Relaxed Lasso

Mohamad Lakkis

March, 2024

This section is to explain some of the coding part in the *Relaxed Lasso*

General Overview and Description of the Code

At first, we are fitting many Lasso models with different λ values and calculating the corresponding 5-fold CV error. Then once we have all of these values we will choose the λ which corresponds to the lowest CV error.
In our case of these data, the $\lambda \approx 0.012$, which is consistent with what we found in the Data analysis problem if we chose the best λ based on the lowest CV Error for the lasso model (See figure 3.7), but of course keep in mind that since we are using CV and with 5 fold instead of 10 folds we might get slightly different λ, and since the range of λ that we are choosing from in data Analysis problem is limited.
Now Since we have our desired λ we will fit the lasso model on the scaled data (better to do that in lasso models), and on this particular λ. After that we will get the coefficients of each variable (feature), and we will remove the variables with coefficients = 0 (Notice that no need to scale back the coefficients since $\hat{\beta}_i = \frac{\hat{\beta}_{i,sc}}{\sigma_i}$ as mentioned in the data analysis report). Say Now we have X_{train} modified. This matrix will have dimension n x k, with $k \leq p$, if the original matrix was n x p.
Now we can choose to fit the new lasso models on the X test modified or X train modified, but as mentioned in the question we will fit it based on the X train modified. So using similar steps we get the best λ. And now we fit step2 Lasso model on this lambda and on the X train modified. After that we can now easily predict new data by first scaling them based on X train modified scaler(or the initial X train scaler it won't have a difference since they are the same except for missing columns which won't affect the scaling for the new X modified), and then modifying these new data to remove the 0 coefficients features and predicting new results !
Notice that the test error in the case of the relaxed lasso 0.51 which is slightly bigger than the one we got in the first lasso 0.49 !
Finally we have the second λ is smaller than the first λ value which is consistent with what was said in the problem statement. Plus the first coefficients are smaller than the second coefficients which is also consistent !