Question 1

What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for

both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer 1

The optimal value of alpha for Ridge regression is 10.

Doubling the alpha value increases regularization strength, shrinking coefficients more aggressively.

This may lead to simpler models with fewer influential predictors.

The most important predictors will likely change, favouring features less affected by regularization.

Question 2

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?

Answer 2

I would choose to apply the Lasso regression model with the optimal lambda value.

Lasso has the advantage of performing feature selection by driving some coefficients to exactly zero. This helps in creating a more interpretable and simpler model by excluding less relevant features.

Ridge regression, on the other hand, only shrinks coefficients towards zero but does not force them to be exactly zero.

Question 3

After building the model, you realised 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?

Answer 3

Since the five most important predictor variables from the Lasso model are not available in the incoming data, the next five most important variables will take their place. To identify these new important variables, I would re-run the Lasso regression model without the original top five predictors and observe the coefficients of the remaining variables. The new five most important predictor variables will be those with the highest absolute coefficients in the updated model.

Question 4

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

Answer 4

To ensure a model is robust and generalizable:

Cross-Validation: Use methods like k-fold cross-validation to test the model on different data subsets, avoiding fitting to one dataset.

Feature Selection: Choose important features and skip noise to keep the model stable and avoid overfitting.

Regularization: Methods like Ridge and Lasso keep model coefficients in check, preventing extreme values that could hurt generalization.

Data Separation: Keep test data entirely separate during prep to avoid the model learning from it.

Enough Data: More data means better pattern capture and less reliance on noise.

Hyperparameters: Find the best settings with grid search and cross-validation.

Benefits are accurate predictions on new data and trustworthiness for real-world use. Robust models work consistently on various data, settings, and times, avoiding poor predictions on new data. Models without this struggle with new situations, giving wrong answers and bad decisions.