Bayesian Adaptive Lasso binary regression with ridge parameter

Haider Kadhim Abbas Hilali and Rahim Alhamzawi

Department of Statistics, College of Administration and Economics, University of Al-Qadisiyah, Iraq

Email: haiderhilali@gmail.com

Abstract. The variable selection (VS) characteristic was considered very important in the data analysis. Regularization technique is one gorgeous way that has proven effective for dealing with high dimensional data. In previous years, statistical researchers have made great efforts in developing procedures of regularization to solved problems of VS. In this paper, we have proposed a new technique for model selection in Binary regression. This technique is Bayesian adaptive Lasso Binary regression (BALBR). It has many features that give optimum estimation and VS property. Specifically, we introduced a new hierarchical model. Then, a new Gibbs sampler method is introduced. We also extend the new approach by adding the ridge parameter inside the variance-covariance matrix to avoid the singularity in case of multicollinearity or in case the number of observations less than the number of predictors. A comparison was made with other previous techniques applying the simulation examples and real data. It is worth mentioning, that the obtained results were promising and encouraging, giving better results compared to the previous methods.

Keywords: Bayesian Lasso, binary regression, Ridge parameter, Variable selection (VS).

1. Introduction:

Linear statistical models are widely used in biological, agricultural and social sciences, as well as in economics and engineering. They are useful in both the planning stages of research and analysis of the data results. We are aware of that regression analysis is one of the statistical procedures that illustrate the relationships between explanatory variables and outcome. For each type of data, there is a suitable regression process, if the binary data is available for the dependent outcome variable and the assumptions of the form can be dealt with using the logistic regression procedure. The application of the conventional regression with this type of data will lead to biased parameters on the one hand, and inconsistent on the other hand. Therefore, it is necessary to determine a regression process that is proportional to this data. So, the Binary regression process is appropriate to those data, it is elucidating among the relationship between
the outcome variable (representing zero or one values) and the independent explanatory variables. The general formula of Binary regression is

\[
y_l = \begin{cases} 
1 & y_l^* \geq 0 \\
0 & \text{otherwise}
\end{cases}.
\] (1)

where \(y_l\) is a response outcome variable, and \(y_l^*\) is latent variable. Then \(y_l^*\) is given by

\[
y_l^* = \hat{x}_l \beta + u_l.
\] (2)

where \(y_l^*\) is \(n \times 1\) vector of responses, \(x_l\) is the \(n \times k\) matrix of predictors, and \(u_l\) is \(n \times 1\) vector of error and distributed normal \((0, \sigma^2)\), and \(k\) is a number of independent variables, and \(n\) is the number of observations.

If the explanatory variables values are known, then the regression analysis assistance us to predict the values of the outcome variable. However, when the independent variables are too large or low ratio the number of observations to the number of variables, then the selection of important and influential variables in the model is a very difficult and very important problem, as it is an important objective for many types of statistical modeling, so actual applications with this problem may not have a real model representing the actual variables. In addition, in order to obtain accurate results from specific evidence, the selected form must match the available data as best as possible, and the proposed regression of the data in question will result in results close to reality. However, to get rid of this problem, statistical researchers resorted to the mechanism of selecting important and influential variables, while eliminating as much as possible non-important explanatory variables, this procedure is the variable selection known as (VS). Although all VS procedures have evolved in recent years in linear regression, these procedures can be improved. One of these VS processes is Lasso procedure (Tibshirani, 1996) for parameter estimation. Specifically, Tibshirani (1996) introduced Lasso procedure in order to interpretability of regression models, and get better prediction accuracy. The aim of the Lasso regression is to obtain a subset of the estimations that reduces the prediction error of the outcome variable, by imposing a constraint on model parameters that cause shrank the unimportant explanatory variables and reduced to zero. After that, Park and Casella (2008) explicated that the parameters of the Lasso regression can be estimated by the Bayesian pattern. Hans ((2009)) introduced a new aspect of the Bayesian treatment of Lasso regression, by using a new Gibbs sampler for Bayesian Lasso regression. And then, Mallick and Yi (2014) developed the previous procedures by using a new procedure called a new Bayesian Lasso (NBL), this proposed procedure for VS and coefficient estimation (CE) in linear regression. The findings of Mallick in his research were very good and proved their efficiency from the previous Bayesian processes used. The good results notified in Mallick process encourage us to use the new Bayesian procedure in Binary regression.

2. Bayesian adaptive Lasso binary regression (BALBR):

Statistical researchers have done many ways to develop a variety of Bayesian regularization procedures to make a variable selection in linear regression. One of these procedures is adaptive Lasso, this method has the same feature that in the Lasso regression procedure, it can shrink some coefficients to zero, giving subsequently a selection of attributes by the regularization. This procedure was proposed by Zou (2006), that the Lasso procedure suggested by R. Tibshirani (1996), by proving that his proposed method has the Oracle attributes stated in Fan and Ping (2004), which Lasso does not own. Precisely, the adaptive Lasso
technique was based on the correct model of coefficients that are not equal to zero. The coefficients estimator of adaptive Lasso is given by

$$\hat{\beta}_{\text{lasso}} = \arg\min_\beta (y - X\beta)'(y - X\beta) + \sum_{j=1}^{k} \lambda_j |\beta_j| \; \cdots \; \lambda_j \geq 0.$$  (3)

where $\hat{\beta}_{\text{lasso}}$ are the estimation coefficients by using adaptive Lasso process, and $\lambda_j$ is a positive penalty parameter.

In general, Mallick and Yi (2014) and Abbas and Thaher (2019) methods observed results show that pragmatic well compares with other regression techniques. The good results reported in Abbas and Thaher (2019) technique encourage us to suggest a new Bayesian procedure for Binary regression. Then, we proposed a $BALBR$ procedure for CE and VS.

Following Mallick & Yi (2014) and Abbas & Thaher (2019), the Laplace formulation formula is written as follows,

$$\frac{\lambda_j}{2} e^{-\lambda_j |\beta_j|} = \int_{w_j > |\beta_j|} \frac{\lambda_j}{2} e^{-w_j} dw_j.$$  (4)

In performance, the formula (4) produces the Gibbs samples more effectively and more efficiently than the formula presented by Mallick and Yi (2014).

2.1 $BALBR$ hierarchy model and Prior Distributions:

By consuming formula (2) and formula (4), the $BALBR$ hierarchy model can be formulated as follows

$$y_i = \begin{cases} 1 & y_i^* \geq 0, \\ 0 & \text{otherwise}, \end{cases}$$

$$y^* | X, \beta, \sigma^2 \sim N_0(X\beta, \sigma^2 I_r),$$

$$\beta | \lambda \sim \prod_{j=1}^{r} \text{Uniform} \left( -\frac{1}{\lambda_j}, \frac{1}{\lambda_j} \right),$$

$$w \sim \prod_{j=1}^{r} \text{Exponential} (1),$$

$$\sigma^2 \sim \frac{1}{\sigma^2},$$

$$\lambda_j \sim \text{Gamma}(g, h).$$

where $w = (w_1, \cdots, w_k)$ and $\lambda = (\lambda_1, \cdots, \lambda_k)$.
2.2 Full Conditional Posterior Distributions of BALBR:

The following full conditional distribution (FCD) of $y^*$ is given by

$$
\begin{align*}
 y^* | y = 1 & \sim N_n(X\beta, \sigma^2 I_n) \prod_{i=1}^{\text{length}(y_i)} I \{0 < y_i < \infty\}, \\
 y^* | y = 0 & \sim N_n(X\beta, \sigma^2 I_n) \prod_{i=1}^{\text{length}(y_i)} I \{-\infty < y_i < 0\}.
\end{align*}
$$

Following Abbas and Thaher (2019), we have the following FCDs for $\beta$, $w$, $\sigma^2$ and $\lambda$ as follows

$$\begin{align*}
\pi(\beta, w, \lambda, \sigma^2 | y^*, X) & \propto \pi(y^* | X, \beta, \sigma^2) \pi(\beta | \lambda), \\
\beta | y^*, X, w, \lambda & \sim N_r(\widehat{\beta}_{OLS}, (X'X)^{-1} \sigma^2) \prod_{j=1}^{r} I \left\{-\frac{w_j}{\lambda_j} < \beta_j < \frac{w_j}{\lambda_j} \right\}, \\
\pi(w | \beta, \lambda) & \propto \pi(\beta | w, \lambda) \pi(w), \\
\pi(\sigma^2 | y^*, X, \beta) & \propto \pi(y^* | X, \beta, \sigma^2) \pi(\sigma^2), \\
\sigma^2 | y^*, X, \beta & \sim \text{InvGamma} \left(\frac{n}{2} + 1, \frac{1}{2} (y^* - X\beta)'(y^* - X\beta)\right), \\
\lambda_j | \beta_j & \propto \text{Gamma}(g + 1, h) I \left\{\frac{w_j}{\lambda_j} < \lambda_j \right\},
\end{align*}
$$

where the $\widehat{\beta}_{OLS}$ is ordinary least squares estimators, and $I(\cdot)$ denotes an indicator function.

3. BALBR with ridge parameter:

In practice, the BALBR method performs very well. Following Gupta and Ibrahim (2007), we added the ridge parameter inside variance-covariance matrix to remedy actual challenges that may appear via multicollinearity and overfitting problems.

Now, referring to the equation (6) and adding ridge parameter $\tau$ to the equation

$$
\exp \left\{-\frac{1}{2\sigma^2} (-2y^* X'X + \tau I_r)^{-1} (X'X + \tau I_r) \beta + \beta' (X'X + \tau I_r) \beta) \right\} \prod_{j=1}^{r} I \left\{|\beta_j| < \frac{w_j}{\lambda_j}\right\},
$$
\[ \beta | y^*, X, w, \lambda \sim N_r(\beta_R, (X'X + \tau I_r)^{-1} \sigma^2) \cap \bigcap_{j=1}^r \left\{ \frac{w_j}{\lambda_j} < \beta_j < \frac{w_j}{\lambda_j} \right\}. \]

where the $\beta_R$ is ridge estimators, and $I(\cdot)$ denotes an indicator function.

4. **BALBR computation:**

We specify Gibbs samples for BALBR procedure by initiate with the initial valuations for parameters $\beta, z, \lambda$ and $\sigma^2$, then we carry out the algorithm as follows

**Algorithm (Sampling in BALBR).**

- **Sampling $y$:** We generate $y_i^*$ as follows
  \[ y^* | y = 1 \sim N_n(X\beta, \sigma^2 I_n) \] left truncated at 0.
  \[ y^* | y = 0 \sim N_n(X\beta, \sigma^2 I_n) \] right truncated at 0.

- **Sampling $z$:** We generate the $w_j$ as follows $w_j = w_j^* + |\lambda_j|$, where $w_j^*$ is a standard exponential distribution.

- **Sampling $\beta$:** We generate $\beta$ coefficients from truncated normal with mean $\beta_R$ and variance-covariance $(X'X + \tau I_r)^{-1} \sigma^2$.

- **Sampling $\sigma^2$:** We generate the $\sigma^2$ from the inverse gamma with shape $n/2 + 1$ and rate $\frac{1}{2} (y^* - X\beta)'(y^* - X\beta)$.

- **Sampling $\lambda$:** We generate the $\lambda_j$ from truncated gamma with shape $g + 1$ and rate $h$.

5. **Simulation analysis:**

In this section, we test the new proposed procedure and measure its performance compared to previous procedures for CE and VS. This test is carried out by applying simulation examples to BALBR with ridge parameter, Bayesian Binary regression (**BBr**) by implementing the MCMCpack package (Martin, Quinn, Park, 2018), and the Bayesian Binary quantile regression (**BBrq**) by using the bayesQR package (Benoit, Alhamzawi, 2011). All these packages will be implemented in R language. In order to take Gibbs sampling, we draw 12,000 iterations, and the first 1000 were excluded as burn-in. The process is assessed based on the best classification. We set $g = h = 0.05$, $\tau = 0.01$ and $\text{Tau} = 0.50$.

5.1 **Simulation 1:**

In this simulation, we generate 7 independent variables with observations (25, 50, 100), and without intercept. The pair-wise correlation between each independent variable equalize to $0.5 |\epsilon_1|$, and we set the $\beta^{true}$ coefficients as follows

\[ \beta^{true} = (3, 0, ..., 0)' \]

We simulated $y_i^*$ as follows

\[ y_i^* = 3x_{1i} + \epsilon_i \]

where $\epsilon_i \sim N(0, \sigma^2)$. and $\sigma \in \{1, 2, 3\}$.

We list the results of coefficients estimate as table below.
Table 1: The results of Simulation 1, when $\varepsilon_i \sim N(0,1)$

| Method  | n  | True | False |
|---------|----|------|-------|
| BALBR  | 24 | 25   | 1     |
| BBr    | 21 | 22   | 3     |
| BALBR  | 50 | 50   | 0     |
| BBq    | 49 | 49   | 1     |
| BBq    | 49 | 49   | 1     |
| BBq    | 89 | 89   | 11    |
| BBq    | 89 | 89   | 11    |

Table 2: The results of Simulation 1, when $\varepsilon_i \sim N(0,4)$

| Method  | n  | True | False |
|---------|----|------|-------|
| BALBR  | 23 | 22   | 2     |
| BBr    | 22 | 22   | 3     |
| BBq    | 48 | 47   | 2     |
| BBq    | 47 | 47   | 3     |
| BBq    | 82 | 82   | 18    |
| BBq    | 80 | 81   | 19    |
| BBq    | 80 | 80   | 20    |

Table 3: The results of Simulation 1, when $\varepsilon_i \sim N(0,9)$

| Method  | n  | True | False |
|---------|----|------|-------|
| BALBR  | 20 | 20   | 5     |
| BBr    | 20 | 20   | 5     |
| BBq    | 20 | 20   | 5     |
| BALBR  | 46 | 45   | 5     |
| BBq    | 46 | 45   | 5     |
| BBq    | 73 | 73   | 27    |
| BBq    | 72 | 72   | 28    |

5.2 Simulation 2:

We generate 8 independent variables with observations (25, 50, 100), the pair-wise correlation between each independent variable equalizes to 0.3, and we set the $\beta^{true}$ coefficients as follows

$$\beta^{true} = (1, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)'.$$

We simulated $y_i^*$ as follows
\[ y_i^* = 1 + 2x_{1i} + \varepsilon_i \]

where \( \varepsilon_i \sim N(0, \sigma^2) \), and \( \sigma \in \{1, 2, 3\} \).

We list the results of coefficients estimate as table below

### Table 4: The results of Simulation 2, when \( \varepsilon_i \sim N(0,1) \)

| Method | \( n \) | True | False |
|--------|--------|------|-------|
| BALBR  | 25     | 25   | 0     |
| BB\( r \) | 25    | 25   | 0     |
| BB\( qr \) | 50    | 50   | 0     |
| BALBR  | 50     | 48   | 2     |
| BB\( r \) | 95    | 95   | 5     |
| BB\( r \) | 100   | 93   | 7     |

### Table 5: The results of Simulation 2, when \( \varepsilon_i \sim N(0,4) \)

| Method | \( n \) | True | False |
|--------|--------|------|-------|
| BALBR  | 22     | 22   | 3     |
| BB\( r \) | 22    | 22   | 3     |
| BB\( qr \) | 43    | 43   | 7     |
| BALBR  | 42     | 43   | 7     |
| BB\( r \) | 87    | 87   | 3     |
| BB\( r \) | 100   | 86   | 4     |

### Table 6: The results of Simulation 2, when \( \varepsilon_i \sim N(0,9) \)

| Method | \( n \) | True | False |
|--------|--------|------|-------|
| BALBR  | 23     | 23   | 2     |
| BB\( r \) | 21    | 21   | 4     |
| BB\( qr \) | 22    | 22   | 3     |
| BALBR  | 43     | 43   | 7     |
| BB\( r \) | 43    | 43   | 7     |
| BB\( r \) | 77    | 77   | 23    |
| BB\( r \) | 100   | 76   | 24    |

5.3 Simulation 3:

We generate 8 independent variables with observations (25, 50, 100), and same \( \beta^{true} \) coefficients in example above. But, the pair-wise correlation between \( x_i \) and \( x_j \) is high correlation, and it is equal to 0.8.

We list the results of coefficients estimate as table below
Table 7: The results of Simulation 3, when $\epsilon_i \sim N(0,1)$

| Method    | $n$ | True | False |
|-----------|-----|------|-------|
| BALBR     | 25  | 25   | 0     |
| $BB_{r}$ | 25  | 25   | 0     |
| $BB_{qr}$| 25  | 25   | 0     |
| BALBR     | 50  | 50   | 0     |
| $BB_{r}$ | 50  | 50   | 0     |
| $BB_{qr}$| 50  | 50   | 0     |
| BALBR     | 95  | 95   | 5     |
| $BB_{r}$ | 100 | 94   | 6     |
| $BB_{qr}$| 94  | 94   | 6     |

Table 8: The results of Simulation 3, when $\epsilon_i \sim N(0,4)$

| Method    | $n$ | True | False |
|-----------|-----|------|-------|
| BALBR     | 25  | 22   | 3     |
| $BB_{r}$ | 25  | 22   | 3     |
| $BB_{qr}$| 22  | 22   | 3     |
| BALBR     | 46  | 46   | 4     |
| $BB_{r}$ | 50  | 45   | 5     |
| $BB_{qr}$| 45  | 45   | 5     |
| BALBR     | 86  | 86   | 14    |
| $BB_{r}$ | 100 | 85   | 15    |
| $BB_{qr}$| 85  | 85   | 15    |

Table 9: The results of Simulation 3, when $\epsilon_i \sim N(0,9)$

| Method    | $n$ | True | False |
|-----------|-----|------|-------|
| BALBR     | 22  | 22   | 3     |
| $BB_{r}$ | 25  | 21   | 4     |
| $BB_{qr}$| 21  | 21   | 4     |
| BALBR     | 40  | 40   | 10    |
| $BB_{r}$ | 50  | 39   | 11    |
| $BB_{qr}$| 40  | 40   | 10    |
| BALBR     | 82  | 82   | 18    |
| $BB_{r}$ | 100 | 81   | 19    |
| $BB_{qr}$| 82  | 82   | 18    |

5.4 Simulation 4:
In this simulation, we generate 7 independent variables with observations (25, 50, 100), the pair-wise correlation between each independent variable equalizes to $0.5 |l-j|$, and we set the $\beta^{true}$ coefficients as follows

$$\beta^{true} = (1, 1, 0, 0, 1, 1, 0, 0)^T.$$

We simulated $y_i^*$ as follows

$$y_i^* = 1 + x_{4i} + x_{5i} + x_{5i} + \epsilon_i$$

where $\epsilon_i \sim N(0, \sigma^2)$, and $\sigma \in \{1, 2, 3\}$.

We list the results of coefficients estimate as table below
Table 10: The results of Simulation 4, when $\varepsilon_i \sim N(0,1)$

| Method    | n  | True | False |
|-----------|----|------|-------|
| BALBR     | 25 | 25   | 0     |
| BBr       | 25 | 25   | 0     |
| BBqr      | 25 | 0    | 0     |
| BALBR     | 49 | 1    | 0     |
| BBr       | 50 | 48   | 2     |
| BBqr      | 48 | 2    | 0     |
| BALBR     | 96 | 4    | 0     |
| BBr       | 100| 94   | 6     |
| BBqr      | 95 | 5    | 0     |

Table 11: The results of Simulation 4, when $\varepsilon_i \sim N(0,4)$

| Method    | n  | True | False |
|-----------|----|------|-------|
| BALBR     | 25 | 25   | 0     |
| BBr       | 25 | 25   | 0     |
| BBqr      | 25 | 0    | 0     |
| BALBR     | 43 | 7    | 0     |
| BBr       | 50 | 42   | 8     |
| BBqr      | 42 | 8    | 0     |
| BALBR     | 81 | 19   | 0     |
| BBr       | 100| 78   | 22    |
| BBqr      | 79 | 21   | 0     |

Table 12: The results of Simulation 4, when $\varepsilon_i \sim N(0,9)$

| Method    | n  | True | False |
|-----------|----|------|-------|
| BALBR     | 25 | 25   | 0     |
| BBr       | 25 | 25   | 0     |
| BBqr      | 23 | 2    | 0     |
| BALBR     | 44 | 6    | 0     |
| BBr       | 50 | 43   | 7     |
| BBqr      | 43 | 7    | 0     |
| BALBR     | 77 | 23   | 0     |
| BBr       | 100| 74   | 26    |
| BBqr      | 75 | 25   | 0     |

5.5 Simulation 5:
In this simulation, we generate 6 independent variables with observations (25, 50, 100), the pair-wise correlation between each independent variable equals to -0.25, and we set the $\beta^{true}$ coefficients as follows

$$\beta^{true} = (1, 1.5, 0, 1.5, 0, 1.5, 0)' .$$

We simulated $y_i^*$ as follows

$$y_i^* = 1 + 1.5x_{1i} + 1.5x_{3i} + 1.5x_{5i} + \varepsilon_i$$

where $\varepsilon_i \sim N(0, \sigma^2)$, and $\sigma \in \{1, 2, 3\}$. We list the results of coefficients estimate as table below
Table 13: The results of Simulation 5, when $\varepsilon_i \sim N(0,1)$

| Method     | $n$ | True | False |
|------------|-----|------|-------|
| BALBR      | 25  | 25   | 0     |
| BB$\theta$ | 25  | 25   | 0     |
| BB$q\theta$| 25  | 25   | 0     |
| BALBR      | 50  | 50   | 0     |
| BB$\theta$ | 50  | 50   | 0     |
| BB$q\theta$| 49  | 49   | 1     |
| BALBR      | 96  | 96   | 4     |
| BB$\theta$ | 100 | 95   | 5     |
| BB$q\theta$| 95  | 95   | 5     |

Table 14: The results of Simulation 5, when $\varepsilon_i \sim N(0,4)$

| Method     | $n$ | True | False |
|------------|-----|------|-------|
| BALBR      | 24  | 24   | 1     |
| BB$\theta$ | 22  | 22   | 3     |
| BB$q\theta$| 23  | 23   | 2     |
| BALBR      | 44  | 43   | 6     |
| BB$\theta$ | 50  | 43   | 7     |
| BB$q\theta$| 43  | 43   | 7     |
| BALBR      | 85  | 83   | 15    |
| BB$\theta$ | 100 | 83   | 17    |
| BB$q\theta$| 83  | 83   | 17    |

Table 15: The results of Simulation 5, when $\varepsilon_i \sim N(0,9)$

| Method     | $n$ | True | False |
|------------|-----|------|-------|
| BALBR      | 23  | 23   | 2     |
| BB$\theta$ | 22  | 22   | 3     |
| BB$q\theta$| 22  | 22   | 3     |
| BALBR      | 45  | 45   | 5     |
| BB$\theta$ | 50  | 41   | 9     |
| BB$q\theta$| 43  | 43   | 7     |
| BALBR      | 81  | 81   | 19    |
| BB$\theta$ | 100 | 75   | 25    |
| BB$q\theta$| 79  | 79   | 21    |

From all examples of previously applied simulations, we can see that the BALBR method is better than the other methods (BB$\theta$ and BB$q\theta$) used in comparisons in terms of classification. The BALBR method has given the best classification most often. This demonstrates the quality of the performance of BALBR procedure in terms of CE and VS.

6. Real Data:

After we have demonstrated that the simulation study exceeds our proposed method, we will apply that method to the real data and then analyze the results statistically. Agricultural researchers have been interested in increasing wheat production with several years for the importance of this crop economically. Therefore, determining the main factors for increasing this crop is very important and helps researchers
predict an increase the wheat production in the future. Hence the importance of the proposed new method, which tries to identify some important variables and show their effect on the rate of increase in production of this crop. In this section of the research we will make a comparison between the proposed method and the previous methods in real data to produce the wheat crop. These data include 11 variables within 500 observations. The response variable in these data represents an increase in wheat production or not, and the explanatory variables will be mentioned in the table below. These data were taken from the Diwaniyah Directorate of Agriculture in Iraq for the year 2017.

**Table 16: The wheat production data variables**

| symbol | Variables description | Rank | Rank description |
|--------|----------------------|------|-------------------|
| Y      | (Increase of wheat product) | 0    | Non-increased productivity |
|        |                      | 1    | Increased         |
| U      | (Urea fertilizer)    | Numeral | Quantity in kilogram |
|        |                      | 1    | Ideal             |
| Ds     | (Sowing date)        | 2    | Early             |
|        |                      | 3    | Late              |
| Qs     | (Sowing seeds quantity) | Numeral | Quantity in kilogram |
| LT     | (Laser field leveling) | 1    | Unused            |
|        |                      | 2    | Used              |
| NPK    | (Compound fertilizer) | Numeral | Quantity in kilogram |
| SMT    | (Sowing machine technique) | 1    | Unused            |
|        |                      | 2    | Used              |
| SC     | (Planting successive corps) | 1    | Planting         |
|        |                      | 2    | Not planting      |
| H      | (Herbicide)          | Numeral | Quantity in milliliter |
| K      | (Potassium fertilizer) | Numeral | Quantity in kilogram |
| ME     | (Microelements fertilizer) | Numeral | Quantity in gram |

The results shown in Table 17 show that the classification in the proposed method BALBR is much better than the other methods used.

**Table 17: The real data classification results**

| Method | True | False |
|--------|------|-------|
| BALBR  | 537  | 47    |
| BB r   | 535  | 49    |
| BB qr  | 536  | 48    |
Figure 1- Histograms of BALBR coefficients estimation
Figure 2- Trace plots of BLTR coefficients estimation
The BALBR procedure coefficients estimation are based on posteriors samples of 12,000 recurrences. In figure 1, which illustrates the histograms of BALBR coefficients estimation, these histograms displayed that the conditional posteriors of this procedure are stationary for its underlying truncated normal distribution. In figure 2, displayed the trace plots of BALBR coefficients estimation, these plots show a reasonably very good approximation, and the noise has been significantly deviated and the chain has reached stability and the center remains relatively constant. This means that the chain is fully mixed and convergent. In figure 3, the plots displayed the autocorrelations of BALBR coefficients estimation, the 10 covariates in these data are highly correlated, and these MCMC chains in BALBR process was practically well.
Conclusions:

We have presented in this research a new technique for model selection of binary regression, where we proposed BALBR method to estimate the coefficients with VS process. We advanced a new Bayesian hierarchical model for BALBR. Moreover, we have provided the Gibbs samples for this procedure. The extension of our procedures has been included in our research, where the ridge parameter is added within the variance-covariance matrix to prevent the singularity in case of multicollinearity and overfitting problems. We demonstrated the advantage of the new procedure in both simulations and analysis of real data. The results showed that our procedure performed very good in terms of VS and PE. In particular, the BALBR procedure is absolutely the best of all the procedures mentioned above. Through the conclusion of this research, statisticians are assisted by the presence of BALBR technique in statistics, using this new technique to ensures accurate and useful results for the corrects predictions.

References:

[1] Abbas, H. K., & Thaher, R. J. (2019). Bayesian adaptive Lasso Tobit regression. *Journal of Al-Qadisiyah for computer science and mathematics, 11*(1), Page-1. Hans, C. (2009). Bayesian lasso regression. *Biometrika*, 96(4), 835-845.

[2] Alhamzawi, R. (2013). Tobit quantile regression with the adaptive lasso penalty. In *The Fourth International Arab Conference of Statistics, 450 ISSN (1681 6870)*.

[3] Alhamzawi, R., & Ali, H. T. M. (2018). Bayesian tobit quantile regression with penalty. *Communications in Statistics-Simulation and Computation, 47*(6), 1739-1750.

[4] Alhamzawi, R., & Ali, H. T. M. (2018). The Bayesian adaptive lasso regression. *Mathematical Biosciences*.

[5] Alhamzawi, R., & Taha Mohammad Ali, H. (2018). A new Gibbs sampler for Bayesian lasso. *Communications in Statistics-Simulation and Computation, 1*-17.

[6] Alhamzawi, R., Yu, K., & Benoit, D. F. (2012). Bayesian adaptive Lasso quantile regression. *Statistical Modelling, 12*(3), 279-297.

[7] Alhusseini, F. H. H. (2017). New Bayesian Lasso in Tobit Quantile Regression. *Romanian Statistical Review Supplement, 65*(6), 213-229.

[8] Alhusseini, F. H. H., & Georgescu, V. (2018). Bayesian composite Tobit quantile regression. *Journal of Applied Statistics, 45*(4), 727-739.

[9] Alshaybawee, T., Alhamzawi, R., Middi, H., & Allyas, I. I. (2018). Bayesian variable selection and coefficient estimation in heteroscedastic linear regression model. *Journal of Applied Statistics, 1*-15.

[10] Andrews, D. F., & Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B (Methodological), 99*-102.

[11] Andrew D. Martin, Kevin M. Quinn, Jong Hee Park (2011). *MCMCPack: Markov Chain Monte Carlo in R*. Journal of Statistical Software. 42(9): 1-21.

[12] Benoit, D. F., Alhamzawi, R., & Yu, K. (2013). Bayesian lasso binary quantile regression. *Computational Statistics, 28*(6), 2861-2873.

[13] Benoit, D. F., Al-Hamzawi, R., Yu, K., Van den Poel, D., & Benoit, M. D. F. (2017). Package ‘bayesQR’.

[14] Benoit, D. F., & Van den Poel, D. (2012). Binary quantile regression: a Bayesian approach based on the asymmetric Laplace distribution. *Journal of Applied Econometrics, 27*(7), 1174-1188.
[15] Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, 95(3), 759-771.

[16] Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics*, 51(1-2), 79-99.

[17] Christian Kleiber and Achim Zeileis (2008). *Applied Econometrics with R*. New York: Springer-Verlag. ISBN 978-0-387-77316-2.

[18] Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of statistics*, 32(2), 407-499.

[19] Feng, X. N., Wu, H. T., & Song, X. Y. (2017). Bayesian adaptive lasso for ordinal regression with latent variables. *Sociological Methods & Research*, 46(4), 926-953.

[20] George, E. I., & McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, 88(423), 881-889.

[21] George, E. I., & McCulloch, R. E. (1997). Approaches for Bayesian variable selection. *Statistica sinica*, 339-373.

[22] Huang, J., Ma, S., & Zhang, C. H. (2008). Adaptive Lasso for sparse high-dimensional regression models. *Statistica Sinica*, 1603-1618.

[23] Kyung, M., Gill, J., Ghosh, M., & Casella, G. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis*, 5(2), 369-411.

[24] Leng, C., Tran, M. N., & Nott, D. (2014). Bayesian adaptive lasso. *Annals of the Institute of Statistical Mathematics*, 66(2), 221-244.

[25] Mallick, H., & Yi, N. (2014). A new Bayesian lasso. *Statistics and its interface*, 7(4), 571.

[26] Park, T., & Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103(482), 681-686.

[27] Pötscher, B. M., & Schneider, U. (2009). On the distribution of the adaptive LASSO estimator. *Journal of Statistical Planning and Inference*, 139(8), 2775-2790.

[28] Sun, W., Ibrahim, J. G., & Zou, F. (2010). Genome-wide multiple loci mapping in experimental crosses by the iterative adaptive penalized regression. *Genetics*.

[29] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.

[30] Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476), 1418-1429.