

Computer simulation of the difference between Lasso and Lars

Huiyi Xia*

School of Big Data an Artificial Intelligence Chizhou University, Anhui, China

Abstract. In order to correct the error that Efron et al pointed out that the modified Lars is a fast algorithm for lasso problem, computer simulation is carried out with data. The simulation results show that when a variable is removed first and then entered in the lasso problem, the variables selected by lasso and Lars cannot be the same. At this time, no matter how to modify the size of the variables selected by Lars, the lasso problem cannot be solved. The computer simulation method is used to correct the errors of Efron et al in this paper, and contribute to development of world data computing.

1 Introduction

Tibshirani¹ proposed lasso in 1996, which is a method of variable selection. At the same time, he proposed the theory of lasso geometry and lasso soft threshold estimation, and solved the calculation of lasso with quadratic programming algorithm. Efron et al² proposed Lars in 2004 and proved that the modified Lars can be used to solve lasso problem. Due to the fast-computing speed of Lars, Lars quickly replace the original algorithm of lasso and becomes the main algorithm in lasso research. In practical application, Lars has fast calculation speed and lasso has wide application range, which makes lasso and Lars quickly applied to empirical research of many disciplines. The two articles that put forward these two methods, “regression shrinkage and selection via the lasso” and “least angle regression”, have high citation rates.

Fan and Li³ put forward the conjecture that lasso estimation has no “Oracle property” in 2001, and proved that SCAD estimation has “Oracle property”. Zou and Hastie⁴ proposed the elastic net in 2005 and compared it with lasso. Zou⁵ proposed the adaptive lasso in 2006, proved that lasso estimation has not “Oracle property” in some cases, and pointed out that the adaptive lasso estimation has “Oracle property”. Due to the great significance of lasso, Tibshirani⁶ was invited to publish an article "regression shrinkage and selection via the lasso: a retrospective" in “Journal of the Royal Statistical Society” in 2011 to systematically review the research of lasso.

* Corresponding author: xiayz_88@163.com

2 The difference between lasso and Lars

When studying a practical problem, the author found that the drawing drawn by lasso is not good, which is different from the general understanding of lasso. Therefore, the author found the difference between lasso and Lars through research. This case is described below:

Example: The response variable y and observation variables $x1, x2, x3, x4, x5$ and $x6$ are:

$$y=(1291,1409,1591,1806,1997,2063,2053,2112,2170,2250,2359,2778,3167.656945,3366.787659,3863.903802,3800,4681,5032,5357.099843)$$

$$x1=(3.533392,4.819786,6.079373,7.117603,7.897359,8.440228,8.967705,9.921417,10.965517,$$

$$12.033269,13.582276,15.987834,18.493737,21.631443,26.581031,31.404543,34.090281,40.151280,47.319400)$$

$$x2=(4348.95, 5218.10, 6242.20, 7407.99, 8651.14, 9875.95, 11444.08, 13395.23,16386.04, 18903.64,$$

$$21715.25,26396.47,31649.29,38760.20,51321.78,61330.35,68518.30,83080.00,108930.00)$$

$$x3=(3.15,4.69,6.08,7.61,9.10,10.45,11.99,13.46,15.83,18.50,22.12,25.41,29.60,34.56,$$

$$40.34, 47.51,60.62,72.60,85.20)$$

$$x4=(1.351000,1.596882,1.841205,1.913012,1.937881,1.856490,1.795226,1.886782,1.883009,$$

$$1.839700,1.928005,2.147798,2.326065,2.465629,2.574117,2.844399,2.619691,2.871182,3.132459)$$

$$x5=(255.1,276.8,290.3,301.6,311.9,329.9,360.6,383.7,416.3,472.1,514.6,554.2,607.4,667.7,752.5,815.7,895.4,965.2,1046.3)$$

$$x6=(1333.8,1789.4,2337.9,2806.7,2977.2,2995.5,2987.4,3146.2,3306.9,3431.7,3582.4,4039.6, 4631.2,5025.1,5791.1,6700.7,7115.6,8119.5,9833.1)$$

Observe the difference between lasso and Lars.

Solution: (1). Use the commands in Lars' software package to draw Lar diagram and lasso diagram of the problem:

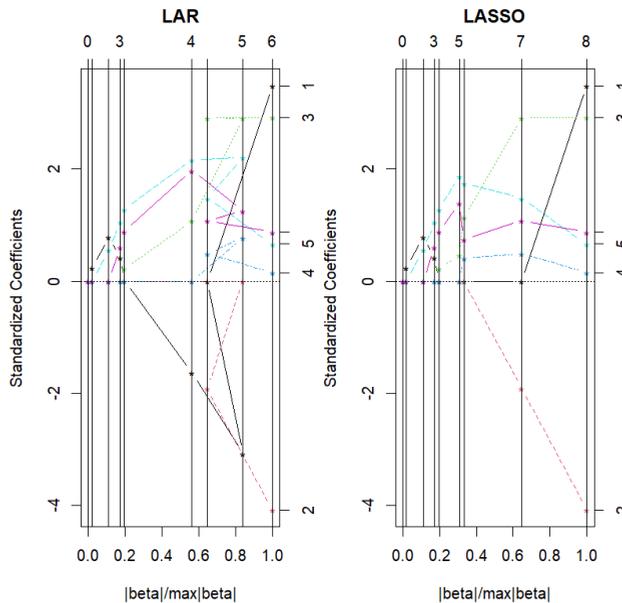


Fig. 1. Graphs of Lars and lasso.

Through figure 1, it is found that the variable selection process of lasso and Lars is different.

(2). The coefficient changes of lasso and Lars are obtained by using the coefficient command of Lars' software package:

Table 1. Variation of Lars coefficient.

	<i>x1</i>	<i>x2</i>	<i>x3</i>	<i>x4</i>	<i>x5</i>	<i>x6</i>
1	0.055	0.000	0.000	0.000	0.000	0.000
2	0.187	0.000	0.000	0.000	0.131	0.000
3	0.100	0.000	0.000	0.000	0.249	0.143
4	-0.385	0.000	0.254	0.000	0.512	0.466
5	-0.725	0.000	0.686	0.184	0.520	0.296
6	0.822	-0.962	0.689	0.036	0.156	0.207

Table 2. Variation of lasso coefficient.

	<i>x1</i>	<i>x2</i>	<i>x3</i>	<i>x4</i>	<i>x5</i>	<i>x6</i>
1	0.055	0.000	0.000	0.000	0.000	0.000
2	0.187	0.000	0.000	0.000	0.131	0.000
3	0.100	0.000	0.000	0.000	0.249	0.143
4	0.000	0.000	0.0523	0.000	0.303	0.209
5	0.000	0.000	0.109	0.000	0.441	0.330
6	0.000	0.000	0.269	0.097	0.413	0.176
7	0.000	-0.450	0.687	0.115	0.350	0.254
8	0.822	-0.962	0.689	0.036	0.156	0.207

The coefficients of 1, 2, 3 and 6 in Table 1 are consistent with those of 1, 2, 3 and 8 in Table 2, and other coefficients are inconsistent.

(3). Use the commands of Lars' software package to obtain the variable selection order of Lars and lasso:

Table 3. Variable selection order of Lars.

	<i>x1</i>	<i>x5</i>	<i>x6</i>	<i>x3</i>	<i>x4</i>	<i>x2</i>
var	1	5	6	3	4	2
step	1	2	3	4	5	6

Table 4. Variable selection order of lasso.

	<i>x1</i>	<i>x5</i>	<i>x6</i>	<i>x3</i>	<i>x1</i>	<i>x4</i>	<i>x2</i>	<i>x1</i>
var	1	5	6	3	-1	4	2	1
step	1	2	3	4	5	6	7	8

Through variable selection order of lasso variable in Table 4, it is found that *x1* is removed first, then entered. The Lars change in Table 3 cannot show this process.

To sum up, Lars, which only modifies the size of variable coefficients, cannot solve the lasso problem.

3 Error of Efron et al.

There are errors in the proof of "Algorithm 3.2 a least angle regression: lasso modification" on pages 76-77 of "the elements of statistical learning"⁷. The proof of the book cannot guarantee that the variables selected by Lars and lasso are consistent. If the variables

selected by lasso and Lars are consistent, the modified Lars according to lasso can solve the lasso problem. If the variables selected by Lars and lasso are inconsistent, the modified Lars according to lasso will get worse and worse and cannot solve the lasso problem.

The data simulation shows that as long as the amount of data becomes larger, the variables selected by lasso and Lars are inconsistent, therefore, as long as Efron et al carefully carry out computer simulation, this error will not appear.

References

1. Tibshirani R 1996 Regression shrinkage and selection via the lasso *Journal of the Royal Statistical Society(Series B (methodological)58(1))* p 267-288
2. Efron B, Hastie T, Johnstone I, Tibshirani R 2004 Least angle regression *The Annals of Statistics (32(2))* p 407-499
3. Fan J, Li R 2001 Variable Selection via Nonconcave penalized likelihood and its oracle properties *Journal of the American Statistical Association (96(456))* p1348-1360.
4. Zou H, Hastie T. 2005 Regularization and variable selection via the elastic net *Journal of the Royal Statistical Society (Series B, 67(2))* p 301-320
5. Zou H 2006 The adaptive lasso and its oracle properties *Journal of the American Statistical Association (101(476))* p 1418-1429
6. Tibshirani R 2011Regression shrinkage and selection via the lasso: a retrospective *Journal of the Royal Statistical Society (Series B, , 73(3))* p 273-282.
7. Hastie T, Tibshirani R, Friedman J 2008 *The Elements of Statistical Learning (Second Edition)* (Spring) p 76-77.