Assignment27

John Bute

2024-11-08

R Markdown

1. As $\lambda \to \infty$, will ^g1 or ^g2 have the smaller training RSS?

The regularization term will dominate as lambda reaches infinity, thus the penalty for non-smoothness becomes increasingly large. Therefore, the third derivative [g^3(x)], will change the curvature of the function g(x) and make it so that the function approaches a quadratic curve , while penalizing the fourth derivative makes the function appraoch a cubic curve. (the idea is that when you take, say, the third derivative of a quadratic function, you end up with 0, therefore as lambda approaches infinity, we force g1, in this case, to become a quadratic function, as the solution that minimalizes the entire function will be quadratic)

Thus, we have a smoother curve(quadratic function) and a less smoother curve (cubic function). Since cubic functions tend to fit data better than quadartic functions, it must have a smaller training RSS, and therefore we will see g2 having a smaller training RSS.

2. As $\lambda \to \infty$, will ^g1 or ^g2 have the smaller test RSS?

g1 will have a smaller test RSS than g2, as we said before, g1 will converge to a quadratic function, while g2 will converge to a cubic function. Therefore, a quadartic function tends to overfit less the data rather than a cubic, and therefore has the lower test RSS.

3. For $\lambda = 0$, will ^g1 or ^g2 have the smaller training and test RSS?

They will both have the same training RSS, as if you remove the regularization term, we allow g1 and g2 to fit the data as closely as possible, leading to overfitting and a OLS regression model. However, since both models are unregularized, we will have the same but way higher test RSS.