

## Research Article

# Method for Solving LASSO Problem Based on Multidimensional Weight

**Chen ChunRong, Chen ShanXiong, Chen Lin, and Zhu YuChen**

*College of Computer & Information Science, Southwest University, Chongqing, China*

Correspondence should be addressed to Chen ShanXiong; [csxpml@163.com](mailto:csxpml@163.com)

Received 15 November 2016; Revised 12 February 2017; Accepted 21 March 2017; Published 4 May 2017

Academic Editor: Farouk Yalaoui

Copyright © 2017 Chen ChunRong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the data mining, the analysis of high-dimensional data is a critical but thorny research topic. The LASSO (least absolute shrinkage and selection operator) algorithm avoids the limitations, which generally employ stepwise regression with information criteria to choose the optimal model, existing in traditional methods. The improved-LARS (Least Angle Regression) algorithm solves the LASSO effectively. This paper presents an improved-LARS algorithm, which is constructed on the basis of multidimensional weight and intends to solve the problems in LASSO. Specifically, in order to distinguish the impact of each variable in the regression, we have separately introduced part of principal component analysis (Part\_PCA), Independent Weight evaluation, and CRITIC, into our proposal. We have explored that these methods supported by our proposal change the regression track by weighted every individual, to optimize the approach direction, as well as the approach variable selection. As a consequence, our proposed algorithm can yield better results in the promise direction. Furthermore, we have illustrated the excellent property of LARS algorithm based on multidimensional weight by the Pima Indians Diabetes. The experiment results show an attractive performance improvement resulting from the proposed method, compared with the improved-LARS, when they are subjected to the same threshold value.

## 1. Introduction

Data mining has shown its charm in the era of big data; it has gained much attention in academia regarding how to mine useful information from mass data by mathematical statistics model [1, 2]. In linear model, model error usually is a result of the lack of key variable. At the beginning of the modeling, generally, the more variables (attribute set) chosen are, the less the model error is. But, in the process of modeling, we need to find the attribute set which has the largest explanatory ability to the response, that is, improving the prediction precision and accuracy of the model through selecting variable [3]. Linear regression analysis is the most widely used of all statistical techniques [4]; the accuracy of that analysis mainly depends on the selection of variables and values of regression coefficients [5]. LASSO is an estimate method which can simplify the index set. In 1996, inspired by the ridge regression (Frank and Friedman, 1993) [6] and Nonnegative Garrote (Breiman, 1995) [7], Tibshirani proposed one new method of variable selection. The idea of this method is minimizing the square of the

residuals with the constraint that the sum of the absolute values of regression coefficient  $\beta$  is less than a constant by construct a penalty function to shrinkage coefficient [8]. As a kind of compression estimates, the LASSO method has higher detection accuracy and better parameter convergence consistency. Efron et al. (2002) proposed the LARS algorithm to support the solution of LASSO [9]. And they proposed improved-LARS algorithm (2004) to eliminate the opposite sign of regression coefficient  $\beta$  and solve LASSO better [10]. The improved-LARS algorithm regresses stepwise; each path keeps the correlation between current residual individual and all the variables the same. It also satisfies the solution of LASSO with the same current approach direction and ensures the optimal results and algorithm complexity.

Zou (2006) introduced the adaptive-LASSO by using the different tuning parameters for different regression coefficients. He suggests minimizing the following objective function [11]:

$$\sum_{j=1}^n (y_j - x_j^T \beta)^2 + n \sum_{i=1}^p \lambda_i |\beta_i|. \quad (1)$$

Keerthi and Shevade (2007) proposed a fast tracking algorithm for LASSO/LARS [12]; it approximates the logistic regression loss by a piecewise quadratic function.

Charbonnier et al. (2010) suggest that  $\beta$  owns an internal structure that describes classes of connectivity between the variables [13]. They present the weighted-LASSO method to infer the parameters of a first-order vector autoregressive model that describes time course expression data generated by directed gene-to-gene regulation networks.

Since the LASSO method minimizes the sum of squared residual errors, even though the least absolute deviation (LAD) estimator is an alternative to the OLS estimate, Jung (2011) proposed a robust-LASSO-estimator that is not sensitive to outliers, heavy-tailed errors, or leverage points [14].

Bergersen et al. (2011) found that a large value of  $w_j$ , the regression coefficient for variable  $j$ , is subject to a larger penalty and therefore is less likely to be included in the model, and vice versa [15]. They proposed to use weighted-LASSO with integrated relevant external information on the covariates to guide the selection towards more stable results.

Arslan (2012) found that, compared with the LAD-LASSO method, the weighted LAD-LASSO (WLAD-LASSO) method will resist the heavy-tailed errors and outliers in explanatory variables [16].

LASSO problem is a convex minimization problem; the forward-backward splitting operator method is important to solving it. Salzo and Villa (2012) proposed accelerated version to improve the method's convergence ability [17].

Zhou et al. (2013) proposed an alternative selection procedure based on the kernelized LARS-LASSO method [18]. By formulating the RBF neural network as a linear-in-the-parameters model, they derived a  $l_1$ -constrained objective function for training the network.

Zhao et al. (2015) added two tuning parameters  $\lambda$  and  $j_0$  to the wavelet-based weighted-LASSO methods. The tuning parameter  $\lambda$  controls the model sparsity. The choice of  $j_0$  controls the optimal level of wavelet decomposition for the functional data. They improved wavelet-based LASSO by adding a prescreening step prior to the model fitting or, alternatively, by using a weighted version of wavelet-based LASSO [19].

Salama et al. (2016) proposed a new LASSO algorithm, the minimum variance distortionless response (MVDR) LARS-LASSO [20], which solves the DOA problem in the CS framework.

In light of superior performance achieved in [10] for solving LASSO problem, a new idea is extended in this paper into the uses of multidimensional weight LARS. Our main contributions are as follows:

- (i) In the solving process of LASSO, each attribute in the evaluation population has different relative importance to the overall evaluation. The relative importance include the following: not all attributes influence the regression results and each individual in the regression model has different weight. When

improved-LARS algorithm calculated the equiangular vector, we distinguish the effect resulting from different attribute variable, considering joint correlation between regression variables and surplus variable.

- (ii) We discuss the method proposed in this paper by the experimental evidence of the Pima Indians Diabetes Data and two sets of evaluation index.

In Section 2, we introduce the LASSO problem and improved-LARS algorithm briefly, including theory and definition. In Section 3 we put forward the LARS algorithm based on multidimensional weighting model, which calculates the direction and variables based on the weighting variables and accelerates the approximation process in promising direction. We introduce the data sets and evaluation indicators when we verify algorithm and discuss the experimental results in Section 4. Section 5 is the summary and prospect of this paper.

## 2. LASSO Problem and Improved-LARS Algorithm

*2.1. The Definition of LASSO.* Suppose that there are the multidimensional variables  $X_j \in R^n$ ,  $j = 1, 2, \dots, m$ , and response  $Y \in R^n$ . Each group of  $X_j$  has a corresponding  $y_j$ . Regression coefficient  $\beta$  is estimated where  $\|\beta\|_{l_1} \leq t$  when the sum of squared residuals is minimal. The LASSO linear regression model is defined by

$$y = X\beta + e. \quad (2)$$

$\beta$  is  $j$ -dimensional column vector, the parameter to be estimated. Error vector  $e$  meets  $E(e) = 0$  and  $\text{Var}(e) = \sigma^2$ . Suppose sparse model  $E(y | x) = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j$ ; most of regression coefficients  $\beta$  are 0 in  $\beta_1, \beta_2, \dots, \beta_j$ . Based on obtaining data, variable selection can identify which coefficient is zero and estimate other nonzero parameters; it is looking for parameters to build a sparse model. The problem we need to solve in matrix is defined by

$$(\alpha, \beta) = \arg \min \|y - X\beta - \alpha\|_{l_2}, \quad \|\beta\|_{l_1} \leq t, \quad (3)$$

where  $t$  is the threshold value of the sum of regression coefficient and  $l_1$  and  $l_2$  are two types of regularization norms.

*2.2. The Improved-LARS Algorithm.* The improved-LARS algorithm can solve LASSO problem well, which is based on the Forward Selection algorithm and Forward Gradient algorithm. The improved-LARS has appropriate forward distance, lower complexity, and more relevance of information. Figure 1 shows the basic steps of algorithm.

- (i) The improved-LARS calculates the correlation between  $X_j$  and  $Y$  constantly and finds the individual  $x_K$  most correlated with the response. It takes the largest step possible in the direction of this individual, using  $x_K$  to approximate  $y_i$ .
- (ii) Until some other individual, say  $x_P$ , has the same correlation with the current residual individual,

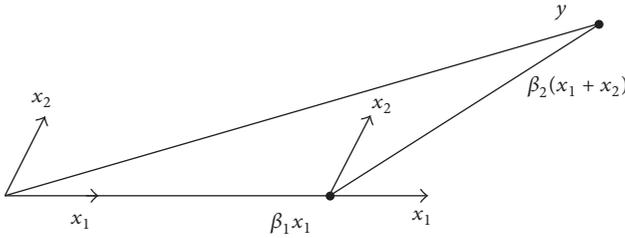


FIGURE 1: Basic steps of the LARS.

$r_{x_K y} = r_{x_P y}$ . Improved-LARS process is in an equiangular direction  $x_U$  ( $x_U$  is the direction between the two predictors  $x_P$  and  $x_K$ ).

- (iii) When a third individual  $x_T$  earns its way into the “most correlated” set, improved-LARS then proceeds equiangularly between  $x_K$ ,  $x_P$ , and  $x_T$ , that is, along the “least angle direction,” until a fourth individual enters, and so forth the direction equiangular means the bisector of each vector in high dimension.
- (iv) The LARS procedure works until the residual error is less than a threshold or all the variables are involved in the approach, the algorithm stop.

In Figure 1, an example of two-dimensional problem, it starts with both coefficients equal to zero, firstly, finding the individual more correlated with the response  $y$ ,  $r_{x_1 y} > r_{x_2 y}$ , approximating  $y$  along the  $x_1$  direction until residual of  $\beta_1 x_1$  and  $y$  has the same correlation with  $x_1$  and  $x_2$ . Then the approximating direction changes to the equiangular between  $x_1$  and  $x_2$ .

### 3. LARS Based on Multidimensional Weight

**3.1. Algorithm Analysis.** In the process of improved-LARS stepwise regression, the angle regression takes all selected variables with the same importance. However, each individual of  $X_j$  has different weight in the regression model that the indicators in its overall evaluation have different relative importance. We take the correlation between individual  $X_j$  and surplus variables into consideration, taking it and the correlation between  $X_j$  and  $y_i$  as the condition to select approximation individual.

In Figure 3, originally, we choose  $x_1$  as the first approximating variable for  $r_{x_1 y} > r_{x_2 y}$ . We take the predictor  $X_j$ 's contribution rate to the whole system as one approximating condition; the new correlation is

$$R_{X_i y} = r_{X_i y} * W_{X_j} \Big|_v^u, \quad (4)$$

where  $W_{X_j}$  is the contribution rate; we will describe the calculation method of it in detail later;  $u$  are  $v$  are constants.

Because of the addition condition, it will inevitably increase the range of values about judgment condition. In order to keep the stability of the system, we limit the product in  $[v, u]$ .

After transformation, it may indicate the possibility that  $x_j$  could have been chosen to be increasing. For example,

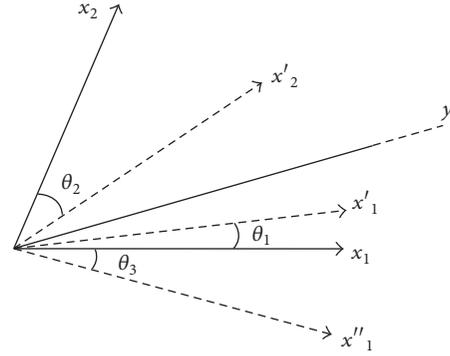


FIGURE 2: The correlation of  $X$  and  $Y$  change when adding the multidimensional weight.

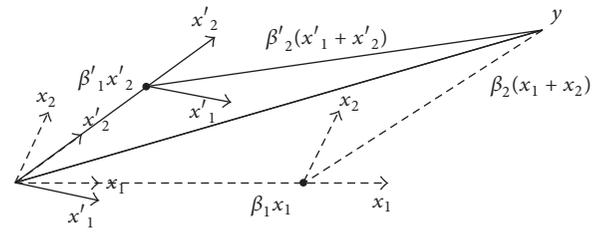


FIGURE 3: The predictor direction when adding the multidimensional weight.

Figure 2 shows that the transformed  $x_2$  gets  $\theta_2$  closer to  $y$ , reaching  $x'_2$ , and the transformed  $x_1$  gets  $\theta_1$  closer to  $y$ , reaching  $x'_1$ . It also may indicate the possibility that  $x_j$  could be selected to be reducing; the transformed  $x_1$  gets  $\theta_2$  far away to  $y$ , reaching  $x''_1$ .

On the basis of Figure 1, the new correlation will significantly change the approximation process. Figure 3 shows the predictor direction when adding the multidimensional weight: the possibility is selected of two predictors' change. It starts with  $x'_2$ ; the correlation of  $x'_1$  is the same as that of residual  $e = y - \beta'_1 x'_1$ ; when moving to  $\beta'_2 x'_2$ , the approximation direction changes to the equiangular direction of  $x'_1$  and  $x'_2$ ; then it moves forward  $\beta'_2 (x'_1 + x'_2)$ ; the approximation process is complete. Approximating path changes, and the regression variables  $\beta_1$  and  $\beta_2$  also change; we can get the new  $\beta$  calculated by improved regression method.

Applying the aforementioned process to multidimensional high-order system, the collected  $m$  feature indicators and  $n$  objects are expressed as

$$X_{N \times m} = \begin{bmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i1} & & x_{ij} & & x_{im} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Nj} & \cdots & x_{Nm} \end{bmatrix}, \quad (5)$$

$$i = 1, 2, \dots, N, \quad j = 1, 2, \dots, m$$

or

$$X = [X_1 \cdots X_j \cdots X_m], \quad j = 1, 2, \dots, m. \quad (6)$$

The collected result response is

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{bmatrix}, \quad i = 1, 2, \dots, N. \quad (7)$$

There are many calculation methods of  $W_{x_i}$  in regression process. Without loss of generality, the calculation methods should be unartificial, relative number, quantitative, and independent. We use part\_PCA, Independence Weight, and CRITIC to control regression process and verify superiority of the algorithm.

**3.1.1. Part of Principle Components Analysis.** PCA uses orthogonal transformation for dimension reduction in statistics [21, 22]. It transforms data into a new coordinate system. The biggest variance of the data projection is in the first coordinate, called the 1st principal component. The second of the data projection is in the second coordinate, called the 2nd principal component, and so on. It keeps the low-order principal component of data set, ignores the high-order principal component after transform, identifies the dominating factors, and keeps top  $m$  principal component whose overall information utilization rate is higher than 85%. Inspired by the principle of PCA, we preserve the value of all the components simultaneously and identify the variance contribution of each component; the part\_PCA algorithm steps are as follows.

*Sample Standardization*

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad i = 1, 2, \dots, N; \quad j = 1, 2, \dots, m, \quad (8)$$

where

$$\begin{aligned} \bar{x}_j &= \frac{\sum_{i=1}^n x_{ij}}{n}, \\ s_j^2 &= \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}. \end{aligned} \quad (9)$$

The correlation coefficient matrix  $R$  is information utilization of each feature

$$R = [r_{ij}]_m \quad xP = \frac{Z^T Z}{n-1}, \quad (10)$$

where

$$r_{ij} = \frac{\sum z_{kj} \cdot z_{kj}}{n-1}, \quad i, j = 1, 2, \dots, m. \quad (11)$$

**3.1.2. Independence Weight.** We sort the multiple correlation coefficient of individual by multiple regression in statistical methods [23, 24]; the greater the multiple correlation coefficient is, the more repeated the information should be, the smaller the information utilization should be, and the smaller the weight should be. Calculation steps are

$$R'_j = \frac{\sum (X_j - \bar{X})(\bar{X} - \bar{X})}{\sqrt{\sum (X_j - \bar{X})^2 \sum (\bar{X} - \bar{X})^2}}, \quad j = 1, 2, \dots, m. \quad (12)$$

$\bar{X}$  is the rest matrix in  $X$  except  $X_j$

$$\bar{X} = \text{mean}(X). \quad (13)$$

For  $R$  is negatively related with weight, we take the reciprocal of multiple correlation coefficient as score, getting the weighting coefficient through normalized processing

$$R = \left[ \frac{1}{R'_1} \cdots \frac{1}{R'_j} \cdots \frac{1}{R'_m} \right], \quad j = 1, 2, \dots, m. \quad (14)$$

**3.1.3. CRITIC.** Based on Independence Weight, CRITIC is a kind of objective weighting method proposed by Diakoulaki et al. [25]. It is based on identifying the objective weight of individual to evaluate the contrast and conflict between indicators [26, 27]. The standard deviation indicates the gap of each scheme in the same index, which can express contrast intensity. The conflict between the indicators is based on the correlation between indicators; relevancy can express conflict.

The quantitative conflict between the  $j$  indicator and other indicators is

$$\sum_{t=1}^m (1 - r_{tj}). \quad (15)$$

$r_{tj}$  is the correlation coefficient between  $X_t$  and  $X_j$

$$r_{tj} = \frac{\sum (X_t - \bar{X})(X_j - \bar{X})}{\sqrt{\sum (X_t - \bar{X})^2 \sum (X_j - \bar{X})^2}}. \quad (16)$$

$C_j$  is the amount of information the  $j$  indicator includes

$$C_j = \sigma_j \sum_{t=1}^m (1 - r_{tj}), \quad j = 1, 2, \dots, m. \quad (17)$$

The greater  $C_j$ , the greater amount of information and the more relative important of that indicator. The objective weight of  $j$  indicator is

$$R_j = \frac{C_j}{\sum_{t=1}^m C_j}, \quad j = 1, 2, \dots, m. \quad (18)$$

**3.2. Algorithm Steps.** In order to obtain the numerical solution of stability,  $X$  and  $y$  in (3) are standardized and preprocessed to omit  $\alpha$ , so that  $\sum_j x_{ij}/N = 0, |X|_{l1} = 1, \bar{Y} = 0$ .

For  $A = \{s_{j1}x_{j1}, \dots, s_{jl}x_{jl}, \dots, s_{jk}x_{jk}\} \subseteq \{1, 2, \dots, m\}$ , define the matrix

$$X_A = [s_{j1}x_{j1}, \dots, s_{jl}x_{jl}, \dots, s_{jk}x_{jk}] \in R^{n \times k}. \quad (19)$$

$X_A$  is the column vector selected from  $X$  to satisfy  $A$ , with the same direction to current  $Y$ , where the signs  $s_{jl}$  equal to

$$s_{jl} = \begin{cases} 1, & \langle x_{jl}, y \rangle > 0, \\ 0, & \langle x_{jl}, y \rangle = 0, \\ -1, & \langle x_{jl}, y \rangle < 0. \end{cases} \quad (20)$$

The equiangular  $u_A$  is as the following:

$$\begin{aligned} G_A &= X_A^T X_A, \\ A_A &= (1_A^T G_A^{-1} 1_A)^{-1/2}, \\ w_A &= A_A G_A^{-1} 1_A, \\ u_A &= X_A w_A. \end{aligned} \quad (21)$$

$u_A$  is the unit vector making equal angles, less than  $90^\circ$ , with the columns of  $X_A$ .  $1_A$  is a vector of 1's of length equaling  $|A|$ , the size of  $A$ .  $w_A$  is the equiangular contribution of each attribute  $X_i$  which is selected in  $X_A$ .  $w_A$  is processed by weighting to change approach direction and approach variable selection.

We now can further describe the improved-LARS based on multidimensional algorithm; we begin at  $u_0 = 0$  and build up new  $u$  by steps. Suppose that  $\hat{u}_A$  is the current computation and  $c$  is the current correlation between predictor and response vector

$$\begin{aligned} c &= X^T (y - \hat{u}_A) \\ \text{or } c_j &= \langle x_j, y - \hat{u}_A \rangle. \end{aligned} \quad (22)$$

The *active set*  $A$  is the set of indices corresponding to covariates with the greatest absolute current correlations, when  $= \max_j \{|c_j|\}$ ;  $u_A$  corresponding to  $A$  is the approximate direction. Let

$$\begin{aligned} s_j &= \text{sign}\{C\}, \quad j \in A, \\ a &= X^T u_A \\ \text{or } a_j &= \langle x_j, u_A \rangle. \end{aligned} \quad (23)$$

The length of approach along the  $u_A$  direction now is

$$\gamma = \min_{j \notin A}^+ \left\{ \frac{C - c_j}{A_A - a_j}, \frac{C + c_j}{A_A + a_j} \right\}. \quad (24)$$

" $\min^+$ " indicates the calculation of the minimum of positive components within each choice of  $j$  in this approximating process. Each predictor in  $A$  corresponds to increase in  $\gamma w_A$ ; we add weight to control approach direction.

Part\_PCA, Independent Weight evaluation, and CRITIC are added in the process of improved-LARS algorithm, respectively. Three parallel algorithms are established to calculate the weight of each attribute.

$$R_{j(\text{part\_PCA})} = \text{part\_PCA}(X(1 : \text{row}(w), :)), \quad (25)$$

$$R_{j(\text{IW})} = \text{IW}(X(1 : \text{row}(w), :)), \quad (26)$$

$$R_{j(\text{CRITIC})} = \text{CRITIC}(X(1 : \text{row}(w), :)). \quad (27)$$

The centralized weight  $R_j$  indicates the impact of weight on the approach direction after the aforementioned three methods.

$$\begin{aligned} R_j &= \text{centralization}(R_{j(\text{part\_PCA}/\text{IW}/\text{CRITIC})}), \\ w' &= w \times R_j, \\ \beta_A &= \beta_A + \gamma w'. \end{aligned} \quad (28)$$

$R_j$  is the same-dimension weight matrix from  $R_{j(*)}$ ; the approach direction estimate is as the following:

$$\tilde{u}_{A+} = \hat{u}_A + \gamma u_A R_j. \quad (29)$$

Then the new *active set* is

$$A_+ = A \cup \{j'\}. \quad (30)$$

$j'$  is the minimizing index in (25) of  $j$ . In order to conform to the requirements of the LASSO solution that the track should keep the same direction with current approach direction, the step size of the first opposite sign is

$$\hat{\gamma} = \min_{j \notin A}^+ \left\{ -\frac{\beta_j}{w_j} \right\}. \quad (31)$$

When  $\hat{\gamma} < \gamma$ , there is opposite sign, removing  $j'$  from  $A$ . Then the algorithm enters the next approximation process. When using  $A_+$  instead of  $A$ , repeat the above steps, until the residual error is less than a threshold or all the variables are involved in the approach. The pseudocode is shown in Pseudocode 1.

This improved algorithm increases the calculation steps for adding the weighting analysis, so the calculation time increases. But the approach mechanism of each variable in the statistical model stays the same, so the space complexity is consistent with the original algorithm.

## 4. Experiment and Result Analysis

**4.1. Introducing Data Set.** For the characteristics of the compression of the regression coefficient, the experiment set should be sparse, as well as one dependent variable which is easy to distinguish. We take Pima Indians Diabetes Data Set provided by Applied Physics Laboratory of the Johns Hopkins University, for example [28]. The data record 768 performance descriptions, negative and positive diabetes sample, including 8 attributes variables and one classification value.

```

Input. Response  $y$  dependent on predictor  $X$ , error tolerance  $\varepsilon$ 
Output. LASSO solution  $\beta$  for (3)
(1) begin
(2) data preprocessing: Normalization ( $X, y$ )
(3) initialization:  $u = \mathbf{0}$ ,  $\hat{y} = y - u$ 
(4)  $c = X^T \hat{y}$ 
(5)  $C = \max_j \{|c_j|\}$ 
(6)  $\hat{j} = \arg \max_j \{|c_j|\}$ ,  $A = \{\hat{j}\}$ 
(7) calculating weight by  $R_{\text{weight}} = \text{part\_PCA}(X)$  or  $R_{\text{weight}} = \text{IW}(X)$  or  $R_{\text{weight}} = \text{CRITIC}(X)$ 
(8) centralization  $R_{\text{weight}}$ ;
(9) when  $\|\hat{y}\|_{l_2} < \varepsilon$  and  $|A| \leq m$  do  $\beta$  cycle;
(10) in LARS cycle
      rate =  $R_{\text{weight}}(1 : \text{row}(w), \cdot)$ 
       $w' = w * \text{rate}$ 
(11) end cycle;
(12) return  $\beta$ .

```

PSEUDOCODE 1

In the classification value, “1” represents that the diabetes test result is positive; “0” represents negative. We verify algorithm by predictor of 8 attributes variables and response of one classification value. The goal of this test is improving on veracity comparing to original LARS algorithm.

**4.2. Verification Condition.** We use ROC curve to show results in order to evaluate more intuitively the performance of the proposed method. That is a binary classification problem whether the participants’ diabetes is positive or negative; the testing results have the following four types:

TP (true positive), the testing results are positive and are positive actually

FP (false positive), the testing results are positive but are negative actually

TN (true negative), the testing results are negative and are negative actually

FN (false negative), the testing results are negative but are positive actually

We take the following three characters as the inspection standards through basic four types of statistics by ROC space:

ACC (accuracy),  $ACC = (TP + TN)/(P + N)$

TPR (true positive rate),  $TPR = TP/P = TP/(TP + FN)$

NPV (negative predictive value),  $NPV = TN/(TN + FN)$

NPV is the proportion of correct detecting about negative; it means that the people who tested negative actually are negative. ACC represents the proportion of correct estimating in the sum of positive and negative. NPV is the proportion of people who tested negative in actually negative population. Compared to NPV, TPR is also called Hit Rate; it is the proportion of correctly detecting the people who

actually are positive in the tested positive population. ACC, TPR, and NPV tell us the result is better or worse than LARS.

Another character we judge the result with is SSR. The smooth turning point of SSR is corresponding to the optimal regression coefficient of predictor variable.

**4.3. Experimental Result.** Threshold  $t$  starting from 0 increases to 1 for the step length 0.01; we draw the changing curve of ACC, TPR, and NPV, with the negative and positive as dependent variables and the 8 attribute variables as the independent variables.

Figure 4 shows the accuracy and the comprehensive optimal value of three inspection standards after each cycle of LASSO with Pima Indians Diabetes.

Figure 4 shows that NPV are all improved when adding the weighting to LASSO’s solution; NPV is improved 5.16% when adding part\_PCA, 5.58% when adding Independence Weight, and 5.1% when adding CRITIC. We can find that TPR is improved 13% when adding part\_PCA, 14% when adding Independence Weight, and 13% when adding CRITIC for those methods changing the approach direction of algorithm. It is observed that ACC is improved 0.32% when adding part\_PCA and Independence Weight. And there is no effect to ACC when adding CRITIC so the ACC keeps the original number.

Therefore, the solutions of LASSO when adding the Independence Weight are optimal, followed by the CRITIC and part\_PCA. This improved algorithm significantly increases the NPV and TPR ensuring the ACC is not reduced. The veracity of LASSO’s solutions is improved through changing approach direction when adding the weighting.

Figure 5 shows the SSR (Sum of Squared Residuals) of response and equiangular direction after each cycle of LASSO with Pima Indians Diabetes. It is obvious that the general trend of SSR completely remains consistent and the residual of optimal coefficient almost remains consistent when adding the different weight judgment. It will not change the advantages of LASSO when adding the weighting. If the

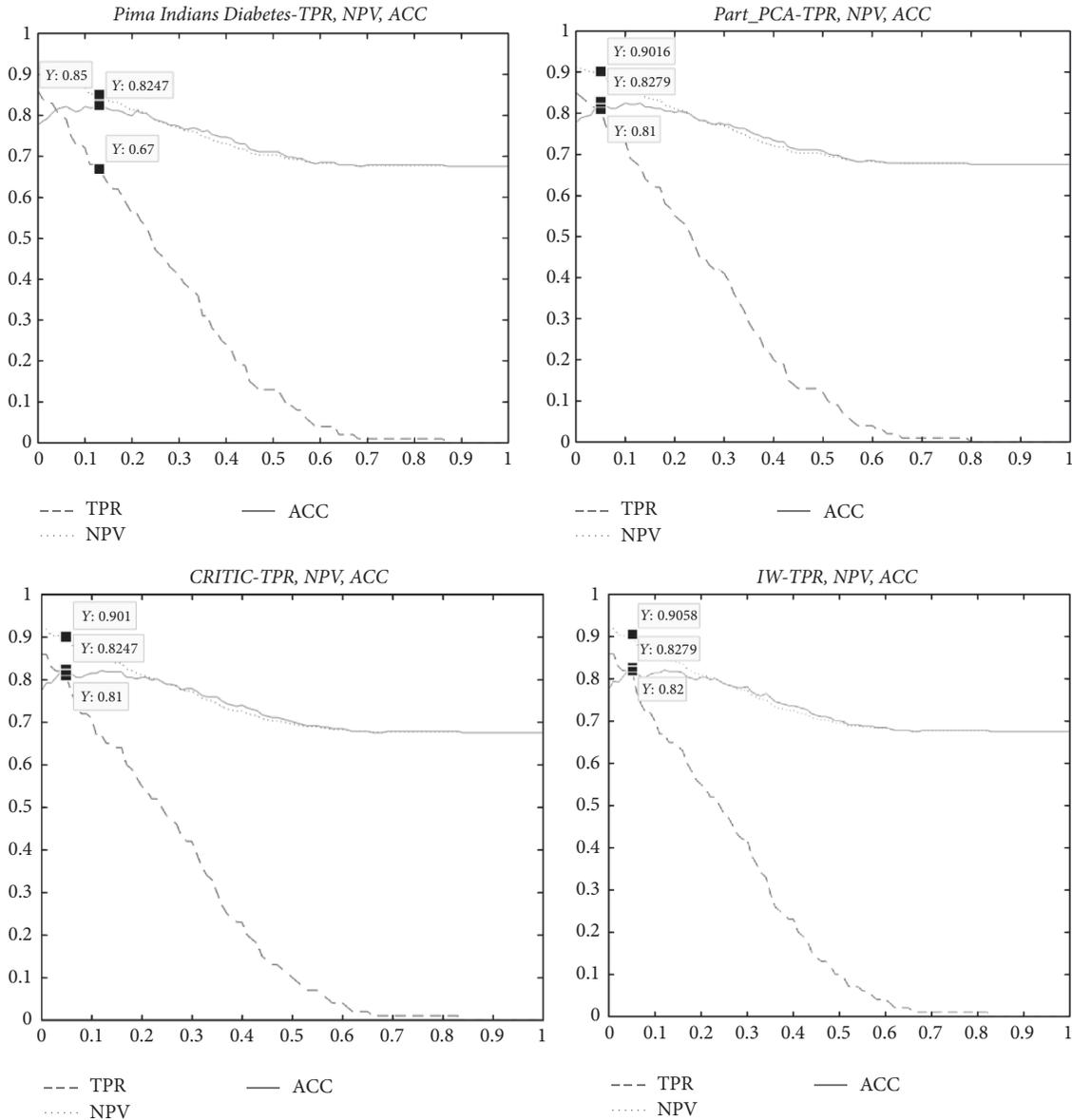


FIGURE 4: Curves about three kinds of inspection standard.

optimal regression coefficient of predictor increases significantly, the SSR will change too. It increases 0.051 residual of LASSO solution when adding part\_PCA, reduces 0.021 residual when adding CRITIC and Independence Weight. The results show that the final regression result is closer to the real response when adding CRITIC and Independence Weight.

It can be found when synthesizing these three inspection standards that, adding the approach with multidimensional weight, the threshold value of the optimal solution is mainly reduced (except CRITIC), which means that the sum of absolute values of system's regression coefficient is less than a smaller threshold; this algorithm meets the requirements in more extreme threshold range.

Table 1 shows the original regression coefficient  $\beta$ ,  $\beta_{part\_PCA}$  which adds part\_PCA,  $\beta_{IW}$  which adds Independence Weight, and  $\beta_{CRITIC}$  which adds CRITIC.

Table 2 shows the difference and innovation of the multidimensional weight LARS, improved-LARS associated with part\_PCA; Independence Weight and CRITIC act on  $\beta$ , changing the regression track and getting more accurate results.

### 5. Conclusions

In this paper, a method considering the variables choosing and the approach direction of LARS algorithm is used to solve LASSO; we propose the LARS algorithm based on multidimensional weight to improve the veracity of LASSO's

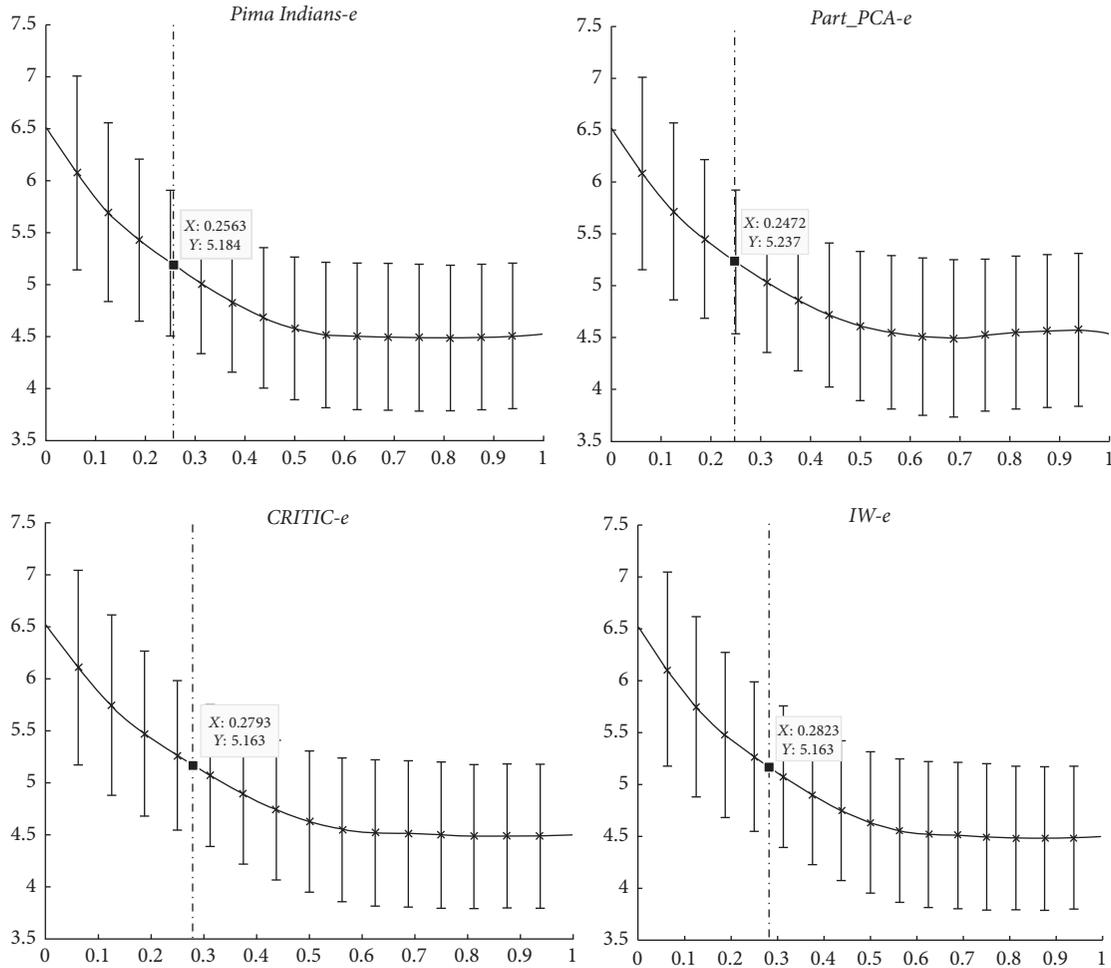


FIGURE 5: The curve about the SSR of response and equiangular direction.

TABLE 1: Regression coefficient.

	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$
B	1.7211	3.5409	-0.4862	-0.3665	-0.6504	2.2141	0.7071	-0.1810
Bpart_PCA	1.5300	3.5311	-0.4269	-0.3018	-0.6181	2.0409	0.6727	-0.0567
$\beta_{IW}$	1.5050	3.5294	-0.4253	-0.2863	-0.6072	2.0066	0.6651	-0.0356
$\beta_{CRITIC}$	1.6449	3.4482	-0.5019	-0.3705	-0.6302	2.1668	0.5471	-0.1597

TABLE 2: Principles of  $w$  in different algorithm.

Comparison of $w$	
Improved-LARS	$w = w_A = A_A G_A^{-1} 1_A$
Multidimensional weight LARS	
part_PCA	$w = w * \text{centralization}(\text{part\_PCA}(X))$
Independence Weight	$w = w * \text{centralization}(IW(X))$
CRITIC	$w = w * \text{centralization}(CRITIC(X))$

solutions and keep the advantage of LASSO’s parameter estimation, which has stable regression coefficient, reduces the number of parameters, and has good consistency of parameter convergence. We verify the efficiency of the algorithm with Pima Indians Diabetes Data Set. The precision of the calculated weight was flawed for the greater dimension of individual, so we need to further optimize the embedding weight algorithm in the later studies, to improve the accuracy and precision of regression algorithm in approach variable and direction choosing which is changed by weighting.

## Conflicts of Interest

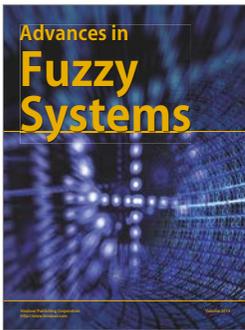
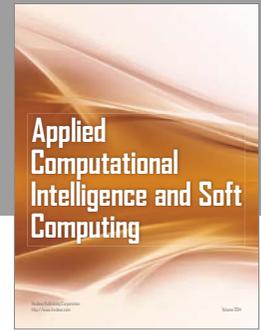
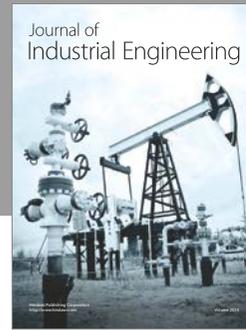
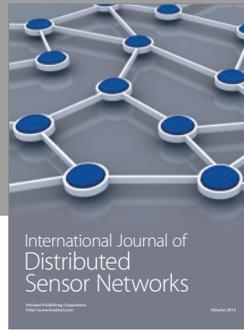
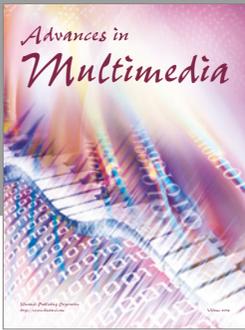
The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by the National Natural Science Foundation of China (61170192, 41271292), Chinese Postdoctoral Science Foundation (2015M580765), the Fundamental Research Funds for the Central Universities (XDJK2014C039, XDJK2016C045), Doctoral Fund of Southwestern University (swu114033), and the Project of Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ1403106).

## References

- [1] A. John and J. A. Rice, *Mathematical Statistics and Data Analysis*, Mathematical statistical physics Elsevier, 2006.
- [2] E. P. G. Box, S. J. Hunter, and W. G. Hunter, "Statistics for experimenters: an introduction to design, data analysis, and model building," *Statistics for Experimenters an Introduction to Design*, vol. 73, no. 10, article S229, 2014.
- [3] Z. L. Ke, *The Application of LASSO and Other Related Methods in Multiple Linear Regression Modle*, Jiaotong University, Beijing, China, 2011.
- [4] <http://people.duke.edu/~rnau/regintro.htm#assumpt>.
- [5] S. A. van de Geer, "High-dimensional generalized linear models and the lasso," *The Annals of Statistics*, vol. 36, no. 2, pp. 614–645, 2008.
- [6] L. E. Frank and J. H. Friedman, "A statistical view of some chemometrics regression tools," *Technometrics*, vol. 35, no. 2, pp. 109–135, 1993.
- [7] L. Breiman, "Better subset regression using the nonnegative garrote," *Technometrics*, vol. 37, no. 4, pp. 373–384, 1995.
- [8] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *Journal of the Royal Statistical Society*, vol. 58, no. 1, pp. 267–288, 1996.
- [9] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," Tech. Rep., Stanford University, Stanford, Calif, USA, 2002.
- [10] B. Efron and R. Tibshirani, "Least angle regression," *Mathematics*, vol. 32, no. 2, pp. 407–499, 2004.
- [11] H. Zou, "The adaptive lasso and its oracle properties," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1418–1429, 2006.
- [12] S. S. Keerthi and S. Shevade, "A fast tracking algorithm for generalized LARS/LASSO," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1826–1830, 2007.
- [13] C. Charbonnier, J. Chiquet, and C. Ambroise, "Weighted-LASSO for structured network inference from time course data," *Statistical Applications in Genetics & Molecular Biology*, vol. 9, no. 1, article 15, 2010.
- [14] K. M. Jung, "Weighted least absolute deviation LASSO estimator," *Communications of the Korean Statistical Society*, vol. 18, no. 6, pp. 733–739, 2011.
- [15] C. L. Bergersen, K. I. Glad, and H. Lyng, "Weighted LASSO with data integration," *Statistical Applications in Genetics & Molecular Biology*, vol. 10, no. 1, pp. 1–29, 2011.
- [16] O. Arslan, "Weighted LAD-LASSO method for robust parameter estimation and variable selection in regression," *Computational Statistics & Data Analysis*, vol. 56, no. 6, pp. 1952–1965, 2012.
- [17] S. Salzo and S. Villa, "Inexact and accelerated proximal point algorithms," *Journal of Convex Analysis*, vol. 19, no. 4, pp. 1167–1192, 2012.
- [18] Q. Zhou, S. Song, C. Wu, and G. Huang, "Kernelized LARS-LASSO for constructing radial basis function neural networks," *Neural Computing and Applications*, vol. 23, no. 7-8, pp. 1969–1976, 2013.
- [19] Y. Zhao, H. Chen, and R. T. Ogden, "Wavelet-based weighted LASSO and screening approaches in functional linear regression," *Journal of Computational & Graphical Statistics*, vol. 24, no. 3, 2015.
- [20] A. A. Salama, M. O. Ahmad, and M. N. Swamy, "Under-determined DOA estimation using MVDR-weighted LASSO," *Sensors*, vol. 16, no. 9, p. 1549, 2016.
- [21] I. T. Jolliffe, *Principal Component Analysis*, vol. 87, Springer, Berlin, Germany, 1986.
- [22] P. Zhang, *Research of Comprehensive Evaluation Based on Principal Component Analysis*, Nanjing University of Science and Technology, 2004.
- [23] H. Cai and W. L. Shen, *The Weight of Comprehensive Benefit Evaluation in Hospital (2)—Independence Weight*, Chinese Hospital Statistics, 1997.
- [24] J. He, E. S. Gao, and L. Chaohua, "The study of the weight coefficient and standardized method of the comprehensive evaluation," *Journal of Public Health in China*, vol. 17, no. 11, pp. 1048–1050, 2001.
- [25] D. Diakoulaki, G. Mavortas, and L. Papayanakis, "Determining objective weights in multiple criteria problem: the CRITIC method," *Computer & Operation Research*, no. 22, pp. 763–770, 1995.
- [26] U. E. Choo, B. Schoner, and W. C. Wedley, "Interpretation of criteria weights in multicriteria decision making," *Computers and Industrial Engineering*, vol. 37, no. 3, pp. 527–541, 1999.
- [27] B. Srdjevic, Y. D. P. Medeiros, and A. S. Faria, "An objective multi-criteria evaluation of water management scenarios," *Water Resources Management*, vol. 18, no. 1, pp. 35–54, 2004.
- [28] V. Sigillito, *UCI Machine Learning Repository*, The Johns Hopkins University, Applied Physics Laboratory, School of Information and Computer Science, 1990.



# Hindawi

Submit your manuscripts at  
<https://www.hindawi.com>

