 

Advanced Regression Assignment

Author: Anish Mahapatra  
Machine Learning II > Module 3

Q1. What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

A: The optimal values for alpha in ridge and lasso regression are as follows:

* Ridge Regression: Optimal value of alpha is 50
* Lasso Regression: Optimal value of alpha is 0.03

If we were to double the value of alpha in Ridge Regression, then the Negative Mean Absolute error of the Train Score decreases and the Negative Mean Absolute error increases. The bias of the model decreases and there may be an increase of variance of the model.

If we were to double the value of alpha in Lasso Regression, then the R^2 value of the *train* dataset reduces from around 84% to 77% and the value f R^2 score reduces from 82% to 77%. This implies that there is a reduction in the bias and variance of the model, hence, it would not be optimal.

Q2. You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?

A-

Q3. After building the model, you realized that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?

A-

Q4. How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?