A simulation Study to Evaluate the Performance of the Two Regulatory Methods (SEA - Lasso) and the (MCP) Method in the Multiple Regression Model and Selection of the Best

Ashwaq Abdul Sada Kadhim
Technical Institute of Al Diwaniyah, AI Furat Al-Awsat Technical University) (ATU), Iraq.
ashwaq.sada.idi12@atu.edu.iq . mobile 00964711509018

ABSTRACT

Variable selection is an important topic in linear regression analysis. In practice, a large number of predictors are usually introduced at the initial stage of model construction to mitigate potential model biases. On the other hand, to enhance the ability to predict and select important variables.

The statisticians made great efforts in developing regularization procedures to solve the problems of V.S. These actions automatically facilitate Variable selection (V.S) by setting specific coefficients to zero and reducing coefficient estimates, providing useful estimates even if the model contains a large number of variables. In this paper, two methods of regularization were proposed to estimate and select the appropriate variables at the same time in the multiple regression model, which are the Standard Error Adjusted Adaptive LASSO (SEA-LASSO) and Minim ax Concave Penalty (MCP) method and selection of the best. The paper problem focuses on using the best regularization method that works on estimation and appropriate selection of important variables at the same time and addressing the problem of multiple linearity using the SEA-LASSO and MCP method. To get the real model.

This paper aims to evaluate the performance of the method (SEA-LASSO) and method (MCP) in terms of the process of estimation and appropriate selection of important variables and treatment the problem multicollinearity through the simulation study.

A simulation study was conducted to compare between these two methods, which included different cases of the factors and testing the effect of the levels of these factors on the performance of these two ways, as well as determining the value of the control parameter (λ) and the criterion for selecting the best value for it and the basis on which to evaluate the performance of the two methods. The simulation results showed that (SEA-LASSO) method is superior to (MCP) method in terms of percentage of operation to reach the real model measured by (PCT), and it is also better in terms of mean squares error (MSE) because it achieves less (MSE) in most cases. A simulation study was used with the program R.

Keywords: Multiple regression model , Adjusted Adaptive LASSO (SEA-LASSO), Minimax Concave Penalty (MCP)

.

1- Introduction

In some multiple regression applications, the number of predictors have become large, which is why, the analysis of that data has become difficult. For the purpose of dealing with this problem, it is necessary to perform dimensionality shrinkage of data with a few assumptions. When talking about dimensionality reduction, it indicates the fact that there are high dimensions, those dimensions have been referred to as the variables or features. Increasing the number of these variables in the multiple regression model, means that the model will be more difficult to analysis the data. Thus, a problem will be encountered, known as the curse of dimensions where term curse of dimensions was introduced by Bellman (1961) when data is sparse in multi-dimensional spaces. Also in the case of linear correlation between the High-dimensional data (HD). In which the greater the number of the variables makes it more difficult predicting a certain quantity. Those variables might not be all influential or effective, or can be interconnected and thereby, redundant which will require reduction. Which is why, dimensionality reduction process means the conversion of the (HD) into a space of a smaller size. It has a significant impact to solve this problem and by reducing the number of the random variables, in other words, simplifying the understanding of the data only visually or numerically and thus ensuring the integrity of data. Moreover, there are other advantages to the reduction of dimensionality where it operates on data compression and reduces the time of the calculations and there are some methods which do not operate efficiently in the cases of very high dimensions (Lian, 2012). Thus, it is necessary to work on the reduction of the dimension to obtain a simplified model. Therefore, the researchers suggested many regulatory Methods that work on the estimation and selection appropriate of important variables at the same time and treatment the problem multicollinearity, among these methods are the Standard Error Adjusted Adaptive LASSO (SEA-LASSO) and Minim ax Concave Penalty (MCP).

This paper is regularized as follows: A multiple regression model and an explanation showing the SEA-LASSO method and the MCP, which both estimate and select important variables simultaneously, are presented in **Section 2**. In **Section 3**, it was specifically designed to study simulation and summarize its results. In **Section 4** A brief conclusion of this study is included.

2- Multiple Regression Model

A model that contains two or more explanatory variables is called multiple linear regression.

The multiple form has the following form:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + \epsilon_i \qquad i = 1, \ldots, n \ldots\ldots(1)$$

The coefficients, $\beta_0, \beta_1, \ldots, \beta_k$, are unknown, and $\epsilon_1, .., \epsilon_n$ are distributed according to $N(0, \sigma^2)$.

Model (2.1) can be written in matrix form as follows [4] :

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}_i \qquad (2)$$

$$Y = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y} \end{bmatrix}, \boldsymbol{X} = \begin{bmatrix} \mathbf{1} & \boldsymbol{x}_{11} & \boldsymbol{x}_{12} & \cdots & \boldsymbol{x}_{1k} \\ \mathbf{1} & \boldsymbol{x}_{21} & \boldsymbol{x}_{22} & \cdots & \boldsymbol{x}_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{1} & \boldsymbol{x}_{n1} & \boldsymbol{x}_{n2} & \cdots & \boldsymbol{x}_{nn} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_k \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \boldsymbol{\epsilon}_1 \\ \boldsymbol{\epsilon}_2 \\ \boldsymbol{\epsilon}_3 \\ \boldsymbol{\epsilon}_4 \end{bmatrix}.$$

Where:

Y : response vector Y_{n+1}.

X : design matrix $X_{n\times k}$.

β : coefficient vector $\beta_{k\times 1}$.

$\boldsymbol{\epsilon}$: the error vector $\epsilon_{n\times 1}$.

$E(\epsilon) = 0$ and $cov(\epsilon) = \sigma^2 I_n$

Least squares estimation is the well know method to estimate the un know vector of $(\boldsymbol{\beta})$.

2-1 Standard Error Adjusted Adaptive LASSO (SEA-LASSO) Method.

It is a special case of adaptive lasso, this method was proposed by Qi an and Yang (2013)which are taken into account the standard errors of the estimators (OLS) when calculating the elastic weights of the adaptive Lasso method that depends on these weights when variables selection for the purpose of calculating the elastic weights used Zou (2006) ordinary least squares estimations (OLS), but the estimator (OLS) suffers from deficiencies in light of The presence of multi-linearity, which makes the elastic weights unstable. The estimator is obtained by minimizing the following Punitive Least Squares Function(PLSF) :

$$\hat{\beta}^{SEA} = arg \min_{\beta}[\sum_{i=1}^{N}(Y_i - X\,\beta)^2 - \lambda_n \sum_{j=1}^{p} \widehat{W}_j^{SEA}|\beta_j| \quad , \qquad \lambda \geq 0 \;.... \quad (\mathbf{3})$$

Where : $\lambda_n \sum_{j=1}^{p} \widehat{W}_j^{SEA}|\beta_j|$, is called SEA-LASSO penalty function. λ_n: The tuning parameter

$\widehat{W}_j^{SEA}$: Estimated flexible weights. The elastic weights of the estimator SEA-LASSO are calculated as follows: $\widehat{W}_j^{SEA} = \left(\frac{sd_j^{SEA}}{|\hat{\beta}_j^{OLS}|}\right)^{\gamma}$

Where :

$\hat{\beta}_j^{OLS}$is the ordinary least squares estimator (OLS)for the regression parameter β_j .

sd_j^{SEA}: It is the standard error of the estimator $\hat{\beta}_j^{OLS}$.

Qi an and Yang (2013) used tuning parameter value ($\gamma = 1$) When the (SEA-LASSO) method is implemented in the current simulation . Qian and Yang (2013) has shown that by choosing a suitable parameter the estimator (λ_n) is characterized as an oracle, i.e. consistent in selecting variables and Asymptotically normal.

2-2 Minimax Concave Penalty (MCP) method.

MCP, a fast, continuous, nearly unbiased and accurate method of penalized variable selection in high-dimensional linear regression suggested by Zhang 2010 .The MCP provides the convexity of the penalized loss in sparse regions to the greatest extent given certain thresholds for variable selection and unbiasedness. (Zhang, 2010)**.** MCP solves:

$$minimize_{\beta}\ \frac{1}{2}\,\|y - x\beta\|_2^2 + P(\beta)\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

where the derivative of the penalty function is [5] :

$$P'(\beta) = \begin{cases} sgn(\beta)\left(\lambda - \frac{|\beta|}{\gamma}\right) & if\ |\beta| < \lambda\gamma \\ 0 & otherwise \end{cases} \ldots\ldots\ldots\ldots(5)$$

Where : $\gamma > 1$. The penalty function can be written explicitly:

$$p\,(|\beta|) = \begin{cases} \lambda\left(|\beta| - \frac{\beta^2}{2\gamma}\right), & if\ \ |\beta| < \lambda\gamma \\ \frac{\gamma\lambda^2}{2}, & otherwise \end{cases} \ldots\ldots\ (6)$$

where γ

: is a second positive hyperparameter that is determined by the researcher, in the simulation study it was set ($\gamma = 3$). The MCP provides the sparse convexity to the broadest extent by minimizing the maximum concavity (Wang et al. (2018)).

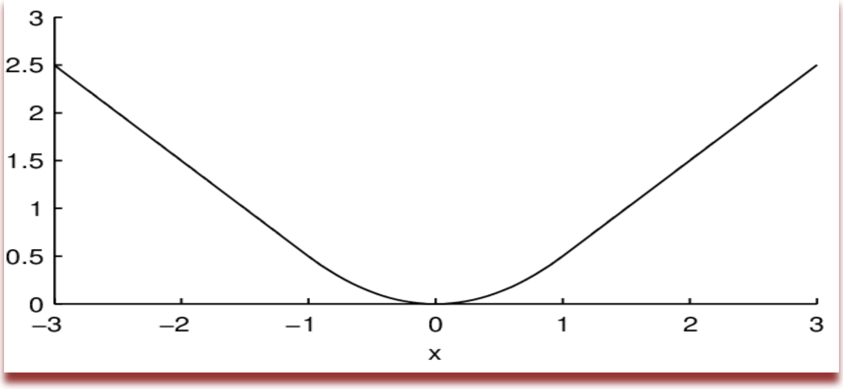

Figure (1): Minim ax Concave Penalty (MCP) method

3. SIMULATION STUDY :

The comparison between method (SEA-LASSO) and method (MCP) was made using a simulated study that included different cases , and testing the effect of the levels of these factors on the performance of these two ways. The program (R) was used when studying the simulation . The simulation study is presented as following:

■ Four levels of sample size were chosen (25 ,50 ,200,1000).

■ Three levels of random error bound variance were used to compare these two methods .

(σ^2= 0.25, σ^2 = 1 , σ^2= 3) .

■ Five levels of linear correlation were used for the explanatory variables (ρ = 0.5 , ρ = 0.7

ρ = 0.9 , ρ = 0.95 , ρ = 0.99) .

■ Two levels of real model regression parameter values were used, and at both levels the number of real model variables were set P_1= 2 , and the total number of variables P = 8 is as follows :

The first level: The regression parameters of the real model variables are equal β_j =6 *for* j =1,2

The second level: The different regression parameters of the real model variables β_j = 6 − j *for* j=1,2 .

At both levels, it is: β_j= 0 *for* j = 3,4,..,8

■ Based on the levels of the factors of the previous study, 1000 samples were generated from the linear regression model: Y = Xβ +ε ,
For each of the study cases, as follows:

1- The random error term was generated from the normal distribution ε~ N (0, $\sigma^2 I_n$) .

2- The design matrix was generated from a multivariate normal distribution with zero mean vector ($0_{P\times1}$) and a covariance matrix equal to the unit matrix (I_p). In order to obtain the required degree of linear correlation between the explanatory variables, without any approximation, which makes it possible to observe the real effect of the required linear correlation on the performance of the two methods.

3- Depending on (1) and (2) the response vector is obtained : Y = Xβ +ε .

■ Qi an and Yang ,(2013) used the BIC criterion (Schwarz, 1978)to choose the tuning parameter (λ) for the SEA-LA SSO method as well as for (MCP) method, which was calculated for each sample for each case of the study as follows:

$$\text{BIC}(\boldsymbol{x}\hat{\beta}) = \ln\left(\frac{\|y - x\hat{\beta}\|^2}{\boldsymbol{n}}\right) + \frac{\ln(\mathbf{n})}{\boldsymbol{n}} \times \hat{d}f\,(\boldsymbol{x}\hat{\beta}) \quad \ldots\ldots\ldots(7)$$

Where $\hat{d}f\,(\boldsymbol{x}\hat{\beta})$: is the degree of freedom .

Many researchers agree, for example (Efron et al, 2004; Zou et al, 2007) that $\hat{d}f$ $(\boldsymbol{x}\hat{\beta})$is equal to the number of estimates of non-zero regression coefficients. The best value for the tuning parameter (λ) is chosen, which makes the BIC criterion as low as possible.

A comparison was made between the two methods based on the following :

1- Extracting the percentage of runs (PCT), which works on evaluating the behavior of the two methods in terms of their ability to reach the real model under a specific case, and it is defined as follows:

$$\text{PCT} = \frac{\text{The number of times the real model is reached}}{\text{n}} \times 100 \ldots\ldots(8)$$

Where : n = number of samples = 1000 sample and $0 \leq$ PCT ≤ 100

If it is (PCT = 0), this means that the Penalized method was not able to reach the true model in all samples, and when (PCT=100) this means that the Penalized method was able to reach the true model in all samples. This means that the Penalized method that gives the highest percentage (PCT) is the best in the light of the concerned case (Kālu, 2014) .

2- To evaluate the performance of the two methods, it was relied on Mean of Squared Errors (MSE) in terms of the accuracy of the estimate, for each case of the study, as follows:

$$\text{MSE} = \frac{\sum_{r=1}^{1000}\|\hat{\beta}_r - \beta\|^2}{1000} \ldots\ldots\ldots(9) .$$

Where $\hat{\beta}_r$: It is the estimation of the regression parameters vector for sample No. (r) for either of the two methods. The method that gives less (MSE) is preferable to the given case.

3.1 Study results analysis.

In this section, we discuss the ability of the two methods to reach the real model, and evaluate the performance of the two methods in terms of the accuracy of the estimate, as shown in the following tables.

Table (1): The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods (SEA-LASSO and MCP), when using the first level, if the regression coefficients of the real model variables are equal, when the sample size is small (n =25) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance (σ^2= 3), using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

Case sequence	levels of correlation (ρ)	Variance(σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	62	100	0.0821	0.0323
2	0.7	0.25	54.2	100	0.2319	0.0497
3	0.9	0.25	56.1	100	0.3717	0.1404
4	0.95	0.25	50.1	100	0.7124	0.3436
5	0.99	0.25	53.9	96.6	3.7587	3.7069
6	0.5	1	58.9	96.9	0.2906	0.1343
7	0.7	1	52.9	97.9	0.4972	0.2002
8	0.9	1	52.8	96.8	2.4617	0.6550
9	0.95	1	53.4	92.9	2.9401	1.6619
10	0.99	1	51.2	60.3	15.8373	19.2869
11	0.5	3	59.9	72.2	0.7589	0.6731
12	0.7	3	52.9	69.9	1.9668	1.0822
13	0.9	3	50.8	70.9	3.7018	3.1759
14	0.95	3	51.9	65.9	8.8726	7.9505
15	0.99	3	45.2	19.9	56.6079	63.1922

We note from Table (**1**) that the (SEA-LASSO) method is superior to the MCP method when using the first level of the vector (β), when the sample size is small and the variance is low (σ^2= 0.25). The (Sea-Lasso) method has the lowest level (MSE) when using this level. Also (Sea-Lasso method) was able to achieve the percentage (PCT = 100) in most of the cases and the MCP method could not reach this percentage. But when the variance is medium σ^2= 1), the (Sea-Lasso) method has the lowest (MSE) in most cases and gives the highest (PCT) than the (MCP) method. In the case of high variance (σ^2= 3), the (Sea-Lasso) method has the lowest (MSE), and gives the highest percentage (PCT) than the (MCP) method in most cases.

Table (2). The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods, when using the first level, , if the sample size is medium (n =50) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance (σ^2= 3),using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

.

Case sequence	levels of correlation (ρ)	Variance(σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	73.8	100	0.0196	0.0179
2	0.7	0.25	71.2	100	0.0239	0.0240
3	0.9	0.25	72.3	100	0.2113	0.0581
4	0.95	0.25	73.1	100	0.3235	0.2677
5	0.99	0.25	75.6	100	1.4116	1.9221
6	0.5	1	72.5	100	0.2258	0.0816
7	0.7	1	73.1	100	0.2064	0.0667
8	0.9	1	71.8	100	0.4992	0.3111
9	0.95	1	72.9	97.8	2.2141	0.5576
10	0.99	1	68.8	84.9	6.1187	9.1315
11	0.5	3	75.9	92.7	0.2931	0.1955
12	0.7	3	70.7	95.4	0.4976	0.2817
13	0.9	3	71.3	91.3	2.0611	0.9567
14	0.95	3	74.2	86.6	2.8627	1.0809
15	0.99	3	62.9	55	21.998	29.068

We note from Table (**2**) that the (SEA-LASSO) method is superior to (MCP) when using the first level of the vector (β) when the sample size is medium (n=50) and the variance is low (σ^2= 0 .25). The (Sea-Lasso) method had the lowest (MSE) in this type of sample when using this level, and (Sea-Lasso) method was able to achieve the ratio (PCT = 100) in all cases and the MCP method could not reach this the ratio.

When the variance is medium(σ^2= 1), the (Sea-Lasso) method has the least (MSE) in this type of sample when using this level, and the (Sea-Lasso) method was able to achieve the ratio (PCT = 100) in most cases, and the MCP method could not reach this percentage.

In the case of high variance, the (Sea-Lasso) method has the lowest (MSE) in this type of sample when using this level, and gives the highest percentage (PCT) than the (MCP) method in most cases.

Table (3): The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods, when using the first level, if the sample size is relatively large (n =200) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance (σ^2= 3) using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

case sequence	levels of correlation (ρ)	Variance (σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	88.9	100	0.0290	0.0155
2	0.7	0.25	86.9	100	0.0199	0.0175
3	0.9	0.25	84.8	100	0.0187	0.0115
4	0.95	0.25	86.7	100	0.0395	0.0302
5	0.99	0.25	87.8	100	0.3901	0.6892
6	0.5	1	93.2	100	0.0279	0.0262
7	0.7	1	85.7	100	0.0412	0.0339
8	0.9	1	86.9	100	0.2071	0.1579
9	0.95	1	88.3	100	0.2811	0.2485
10	0.99	1	85.9	97.9	0.8716	1.9973
11	0.5	3	92.3	100	0.0712	0.0392
12	0.7	3	85.9	100	0.2041	0.1601
13	0.9	3	88.9	100	0.2951	0.2010
14	0.95	3	89.1	100	0.4989	0.3981
15	0.99	3	87.9	92.9	3.1712	4.9249

We notice from Table (**3**) that the (SEA-LASSO) method is superior to (MCP) when using the first level of the vector (β) when the sample is relatively large (n=200) and the variance is low(σ^2= 0 .25), and the (Sea-Lasso) method has less (MSE) In most cases, it also achieved the percentage (PCT = 100) in all cases, and the MCP method was not able to reach this percentage.

In the case of medium variance (σ^2= 1), the (Sea-Lasso) method had the lowest (MSE), and (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in most cases.

When the variance is high, the (Sea-Lasso) method has a lower (MSE) than the (MCP), and it was able to achieve the percentage (PCT = 100) in most of the cases.

Table (4): The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods, when using the first level, if the sample size is very large (n =1000) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance (σ^2= 3) using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

case sequence	levels of correlation (ρ)	Variance (σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	94.9	100	0.0109	0.0225
2	0.7	0.25	96.1	100	0.0112	0.0156
3	0.9	0.25	93.9	100	0.0141	0.0231
4	0.95	0.25	94.8	100	0.0179	0.0411
5	0.99	0.25	94.9	100	0.0343	0.4451
6	0.5	1	94.9	100	0.0125	0.0165
7	0.7	1	95.1	100	0.0161	0.0168
8	0.9	1	96.2	100	0.0228	0.0271
9	0.95	1	95.2	100	0.0290	0.0382
10	0.99	1	95.1	100	0.2371	0.5939
11	0.5	3	94.9	100	0.0214	0.0227
12	0.7	3	94.1	100	0.0271	0.0251
13	0.9	3	95. 8	100	0.0359	0.0399
14	0.95	3	95,2	100	0.0911	0.1191
15	0.99	3	94.8	100	0.5236	1.9299

We note from Table (**4**) that the (SEA-LASSO) method is superior to (MCP) when using the first level of the vector (β) when the sample is very large and the variance is low. The (MCP) method has the lowest (MSE) in this type of sample when using this level, and the (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in all cases.

In the case of the medium variance (σ^2= 1), the (MCP) method has the lowest (MSE) in this type of sample when using this level of the (Sea-Lasso) method, but the (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in all cases.

But when the variance is high (σ^2= 3). The (MCP) method has less (MSE) in most cases than this type of sample when using this level of (Sea-Lasso) method, but also (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in all cases.

Table (5): The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods, when using the second level, i.e. when the regression parameters of the real model variables are not equal. If the sample size is small (n =25) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance (σ^2= 3) using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

case sequence	levels of correlation (ρ)	Variance (σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	59.9	100	0.0617	0.0295
2	0.7	0.25	51.9	100	0.2438	0.0381
3	0.9	0.25	53.8	98.9	0.2942	0.1537
4	0.95	0.25	50.9	98.7	0.7906	0.1978
5	0.99	0.25	53.9	82.1	3.0801	3.9197
6	0.5	1	58.9	81.9	0.3211	0.2951
7	0.7	1	52.9	88.1	0.4961	0.3018
8	0.9	1	54.1	82.9	1.9078	0.8901
9	0.95	1	53.3	77.1	3.0326	1.9713
10	0.99	1	40.3	38.4	19.989	20.9601
11	0.5	3	62.1	66.1	0.7561	0.6810
12	0.7	3	52.9	62.9	1.9653	1.3754
13	0.9	3	51.9	59.2	5.0324	4.0701
14	0.95	3	48.8	52.3	9.8821	9.2342
15	0.99	3	16.1	7.9	56.0172	52.9761

We note from Table (**5**) that the (Sea-LASSO) method is superior to the MCP method when using the second level of the vector (β), when the sample size is small and the variance is low. Where she had lower (MSE) in most cases. Also (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in two cases and achieved the highest percentage in the rest of the cases.

It also outperformed the (MCP) method when the variance is medium and has the lowest (MSE) in most cases, and gives a higher percentage (PCT) than the (MCP) method in most cases.

As for the high variance. The (Sea-Lasso) method had lower (MSE) in most of the cases, and gave the percentage (PCT) slightly higher than (MCP) method in most cases.

Table (6): The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods, when using the second level, if the sample size is medium (n =50) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance (σ^2= 3),using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

case sequence	levels of correlation (ρ)	Variance (σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	75.1	100	0.0498	0.0245
2	0.7	0.25	71.1	100	0.0502	0.0151
3	0.9	0.25	70.8	100	0.1328	0.0591
4	0.95	0.25	70. 9	100	0.3041	0.1815
5	0.99	0.25	75.1	95.9	0.9981	1.9131
6	0.5	1	73.9	97.9	0.2115	0.0551
7	0.7	1	70.8	97.5	0.2937	0.0887
8	0.9	1	71.9	96.8	0.4724	0.1976
9	0.95	1	72.8	93.5	0.9921	0.5791
10	0.99	1	65.6	67.6	5.9661	8.9549
11	0.5	3	76.1	80.4	0.4264	0.1956
12	0.7	3	71.8	82	0.4966	0.3924
13	0.9	3	70.1	80	1.9760	1.7484
14	0.95	3	71	76.9	2.8909	2.2322
15	0.99	3	37.8	27.5	25.9433	31.8783

We note from Table (**6**) that the (SEA-LASSO) method is superior to (MCP) when using the second level of the vector (β) when the sample size is medium and the variance is low. The (Sea-Lasso) method had the lowest (MSE) in this type of sample when using this level, and (Sea-Lasso) method was able to achieve the ratio (PCT = 100) in most cases.

We also note in this table that the (SEA-LASSO) method is superior to (MCP) when

case sequence	levels of correlation (ρ)	Variance (σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	88.5	100	0.0049	0.0043
2	0.7	0.25	86.7	100	0.0079	0.0065
3	0.9	0.25	84.9	100	0.0187	0.0199
4	0.95	0.25	86.8	100	0.0511	0.0299
5	0.99	0.25	87.9	100	0.1919	0.4921
6	0.5	1	90.9	100	0.0192	0.0152
7	0.7	1	86.8	100	0.0417	0.0199
8	0.9	1	86.7	100	0.1119	0.0807
9	0.95	1	88	100	0.2097	0.1702
10	0.99	1	88.1	98.6	0.8968	1.9578
11	0.5	3	89.9	97.9	0.0617	0.0521
12	0.7	3	88.1	99	0.2098	0.0583
13	0.9	3	88.7	98.6	0.1885	0.1594
14	0.95	3	87.8	98.9	0.6381	0.5221
15	0.99	3	85.6	77.9	2.6999	6.9491

the variance is medium. It also has lower (MSE) and higher percentage (PCT) than (MCP) method in most cases.

As for the high variance. The Sea-Lasso method has a lower MSE and a higher PCT rate than MCP in most cases.

Table (7): The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods, when using the second level, if the sample size is relatively large (n =200) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance

(σ^2= 3),using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

We note from Table (**7**) that the (SEA-LASSO) method is superior to (MCP) when using the second level of the vector (β) when the sample is relatively large and the variance is low. It also has less (MSE) in most cases, and was (Sea-Lasso)method was able to achieve the percentage (PCT = 100) in all cases, and the MCP method was not able to reach this percentage.

We also note from the table that the (SEA-LASSO) method is superior to (MCP) when the variance is medium. Whereas, the MSE was lower in most cases, and the (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in most cases.

As for the high variance. The (Sea-Lasso) method had the lowest (MSE) in most cases, and (Sea-Lasso) method was able to give a higher percentage (PCT) than the (MCP) method in most of the cases.

Table (8) :The percentage of runs to reach the real model (PCT) and the mean of the squares of error (MSE) for the two methods, when using the second level, if the sample size is very large (n =1000) and three levels of variance of the random error term (low variance(σ^2= 0.25), medium variance (σ^2= 1), and high variance (σ^2= 3) using five levels of linear correlation between the explanatory variables **(**0.5, 0.7, 0.9, 0.95, 0.99).

case sequence	levels of correlation (ρ)	Variance (σ^2)	PCT		MSE	
			MCP	SEA-LASSO	MCP	SEA-LASSO
1	0.5	0.25	94.9	100	0.0017	0.0135
2	0.7	0.25	96.1	100	0.0025	0.0045
3	0.9	0.25	93.8	100	0.0050	0.0231
4	0.95	0.25	96.1	100	0.0068	0.0436
5	0.99	0.25	94.9	100	0.0432	0.5252
6	0.5	1	94.9	100	0.0028	0.0038
7	0.7	1	95.2	100	0.0049	0.0055
8	0.9	1	96.1	100	0.0128	0.0222
9	0.95	1	95.3	100	0.0298	0.0597
10	0.99	1	95.2	100	0.3420	0.6188
11	0.5	3	96.2	100	0.0201	0.0154
12	0.7	3	94.1	100	0.0199	0.0158
13	0.9	3	95.6	100	0.0380	0.0399
14	0.95	3	94.9	100	0.0643	0.2541
15	0.99	3	94.8	97.9	0.3915	0.8756

We notice from Table (8) that the (SEA-LASSO) method is superior to (MCP) when using the second level of the vector (β) when the sample is very large and the variance is low. The (MCP) method had the lowest (MSE) in this type of sample when using this level, and the (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in all cases.

We also note from the table that the (SEA-LASSO) method is superior to (MCP) when the variance is medium. The (MCP) method had the lowest (MSE) in this type of sample, and (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in all cases.

In the case of high variance, the (MCP) method has less (MSE) in most cases, and the (Sea-Lasso) method was able to achieve the percentage (PCT = 100) in most cases.

4. Conclusion

This study presented a comparison between two punitive methods that were used in this research to estimate and select the appropriate variables at the same time in the multiple regression model and to choose the best, namely (SEA-LASSO) method and (MCP) method. The results showed that the (SEA-LASSO) method outperformed the (MCP) method in terms of accessing the real model measured by (PCT), and estimation accuracy, as it achieved the lowest (MSE) in most cases for small samples (n = 25) and medium samples (n = 50) in the first and second levels.

The (SEA-LASSO) method is best in the cases of relatively large samples (n = 200) in view of the medium and high variance($\sigma^2 = 1$, $\sigma^2 = 3$) , regardless of the equality or difference in the regression coefficients of the real model variables.

In the case of very large samples (n = 1000), the MCP method had the lowest (MSE), but it did not outperform the (SEA-LASSO) method in terms of access to the real model (PCT), where the (SEA-LASSO) method had achieved the highest percentage. Thus, the (SEA-LASSO) method is considered the best in terms of access to the real model (PCT) and in terms of estimation accuracy because it achieved the lowest (MSE) in most cases by studying the simulation results.
We recommend the use of some other variable selection methods such as Group Lasso, and other methods of regularization that work on estimating and appropriate selection of variables at the same time, addressing the polyline problem, and solving other statistical problems

REFERENCES:

1- Bellman, R. E. (1961). Adaptive Control Processes. Princeton University Press, Princeton, New Jersey.
2- Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. {2004). Least angle regression.
The Annals of Statistics, 32, 407-499.
3- Fan, J. and Li, R. (2001), "Variable Selection via Non concave Penalized Likelihood and its Oracle Properties," Journal of the American Statistical Association, **96**, pp-.1348-1360.

4- http://faculty.ksu.edu.sa/sites/default/files/mhdr_4_thlyl_lnhdr_lkhty_lmtdd_1.pdf.
5- https://statisticaloddsandends.wordpress.com/2019/12/09/the-minimax-concave-penalty-mcp/.

6- Kaul, A. (2014), "LASSO with Long Memory Regression Errors," Journal of Statistical Planning and Inference, **153** Annals of the Institute of Statistical Mathematics, **65**, pp.295-318

7- Schwarz, G. (1978), "Estimating the Dimension of a Model," The An-nals of Statistics, **6**, pp.461-464, pp.11-26.
8- Lian, H. (2012), "Variable Selection in High-Dimensional Partly Li-near Additive Models," Journal of Nonparametric Statistics, **24**, pp. 82-5-839.
9- Qian, W. and Yang, Y. (2013), "Model Selection via Standard Error Adjusted Adaptive LASSO,"
.
10- Wang, X., Wei, M., & Yao, T. (2018, July). Minimax concave penalized multiarmed bandit model with high-dimensional covariates. In International Conference on Machine Learning (pp. 5200-5208). PMLR.
11- Zou, H., Hastie, T. and Tibshirani, R. {2007). On the degrees of freedom of the lasso.
The Annals of Statistics, 35, 2173-2192.

12- Zhang, C. H. (2010). Nearly unbiased variable selection under minimax concave penalty. The Annals of statistics, 38(2), 894-942.

13- Zou, H. (2006), "The Adaptive L-ASSO and its Oracle Properties," Journal of the American Statistical Association, **101**, pp.1418-1429.

المستخلص

يعد اختيار المتغير موضوعًا مهمًا في تحليل الانحدار الخطي. من الناحية العملية ، عادةً ما يتم تقديم عدد كبير من المتنبئين في المرحلة الأولى من بناء النموذج للتخفيف من تحيزات النموذج المحتملة. من ناحية أخرى ، لتعزيز القدرة على التنبؤ بالمتغيرات المهمة واختيارها.

بذل الإحصائيون جهودًا كبيرة في تطوير إجراءات التنظيم لحل مشكلات (V.S). تسهل هذه الإجراءات تلقائيًا الاختيار المتغير (V.S) عن طريق تعيين معاملات محددة إلى الصفر وتقليل تقديرات المعامل ، مما يوفر تقديرات مفيدة حتى لو كان النموذج يحتوي على عدد كبير من المتغيرات. في هذا البحث ، تم اقتراح طريقتين للتنظيم لتقدير واختيار المتغيرات المناسبة في نفس الوقت في نموذج الانحدار المتعدد ، وهما أسلوب LASSO المعياري المعدل للخطأ (SEA-LASSO) و Minimax Concave Penalty (MCP) و اختيار الافضل.

تركز مشكلة الورقة على استخدام أفضل طريقة تنظيم تعمل على التقدير والاختيار المناسب للمتغيرات المهمة في نفس الوقت ومعالجة مشكلة الخطية المتعددة باستخدام طريقة SEA-LASSO و MCP. للحصول على النموذج الحقيقي.

تهدف هذه الورقة إلى تقييم أداء طريقة (SEA-LASSO) وطريقة (MCP) من حيث عملية التقدير والاختيار المناسب للمتغيرات الهامة ومعالجة مشكلة الخطية المتعددة من خلال دراسة المحاكاة.

أجريت دراسة محاكاة للمقارنة بين هاتين الطريقتين والتي اشتملت على حالات مختلفة من العوامل واختبار تأثير مستويات هذه العوامل على أداء هاتين الطريقتين ، وكذلك تحديد قيمة معامل الضبط (λ) ومعيار اختيار أفضل قيمة لها والأساس الذي يتم على أساسه تقييم أداء الطريقتين. أظهرت نتائج المحاكاة أن طريقة (SEA-LASSO) تتفوق على طريقة (MCP) من حيث النسبة المئوية للتشغيل للوصول إلى النموذج الحقيقي المقاس بواسطة (PCT) ، كما أنها أفضل من حيث متوسط مربعات الخطأ (MSE) لأنه يحقق أقل (MSE) في معظم الحالات. تم استخدام دراسة محاكاة مع برنامج R.

الكلمات الرئيسية: ، Multiple regression model , Adjusted Adaptive LASSO (SEA-LASSO), Minimax Concave Penalty (MCP).