HW3 - REPORT

Vito Založnik

I. Introduction

Our task was to implement Ridge and Lasso regression algorithms. For Ridge regression, we used closed form solution, and for Lasso regression, we minimized the β matrix defined as

$$\beta = argmin_{\beta}(\sum_{i=1}^{n}(\beta^T x_i - y_i)^2 + \lambda \sum_{i=1}^{k} |\beta_i|).$$

Then we had to apply it to the Superconductor dataset and find optimal λ values. The dataset contained 81 numerical features that we used to predict desired critical temperature. Then we had to evaluate the selected Ridge and Lasso regression models on the last 100 samples from the dataset. For the score function, we used root mean squared error (RMSE). Uncertainty was evaluated with standard error (SE).

II. Data preprocessing

As both regularizations impact high linear coefficients, that depend on the scale of our features, we standardized all of our feature values. We used standardization only when the number of training samples was higher than 30. We also experimented with applying PCA to desired percentage of variance to reduce the difference between weights between both regressions using λ parameter 0. This also improved the time complexity of the models.

Given the dataset of 300 samples, we split the dataset into test and train sets, using the first 200 samples for a train set and the last 100 for a test set. We used 10-fold cross-validation and RMSE to select the best λ parameter for each regression. Then we bootstrapped train data 200 times, trained the model, and evaluated obtained Ridge model on a whole test dataset so that we obtained the mean RMSE and its standard error.

III. Results

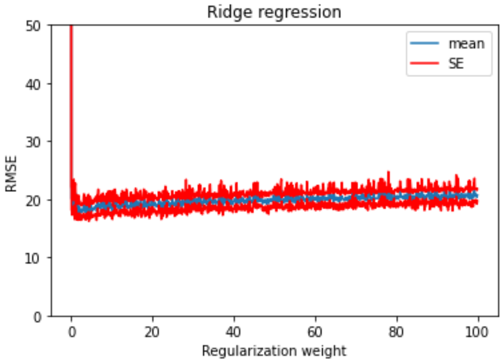

Fig. 1. RMSE on train data with respect to regularization weight

As we did not want our models to result in underfitting, we decided to compare only λ values lower than 5.

In figure 1 we can see the changing mean RMSE and its SE with respect to regularization weight, using Ridge regression. The minimmum of mean RMSE was obtained using $\lambda = 2$.

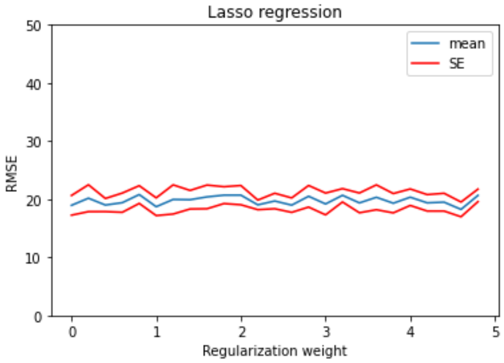

Fig. 2. RMSE on train data with respect to regularization weight

In figure 2 we can see the changing mean RMSE and its SE with respect to regularization weight, using Lasso regression. The minimum of mean RMSE was obtained using $\lambda = 0.5$

Model	λ	Mean RMSE $\pm SE$
Lasso regression	0.5	20.0371 ± 1.4987
Ridge regression	2	18.3624 ± 1.5107

TABLE I
OPTIMAL SELECTED WEIGHTS

In the table I, we can see calculated optimal lambda values based on 10-fold cross-validation.

IV. Model evaluation

With the optimal weights, we evaluated the final RMSE on the test data set. To measure uncertainty we fitted our model 50 times with 200 bootstrapped samples of our training data set.

Model	λ	Mean RMSE $\pm SE$
Lasso regression	0.5	24.8453 ± 1.1856
Ridge regression	2	22.7999 ± 0.7369

TABLE II
BOOTSTRAPED TEST DATA - WITHOUT USING PCA

Model	λ	Mean RMSE $\pm SE$
Lasso regression	0.5	17.3893 ± 0.1065
Ridge regression	2	17.4096 ± 0.1034

TABLE III
BOOTSTRAPED TEST DATA - WITH USING PCA

Without applying PCA, the ridge regression model performed better out of the two, with better RMSE and standard error. After applying the same procedure with only adding PCA (with 80% of variance kept) to the data preprocessing process, we obtained better and much more certain results. Both models performed roughly the same.