is due to random noise in the data and does not change with the choice of model or λ value.