

IT585 Advanced Machine Learning
Lab4
Understanding Shrinkage in Constrained/Regularized Regression

Instructions:

1. You have to code the solution in Google colab
2. You can use inbuilt libraries from python
3. Your plots, code, any insights, observations written as text should be submitted as one ipynb file to google classroom
4. Deadline : February 19,2024 11:59 PM IST
5. Name of your file should be : yourrollno_lab4.ipynb

In this lab we will analyze the effects of changing the parameter lambda on the lasso and ridge regression coefficients.

The solution vector of the Ridge Regression is given as

$$\hat{\beta}^{\text{ridge}} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p \beta_j^2 \right\}.$$

Here beta is a (p+1) dimensional vector including the bias term beta_0. X is your data matrix and y is the response vector.

Similarly solution vector for the lasso regression is given as

$$\hat{\beta}^{\text{lasso}} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\}.$$

As discussed, the regularization parameter lambda can be thought of as corresponding to the shrinkage of the solution space.

Take the diabetes dataset and the boston house-prices dataset from the sklearn library. Make sure that each column is normalized to have zero mean and unit norm. Solve the ridge and lasso regression on both the datasets for a number of different values of Lambda . Plot the values of each coefficient of the solution vector beta (i.e beta_1, beta_2 ,... , beta_p) separately (on the Y axis) vs the value of Lambda (on the X axis) for both LASSO and Ridge Regression and for both the datasets. You get a sequence of values for each coefficient of the beta vector

for different values of λ . This is called the regularization path. Comment on your observations and insights that you get from regularization paths of ridge and LASSO regression. Compare and contrast them. Next read about the [Elastic net regularized regression](#) and also do the same experiment for that for different combinations of λ_1 and λ_2 . Provide appropriate plots.

A very good discussion is given in the book “Elements of Statistical Learning” by Tibshirani et al. in the chapter “Linear Methods for Regression” which you can refer to with proper citations.