Yu Jinghuan
5529 0708
CS 5487 Machine Learning
2018年12月2日

CS5487 Programming Assignment 1

Part 1 Polynomial function

a) Implementation

In this Assignment, I use the *numpy* and *scipy* as the basic framework and implement the LS, RLS, RR. Reference coordinate descent method to implement the LASSO. Each Python file is runnable.

b) Regression Results

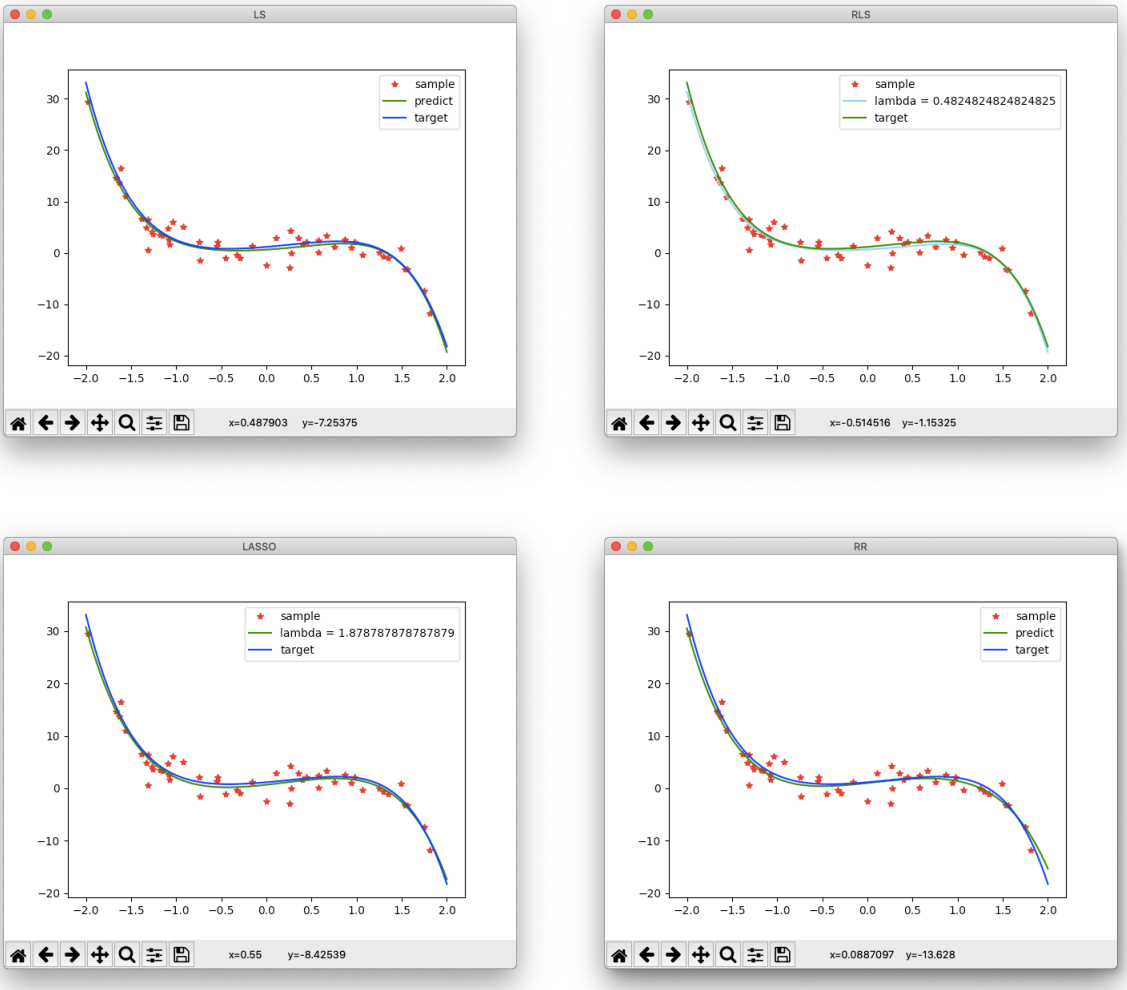

Fig 1 . Regression Result using LS, RLS, LASSO and RR

Fig 1. Shows the result of each regression algorithms, use red dots to represent the sample points, as for regression methods RLS and LASSO, traversing the array from generally feasible ranges and choose the hyper-parameters with the least mean-squared error. The mean-squared error of different regression methods are listed in the Table 3.

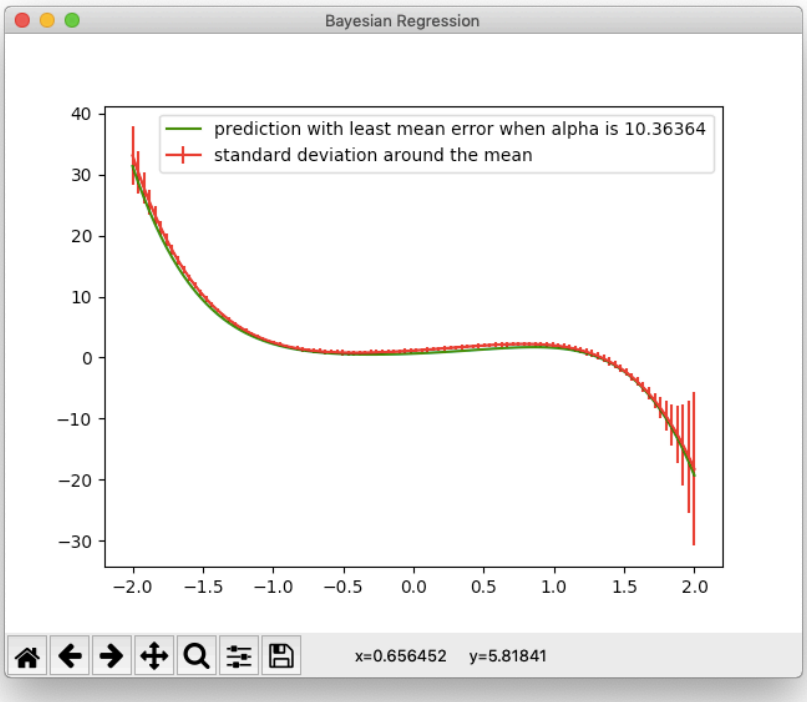

Fig 2. Bayesian regression With Standard Deviation

Bayesian regression and standard deviation around the mean are shown in the Fig 2. Same as the RLS and LASSO, traversing the possible alpha to detect the best behaved hyper-parameters.

Table 3. Regression Result Analysis

Algorithm	Mean Squared Error	Hyper Paramemter
LS	0.41	NA
RLS	0.41	0.48
LASSO	0.43	1.88
RR	0.77	NA
BR	0.41	10.36

c) Regression Results with reduced amount of training data

For this question, the amount of training data is reduced to 10%, 25%, 50% and 100% of original training data. Figure 4~7 shows the different results of each datasets and algorithms.

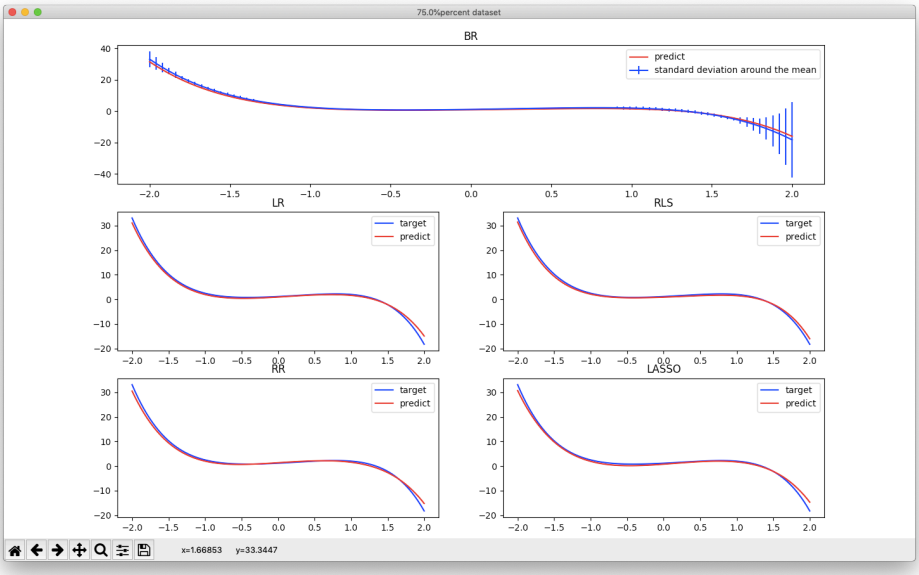

Fig 4. Regression Results with 75% data

In order to observe the overfitting stably. We choose the training set from small to large of the total sample data's X-axis values. But during statistics of the errors, we use random sets and repeat 1000 time to get the average.

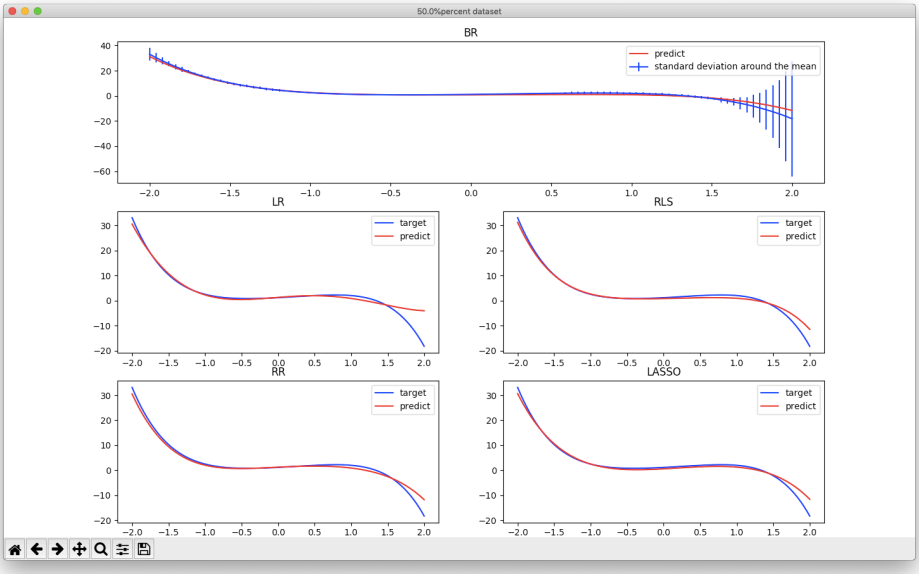

Fig 5. Regression Results with 50% data

Comparing the Bayesian Regression results shown in Figure 4 and Figure 5. The deviation around the mean increased at the end of the dataset, this means this regression method has the trend of the training sets.

Other regression algorithms also have more gap between the target values and predict values. The most obvious one is the LS regression.

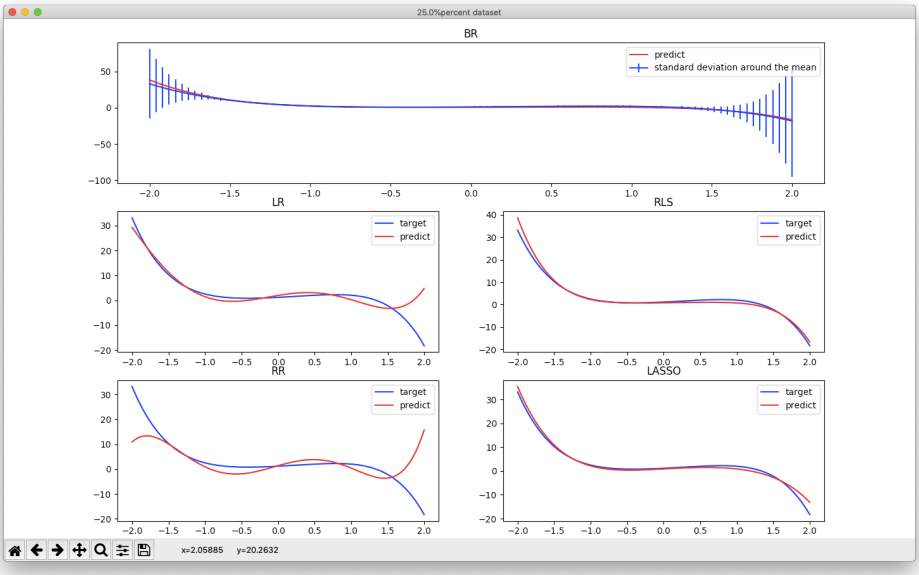

Fig 6. Regression Results with 25% data

When the data set is further reduced, the overfitting becomes more and more severe. LS and RR are becoming twisted.

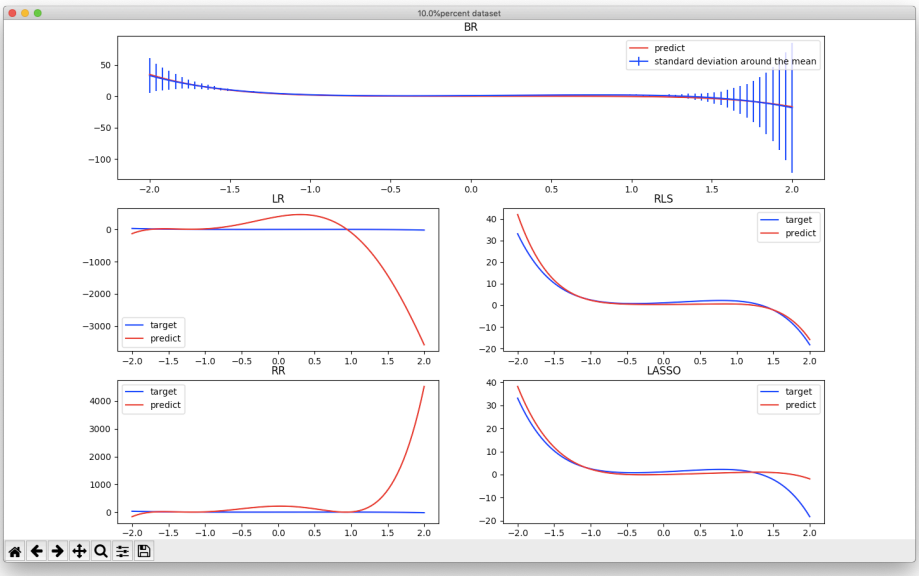

Fig 7. Regression Results with 10% data

Figure 8. statistics and depicts the MSE of different algorithms under different data sets. Since the MSE of LS and RR algorithm in small training set are very big, the scope of y axis is limited to receive meaningful results.

From the figure we can conclude: LS model and RR model present the worst performance in small datasets, for they have no regulation item to avoid overfitting, which means they are very sensitive. BR may be the most robust with less data.

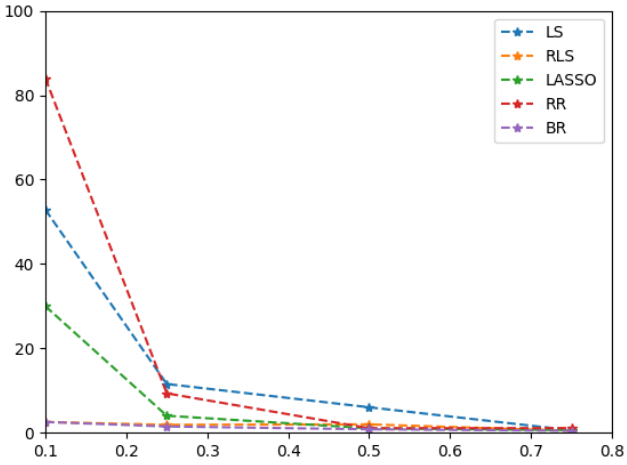

Fig 8. MSEs of each algorithms

d) Regression Results with outliers output value

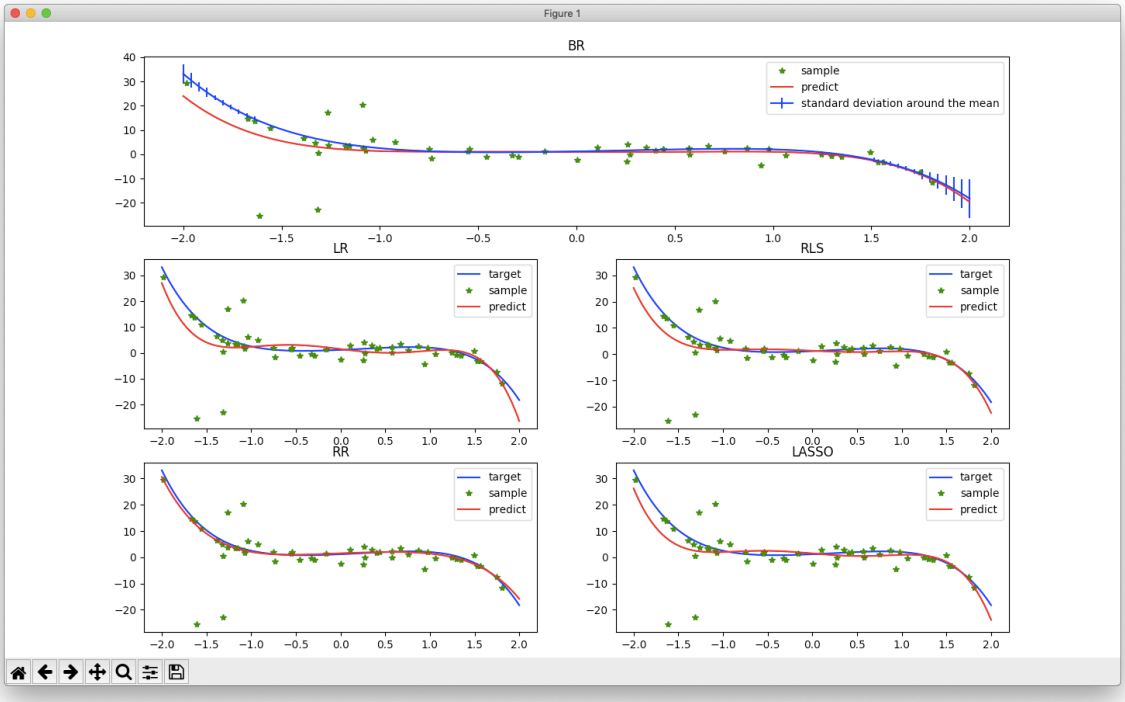

Fig 9. Regression Results with outliers output values

For this problem, some random outliers are added. The ratio of the outliers over training data is 10% and the range for outliers is from -50 to 50. RR is the most robust for it use L1 norm as the objective function while other regressions are using L2 norm, which will amplify the value of deviation.

e) Regression Results with higher-order polynomial

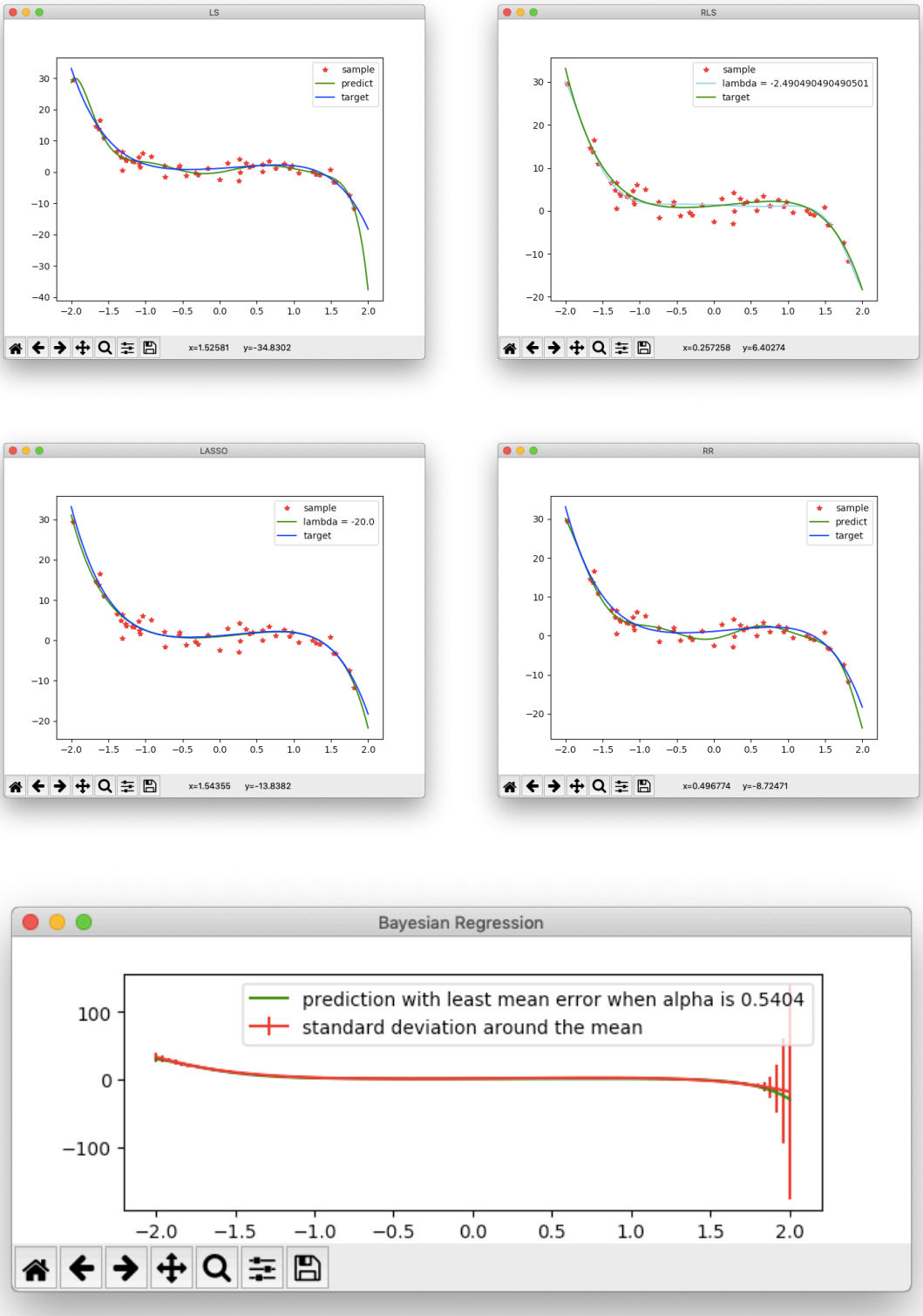

Fig 10. Regression Results with outliers output values

Table 11. Regression Result Analysis

Algorithm	Mean Squared Error (10th)	Mean Squared Error (5th)	Hyper Paramemter

Table 11. Regression Result Analysis

LS	7.98	0.41	NA
RLS	0.57	0.41	-2.49
LASSO	0.62	0.43	-20.00
RR	1.75	0.77	NA
BR	2.87	0.41	0.54

10th order polynomial is estimated and the results are presented in Figure 10. The results show that RLS is most robust. All the regressions perform worse than 5th order polynomial for result in higher MSE.

Part 1 Polynomial function

a) Using the features directly

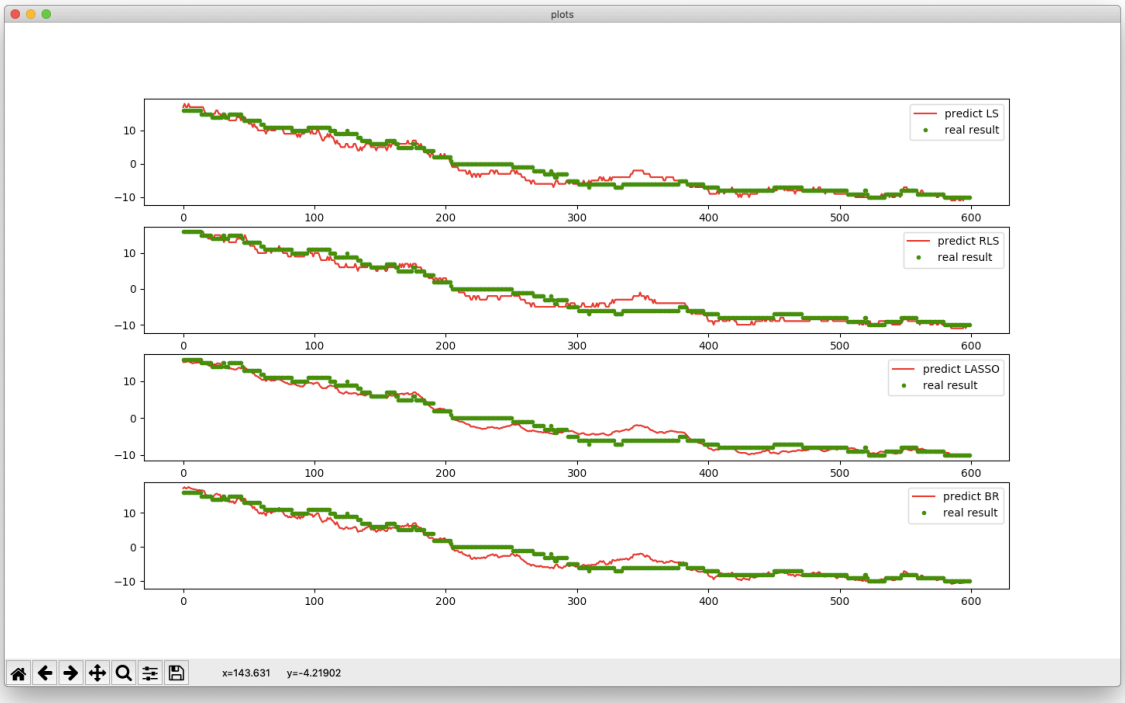

Fig 12. Regression Result of four main regression.

Table 13. Regression Result Analysis

	LS	RLS	LASSO	BR
MSE	3.20	2.78	2.36	3.04
MAE	1.34	1.30	1.24	1.35

Fig 13. Bar graph of different MAE and MSE

Table 2 shows the results of MSE and MAE for the 5 estimations. All the regression methods don’t perform very well. But LASSO is a little better than others.