ORIGINAL RESEARCH ORİJİNAL ARAŞTIRMA DOI: 10.5336/biostatic.2024-106402

A Comparison of Penalized Regression Methods on Model Estimation and Variable Selection: A Simulation Study

Cezalı Regresyon Yöntemlerinin Model Tahmini ve Değişken Seçimindeki Etkisinin Karşılaştırılması: Simülasyon Çalışması

Ali Türker ÇİFTÇİ[a], Didem DERİCİ YILDIRIM[b], Damla Hazal SUCU[b]

[a]Niğde Ömer Halisdemir University Faculty of Medicine, Department of Biostatistics and Medical Informatics, Niğde, Türkiye
[b]Mersin University Faculty of Medicine, Department of Biostatistics and Medical Informatics, Mersin, Türkiye

This study was presented as an oral presentation at 1st National Health Sciences Graduate Student Symposium, 28-29 April 2022, Mersin, Türkiye.

ABSTRACT Objective: The aim of this study is to determine which variables to include in the model and to examine how successful the model estimation is with the Least Angle Regression (LAR), Least Absolute Shrinkage Selection Operator (LASSO) and Elastic Net (EN) regression methods, which are alternatives to unbiased methods. **Material and Methods:** In this study, variables that LAR, LASSO and EN regression methods, which are among the biased methods, take in model selection and their model prediction success are compared. For this purpose, data sets were generated in different scenarios in R program. The results obtained after data sets produced at the end of the simulation with standard normal distribution, sample sizes n=50, n=100, n=200, number of independent variables p=16, p=18, p=20 and correlation coefficients r=0.10; r=0.60; r=0.90 were recorded. **Results:** Model predictions of the methods were recorded in the study results. While the model predictions of the LAR and LASSO methods were close to each other, the EN regression method differed in model prediction. When analyzed in terms of Mean Square Error (MSE) and Coefficients of Determination, close values were observed. **Conclusion:** While model prediction success is high in data sets with low sample size, model prediction success decreases and MSE values increase when the sample size increases gradually. For this reason, it has been observed that these methods are more useful and provide better model prediction in cases where there is a multicollinearity problem in the data sets and in scenarios where the sample size is small.

Keywords: Penalized regression; variable selection; least angle regression; least absolute shrinkage selection operator; elastic net regression

ÖZET Amaç: Bu çalışmanın amacı, hangi değişkenlerin modele dâhil edileceğini belirlemek ve yansız yöntemlere alternatif olan En Küçük Açı Regresyonu [Least Angle Regression (LAR)], En Küçük Mutlak Büzülme Seçim Operatörü [Least Absolute Shrinkage Selection Operator (LASSO)] ve Elastik Ağ [Elastic Net (EN)] regresyon yöntemleri ile model tahmininin ne kadar başarılı olduğunu incelemektir. **Gereç ve Yöntemler:** Bu çalışmada, yanlı yöntemlerden olan LAR, LASSO ve EN regresyon yöntemlerinin model seçiminde hangi değişkenleri aldığı ve model tahmin başarıları karşılaştırılmıştır. Bu amaçla R programında farklı senaryolarda veri setleri üretilmiştir. Standart normal dağılıma sahip, örnek genişlikleri n=50, n=100, n=200, bağımsız değişken sayıları p=16, p=18, p=20 ve korelasyon katsayıları r=0,10; r=0,60; r=0,90 olacak biçimde yapılan simülasyon sonunda üretilen veri setlerinin ardından elde edilen sonuçlar kaydedilmiştir. **Bulgular:** Çalışma sonuçlarında yöntemlerin model tahminleri kaydedilmiştir. LAR ve LASSO yöntemlerinin model tahminleri birbirine yakınken, EN regresyon yöntemi, model tahmininde farklılık göstermiştir. Hata Kareler Ortalaması (HKO) ve Belirtme Katsayıları ile incelendiğinde ise birbirlerine yakın değerler gözlenmiştir. **Sonuç:** Örnek genişliği düşük olan veri setlerinde model tahmin başarısı yüksek iken, örnek genişliğinin giderek arttığı durumlarda model tahmin başarısı azalmakta, HKO değerleri artmaktadır. Bu sebeple veri setlerinde çoklu bağlantı sorunu olduğu durumda ve örnek genişliğinin küçük olduğu senaryolarda bu yöntemlerin daha kullanışlı olduğu ve daha iyi model tahmini yaptığı görülmüştür.

Anahtar kelimeler: Cezalı regresyon; değişken seçimi; en küçük açı regresyonu; en küçük mutlak büzülme seçim operatörü; elastik ağ regresyon

TO CITE THIS ARTICLE:
Çiftçi AT, Derici Yıldırım D, Sucu DH. A comparison of penalized regression methods on model estimation and variable selection: A simulation study. Turkiye Klinikleri J Biostat. 2025;17(1):1-15.

Correspondence: Ali Türker ÇİFTÇİ
Niğde Ömer Halisdemir University Faculty of Medicine, Department of Biostatistics and Medical Informatics, Niğde, Türkiye
E-mail: turker_ciftci_42@hotmail.com

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

Received: 21 Oct 2024 ***Received in revised form:*** 25 Feb 2025 ***Accepted:*** 26 Feb 2025 ***Available online:*** 14 Mar 2025

Regression analysis is used in clinical studies to determine the relationship between 2 or more independent variables and a dependent variable that have a cause-and-effect relationship and to make predictions or forecasts on the topic based on this relationship. The primary aim of regression analysis is to develop the predictive model that explains the dependent variable with the fewest possible variables.[1]

In clinical studies, interpreting the model becomes challenging when there are a large number of independent variables. In such cases, variable selection must be performed to identify the best model with fewer variables. While selecting variables, the relationship between independent variables is the most critical aspect to consider. One of the most important assumptions of multiple linear regression is the absence of relationships among the independent variables. If this assumption is violated, multicollinearity arises. In the case of multicollinearity, models estimated using the Ordinary Least Squares (OLS) method lead to erroneous results.[2] To address this problem, shrinkage estimation methods, which are biased estimators, have been proposed as an alternative to OLS.[3] Among these methods, the most commonly used are Ridge Regression, Least Angle Regression (LAR), Least Absolute Shrinkage and Selection Operator (LASSO), and Elastic Net (EN) regression.

Although penalized regression methods are biased estimators, they are used as alternatives to unbiased estimation methods due to their ability to meet theoretical expectations and reduce variance.[4] Another function of these methods is to perform variable selection. While estimating the model, these methods exclude independent variables that do not explain the dependent variable, thus performing variable selection. Ridge Regression, one of the penalized regression methods, does not perform variable selection.[5] It includes all independent variables in the dataset in the model. For this reason, the Ridge Regression method is not addressed in this study.

The 1st aim of this study is to determine which variables are included in the model by the LAR, LASSO, and EN regression methods, which are alternatives to unbiased methods. The second aim is to evaluate the coefficients of the included variables and the success of the model estimation using the Coefficient of Determination (R^2) and the Mean Squared Error (MSE).

MATERIAL AND METHODS

LEAST ANGLE REGRESSION

The LAR method was first proposed by Efron et al. as a method that can be used for variable selection in linear models. The LAR method selects the x_j variable with the highest absolute correlation among all the independent variables in the dataset. This x_j performs a simple linear regression with the dependent variable y. Then, it initiates an algorithmic process to select the other independent variables according to a geometric order. This method applies as many steps as the number of variables and forms a model with the variables obtained at the end of n steps to estimate parameters.[6]

The LAR method starts with a model containing no variables. Similar to forward selection methods, it begins the algorithm process with the independent variable x_j that has the highest correlation with the dependent variable. This x_j variable takes a large step in the direction of the residuals $(y-\hat{y})$ with the highest correlation and calculates the β_{j1} coefficient. The algorithm stops when it finds another variable x_{j2} with the highest correlation with the residuals, just as it did with x_{j1}.The second variable x_{j2} is included in the active model.[7] The coefficients of these included variables are adjusted together to maintain and reduce their correlations. At this stage, the LAR algorithm diverges from other methods. Instead of continuing in the direction of the x_{j1} variable alone, it proceeds in an angular direction that balances x_{j1} and x_{j2}. The algorithm continues in this equal-angle direction until it finds a 3rd independent variable that has an equal correlation with the residuals as the vector formed by x_{j1} and x_{j2}.

When there are 2 independent variables, the logic of the LAR method, as shown, is that $\bar{y}$ is the projection of y onto L (x_1, x_2). The LAR procedure starts with $\hat{\mu}_0$=0 (It represents the dependent variable value of the LAR estimator when the model contains no variables). The residual vector $\bar{y}_1$-$\hat{\mu}_1$ ($\hat{\mu}_1$ represents the value of the dependent variable estimated in the 1st step) will have a stronger correlation with x_1 than x_2. For this reason, the LAR estimation is $\hat{\mu}_1=\hat{\mu}_0+\hat{\delta}_1 x_1$. Here, $\widehat{\delta}_1$ is chosen such that the angle between x_1 and x_2 is bisected by $\bar{y}_2$ - $\hat{\mu}_1$. From this, $\hat{\mu}_2=\hat{\mu}_1+\hat{\delta}_1 u_2$ emerges, where u_2 is the angle bisector. The equality $\bar{y}_2=\hat{\mu}_2$ holds only when there are two independent variables.

In the LAR method, the algorithm increases the β estimates in each step to remain close to the β estimates predicted by OLS. The objective of this method is:

$$\min\{e' e+\lambda\beta\} \quad (1)$$

Since LAR is a penalized regression method, it aims to minimize the sum of squared errors using a penalty parameter, denoted by λ .

LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR

Although OLS estimates yield high prediction accuracy, they exhibit high variance, especially in cases of multicollinearity where they become insufficient.[8] OLS estimators often produce predictions with low bias but high variance. Although Ridge Regression is a technique developed to enhance OLS estimates, it complicates interpretability. As an alternative, Tibshirani introduced the LASSO method.[9]

Let x_j represent the independent variables for j=1,2,3,…, m; y_i represent the dependent variable and m denote the number of independent variable. Multiple linear regression is expressed as:

$$\mu=\beta_0+\sum_{j=1}^{m}(x_j\,\beta_j) \text{ (}\mu\text{, the value predicted by the model for the dependent variable)} \quad (2)$$

subject to:

$$\sum_{i=1}^{n} y_i=0,\sum_{j=1}^{n}(x_j=0),\ \sum_{j=1}^{n}(x_j^2=1),\ j=1,2,\ldots\ldots,m \quad (3)$$

The regression coefficients of the independent variable vectors ($\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_{3,\ \ldots,}$ $\hat{\beta}_m$) form the estimate vector $\hat{\mu}$.

$$\hat{\mu}=\sum_{j=1}^{m}(x_j\,\hat{\beta}_j)=X\hat{\beta} \quad [X_{nxm}=(x_1, x_2, x_3, \ldots, x_m)] \quad (4)$$

The total sum of squared errors is given by:

$$S(\hat{\beta})=\|y-\hat{\mu}\|^2=\sum_{j=1}^{n}(y-\hat{\mu}_j)^2 \quad (5)$$

If the absolute form of β is denoted as T($\hat{\beta}$), then:

$$T(\hat{\beta})=\sum_{j=1}^{m}|\beta j| \quad (6)$$

In the above equation, T($\hat{\beta}$), represented by λ, is the penalty (or tuning) parameter in the LASSO method. Thus, LASSO aims to solve the problem:

$$\sum_{j=1}^{m}|\beta j|<\lambda \quad (7)$$

by minimizing:

$$\text{argmin}_\beta \sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{m}(x_j\beta_j))^2 \quad (8)$$

The LASSO estimator is defined as:[10]

$$\hat{\beta}^{LASSO} = \text{argmin}_{\beta} \{\sum_{i=1}^{n}((y_i - \beta_0 - \sum_{j=1}^{m}(x_j\beta_j))^2 + \lambda \sum_{j=1}^{m} |\beta j|\} \quad (9)$$

LASSO is a method that constructs a model by shrinking some coefficients toward zero and setting others exactly to zero. Its primary advantage lies in the quality of its model estimation. By reducing or eliminating some coefficients, it decreases variance.

As a penalized regression method, LASSO differs from the LAR method by combining the penalty parameter with $|\beta|$ instead of β. The objective of this method is formulated as:

$$\min\{e^{'} e + \lambda|\beta|\} \text{ (e, represents the prediction errors of the model)} \quad (10)$$

Although LAR and LASSO are closely related, LAR performs the calculations in a single step, making it faster. However, LASSO's disadvantage is its lack of robustness in cases of strong multicollinearity. When the number of independent variables exceeds the sample size, it selects at most as many variables as the sample size. To address this limitation, the EN regression method was proposed.[11]

ELASTIC NET REGRESSION

EN regression is another penalized regression method proposed by Zou and Hastie. As in other penalized regression techniques, it performs variable selection and estimates coefficients for independent variables.[12]

The EN regression method estimates the model by applying the penalty parameters of both Ridge Regression and LASSO regression. It combines both methods to create a model. It tends to minimize the coefficients by applying penalties to the coefficients, as in Ridge regression, and gives a model estimation that can be more efficient by selecting variables, as in LASSO regression.[13]

In this method, let x_j represent the independent variables j=1,2,3,…,m and y_i denote the dependent variable.

$$\sum_{i=1}^{n} y_i=0 \quad , \sum_{j=1}^{n}(x_j=0) \quad , \sum_{j=1}^{n}(x_j^{2}=1), \; j=1,2,3,\ldots, m \quad (11)$$

The regression coefficients $\hat{\beta}_m$ determine $\hat{\mu}$.

The total sum of squared errors is given by:

$$S(\hat{\beta})=||y-\hat{\mu}||^2=\sum_{j=1}^{n}(y-\hat{\mu}_j)^2 \quad (12)$$

Elastic Net aims to estimate the model as follows:

$$\hat{\beta}^{ENET}=|y-X\beta|^2+\lambda_1|\beta|_1+\lambda_2|\beta|^2 \quad (13)$$

In Equation (13),

the term $\lambda_1 |\beta|_1$ represents the LASSO penalty, while $\lambda_2 |\beta|^2$ corresponds to the Ridge Regression penalty. Similar to LASSO, λ is the penalty parameter used for model estimation.[14,15] In light of all this methodological information, we can summarize the simulation steps we have performed in our study:

SIMULATION STUDY

The following steps were taken in the simulation study:

- Independent variables were generated from the Standard Normal Distribution [n (0.1)] with p=16, 18, and 20.
- Sample sizes of n=50, 100, and 200 were generated for these independent variables.

- Correlations between independent variables were set to r=0.10, 0.60, and 0.90.
- The penalty parameter λ was set within the range of [0.01-100], and its optimal value was determined during the simulation.
- Cross-Validation procedures were applied solely to find the optimal λ penalty parameter.
- R^2 and its standard errors and MSE values and their standard errors were obtained for all combinations.

To ensure the generalizability of regression results, it is desirable to have at least 10 observations per independent variable, ideally between 15 and 20.[16] Consequently, datasets with a sample size of 50 included 3 independent variables, those with a sample size of 100 included 7 independent variables, and those with a sample size of 200 included 13 independent variables. LAR and LASSO methods, which are penalized regression techniques, can only select up to "n" variables when p>n due to the nature of convex optimization problems.[12,13] LAR and LASSO methods never select the correct model in these cases. For this reason, the study focused solely on cases where p<n. Simulations and method comparisons were performed using version 3.6.3 of the R programming language.

Since this study is a simulation study, ethics committee, Declaration of Helsinki and informed consent are not required.

RESULTS

When penalized regression methods were applied to data generated through simulation, the coefficients of the estimated models, R^2 values, and the Mean±Standard Error values of Mean Squared Errors were presented in tables.

When Table 1 is examined, with 16 variables, a correlation coefficient of 0.10, and n=50, the LAR and LASSO methods included the same variables in the model, while the EN method constructed the model with different coefficients. When n=100, the EN method included different variables in the model compared to the other 2 methods, whereas LAR and LASSO selected the same variables and estimated the same coefficients. When n=200, LAR and LASSO included the same variables in the model, while the EN method excluded the variable X_9 and included the variable X_{13}, thus completing the variable selection process differently.

For a correlation coefficient of 0.60 and n=50, LAR and LASSO included only one different variable in their models. While LAR and LASSO calculated similar coefficient values, the EN method estimated smaller regression coefficients. When n=100, LASSO included a different variable compared to LAR, creating distinct predictions. The EN method included other variables in the model, leading to a prediction different from the other 2 methods. For n=200, LAR and LASSO selected the same variables and calculated similar coefficient values for these variables, while the EN method included 2 different variables compared to the other 2 methods.

For a correlation coefficient of 0.90 and n=50, LAR and LASSO included the same variables in the model and arrived at the same prediction. The EN method, however, produced a different prediction. When n=100, LAR and LASSO selected the same variables and completed the variable selection process. For n=200, LAR and LASSO included the same variables in the model, while the EN method differed by including variables X_5 and X_{15} and excluding X_7.

When Table 2 is examined, with 18 variables, a correlation coefficient of 0.10, and n=50, the only variable commonly included by LAR, LASSO, and EN methods was X_{15}. While LAR and LASSO calculated nearly identical regression coefficients, EN produced different estimates. For n=100, LAR, LASSO, and EN included the same three variables in the model, but the inclusion of other variables varied. When n=200, LAR and EN methods differed by only 2 variables in their model predictions, while LASSO included different variables.

TABLE 1: Estimated model when the number of variables is 16 and the correlation coefficient (r) is 0.10, 0.60 and 0.90 respectively

	n=50			n=100			n=200		
	LASSO	**LAR**	**EN**	**LASSO**	**LAR**	**EN**	**LASSO**	**LAR**	**EN**
r=0.10									
β_1	-	-	-	-	-	-	-	-	-
β_2	-	-	-	0.078	0.078	0.066	0.064	0.064	0.054
β_3	0.080	0.080	0.061	0.077	0.077	0.068	0.066	0.066	0.056
β_4	0.082	0.081	-	0.074	0.074	-	0.061	0.061	0.051
β_5	-	-	-	0.086	0.086	-	0.063	0.063	0.049
β_6	-	-	-	0.081	0.081	0.064	0.064	0.064	0.053
β_7	0.110	0.111	-	-	-	-	0.065	0.065	0.054
β_8	-	-	-	-	-	-	0.063	0.063	0.051
β_9	-	-	-	0.085	0.086	0.064	0.062	0.062	-
β_{10}	-	-	-	-	-	0.069	0.062	0.062	0.055
β_{11}	-	-	0.080	-	-	-	0.068	0.068	0.054
β_{12}	-	-	-	-	-	-	0.067	0.067	0.057
β_{13}	-	-	-	-	-	0.071	-	-	0.055
β_{14}	-	-	-	-	-	-	0.065	0.065	0.053
β_{15}	-	-	-	-	-	-	-	-	-
β_{16}	-	-	0.074	0.083	0.083	0.062	0.066	0.066	0.057
r=0.60									
	n=50			**n=100**			**n=200**		
	LASSO	**LAR**	**EN**	**LASSO**	**LAR**	**EN**	**LASSO**	**LAR**	**EN**
β_1	-	-	-	-	-	-	-	-	-
β_2	-	-	-	0.099	0.098	-	0.081	0.081	0.074
β_3	-	-	-	-	0.093	-	0.082	0.082	0.076
β_4	-	-	-	-	-	-	0.082	0.082	0.076
β_5	-	-	-	-	-	0.080	0.079	0.080	0.074
β_6	0.123	0.121	0.101	-	-	0.085	0.080	0.080	-
β_7	-	-	-	0.095	0.093	0.085	0.081	0.081	0.075
β_8	-	-	-	0.097	-	-	0.082	0.081	0.078
β_9	-	-	0.097	-	-	-	0.083	0.084	0.076

β_{10}	-	-	-	0.100	0.099	0.089	-	-	0.077
β_{11}	-	0.111	-	0.096	0.096	0.086	0.080	0.081	0.076
β_{12}	-	-	-	0.102	0.102	0.091	0.078	0.079	0.074
β_{13}	-	-	-	0.097	0.097	0.086	0.082	0.082	-
β_{14}	-	-	-	-	-	-	0.080	0.080	0.074
β_{15}	0.118	0.115	-	-	-	-	0.082	0.082	0.078
β_{16}	0.117	-	0.098	-	-	-	-	-	0.078
r=0.90									
	n=50			**n=100**			**n=200**		
	LASSO	**LAR**	**EN**	**LASSO**	**LAR**	**EN**	**LASSO**	**LAR**	**EN**
β_1	-	-	-	-	-	-	-	-	-
β_2	-	-	0.096	0.097	0.096	0.084	0.084	0.083	0.077
β_3	-	-	-	0.093	0.093	0.083	0.081	0.081	0.074
β_4	-	-	0.098	0.100	0.099	0.084	0.084	0.083	0.077
β_5	0.125	0.123	-	-	-	0.085	-	-	0.077
β_6	0.119	0.117	-	0.102	0.100	-	0.086	0.086	0.076
β_7	-	-	-	0.097	0.096	0.084	0.088	0.088	-
β_8	0.116	0.113	0.094	-	-	-	0.082	0.082	0.075
β_9	-	-	-	-	-	-	0.081	0.081	0.074
β_{10}	-	-	-	0.101	0.100	0.086	0.084	0.082	0.075
β_{11}	-	-	-	0.095	0.094	-	0.083	0.083	0.075
β_{12}	-	-	-	-	-	-	0.083	0.083	0.076
β_{13}	-	-	-	-	-	0.086	0.083	0.083	0.078
β_{14}	-	-	-	-	-	-	0.086	0.085	0.078
β_{15}	-	-	-	-	-	-	-	-	0.077
β_{16}	-	-	-	-	-	-	0.083	0.083	-

LASSO: Least absolute shrinkage and selection operator; LAR: Least angle regression; EN: Elastic-net regression; β: Coefficient of regression; r: Coefficient of correlation; n: Sample size

TABLE 2: Estimated model when the number of variables is 18 and the correlation coefficient (r) is 0.10, 0.60 and 0.90 respectively

	n=50			n=100			n=200		
	LASSO	LAR	EN	LASSO	LAR	EN	LASSO	LAR	EN
r=0.10									
β_1	-	-	-	-	-	-	-	-	-
β_2	-	-	-	-	-	-	0.062	0.061	0.051
β_3	-	-	-	0.076	0.080	0.065	0.062	0.064	0.054
β_4	-	-	-	0.079	0.076	-	0.063	0.061	0.050
β_5	0.079	0.080	-	-	-	-	0.065	0.062	0.052
β_6	-	-	-	-	-	-	0.060	-	-
β_7	-	-	-	-	0.081	0.067	0.059	0.064	0.053
β_8	-	-	-	-	-	-	-	0.056	0.045
β_9	-	-	-	-	-	-	0.063	0.058	-
β_{10}	-	-	-	0.081	-	-	-	0.068	0.056
β_{11}	-	-	-	0.079	0.075	0.060	0.059	0.063	0.050
β_{12}	-	-	-	-	0.076	0.061	-	-	-
β_{13}	-	-	0.070	0.085	-	-	0.060	-	-
β_{14}	0.085	0.085	-	0.074	-	-	0.063	0.063	0.053
β_{15}	0.073	0.073	0.063	-	0.083	0.070	0.062	0.061	0.054
β_{16}	-	-	-	-	-	-	0.064	0.062	0.051
β_{17}	-	-	-	-	-	0.056	-	0.058	0.049
β_{18}	-	-	0.079	0.070	0.068	0.057	0.062	-	0.049
r=0.60									
	n=50			n=100			n=200		
	LASSO	LAR	EN	LASSO	LAR	EN	LASSO	LAR	EN
β_1	-	-	-	-	-	-	-	-	-
β_2	0.103	-	0.087	-	-	-	0.079	0.079	0.074
β_3	-	-	-	-	-	-	0.080	0.080	0.073
β_4	-	-	-	-	0.087	0.079	0.072	0.071	0.067
β_5	-	-	-	-	-	0.080	-	-	-
β_6	-	-	-	0.096	0.096	-	-	-	-
β_7	-	-	-	0.092	0.092	0.082	0.079	0.079	0.073
β_8	-	-	-	0.093	0.092	0.083	0.078	0.078	0.074
β_9	-	-	-	0.095	0.094	0.086	0.074	0.074	0.069
β_{10}	-	-	-	-	-	-	0.078	0.078	-

β_{11}	-	-	-	0.090	0.088	0.079	-	-	-
β_{12}	-	-	0.094	-	-	-	0.074	0.073	0.068
β_{13}	0.109	0.107	0.090	0.091	0.089	0.078	0.074	0.074	0.069
β_{14}	-	-	-	0.092	-	-	0.078	0.078	0.072
β_{15}	-	-	-	-	-	-	0.077	0.077	0.071
β_{16}	-	-	-	-	-	-	-	-	0.072
β_{17}	0.110	0.107	-	-	-	-	0.076	0.076	0.071
β_{18}	-	0.108	-	-	-	-	0.077	0.075	0.070
r=0.90									
	n=50			**n=100**			**n=200**		
	LASSO	**LAR**	**EN**	**LASSO**	**LAR**	**EN**	**LASSO**	**LAR**	**EN**
β_1	-	-	-	-	-	-	-	-	-
β_2	-	-	-	0.094	-	0.078	0.079	0.079	0.072
β_3	-	-	-	0.093	0.092	0.079	0.076	0.076	0.069
β_4	-	-	-	-	-	0.080	0.082	0.082	0.072
β_5	-	-	-	0.089	-	-	0.077	-	0.068
β_6	-	-	-	-	-	-	0.079	0.079	0.071
β_7	-	-	-	-	-	-	0.078	0.076	-
β_8	-	-	-	-	-	-	-	-	0.070
β_9	0.117	-	0.090	0.096	0.094	0.081	0.081	0.080	0.072
β_{10}	-	-	-	-	-	-	-	0.077	0.070
β_{11}	-	0.111	0.089	-	-	-	-	-	-
β_{12}	-	-	-	-	-	-	0.077	0.075	0.069
β_{13}	-	-	-	-	-	-	0.086	0.079	0.071
β_{14}	-	-	-	0.091	0.090	0.078	-	-	-
β_{15}	-	-	-	-	0.089	-	0.079	0.078	0.071
β_{16}	0.115	0.110	0.093	0.094	0.093	0.081	0.076	0.076	-
β_{17}	0.119	0.117	-	-	0.091	0.077	0.074	0.073	0.068
β_{18}	-	-	-	0.090	0.088	-	0.080	0.080	0.074

LASSO: Least absolute shrinkage and selection operator; LAR: Least angle regression; EN: Elastic-net regression; β: Coefficient of regression; r: Coefficient of correlation; n: Sample size

For a correlation coefficient of 0.60 and n=50, LAR, LASSO, and EN methods included some common variables while differing in others to complete their models. For n=100, LAR included variable X_4 but excluded X_{14} compared to LASSO. Furthermore, differences were observed between LAR and EN for variables X_5 and X_6. The LASSO method, on the other hand, included X_{14} instead of X_4. For n=200, LAR and LASSO estimated the same model, while EN included variable X_{16}, differing from the other 2 methods.

For a correlation coefficient of 0.90 and n=50, LAR, LASSO, and EN methods included 3 variables in the model, with only X_{16} being commonly included. For n=100, EN, LAR, and LASSO included different variables in the model, although they calculated similar coefficient values for the included variables. For n=200, penalized regression methods produced similar model predictions.

When Table 3 is examined, with 20 variables, a correlation coefficient of 0.10, and n=50, LAR and LASSO included the same variables in the model, while EN included a different variable compared to the other two methods. For n=100, LAR and LASSO included the same variables in the model with identical regression coefficients.

EN, however, excluded variable X_{17} and included variables X_7 and X_{13}, thus completing the variable selection differently. For n=200, the penalized regression methods estimated the models using different variables and regression coefficients.

For a correlation coefficient of 0.60 and n=50, penalized regression methods included different variables in the model, thereby applying variable selection differently. For n=100, LASSO included variable X_{20} in the model, while LAR included X_{19} instead of X_{20}. EN included variable X_{14} but excluded X_9 compared to the other two methods. For n=200, penalized regression methods included the same variables in the model. While the regression coefficients were similar between LAR and LASSO, some differences were observed in the coefficients estimated by the EN method.

When the correlation coefficient is 0.90 and n=50, LAR, LASSO, and EN methods calculated regression coefficients for the variables in the model with similar values. For n=100, LAR and LASSO methods included the same variables in the model, while the EN method completed variable selection differently from the other methods. For n=200, the LAR method did not include the X_{13} variable in the model but added the X_{18} variable, unlike the LASSO method.

In Table 4, for 16 variables, a correlation coefficient of 0.10, and n=50, the LAR method achieved the highest R^2 value and the lowest error. For n=100, the R^2 values of the LAR and LASSO methods were identical. For n=200, the EN method lagged behind the other methods in model prediction performance. With 16 variables and a correlation coefficient of 0.60 across all sample sizes, the LAR method showed the highest model prediction performance, while the EN method was calculated to have lower model performance. For 16 variables, a correlation coefficient of 0.90, and n=50 or n=100, the LAR and LASSO methods predicted the same R^2 values. For n=200, the LAR method outperformed the others in model prediction performance. For 18 variables, a correlation coefficient of 0.10, and n=50, the LAR method provided the highest model prediction performance. For n=100 and n=200, the LAR and LASSO methods predicted identical results. For 18 variables, a correlation coefficient of 0.60, and n=50, the LAR method estimated a lower error, while the LASSO method performed better for n=100. For n=200, the LAR and LASSO methods provided identical predictions. With 18 variables, a correlation coefficient of 0.90, and n=50, the LAR method predicted the highest R^2 value and the lowest MSE, while the EN method underperformed compared to the other 2 methods. For n=100 and n=200, the LAR and LASSO methods achieved the same model prediction performance. For 20 variables, a correlation coefficient of 0.10, and n=50, the LAR method yielded the best results. For n=100 and n=200, the EN method predicted lower R^2 values and higher MSE compared to the LAR and LASSO methods. With 20 variables, a correlation coefficient of 0.60, and across all sample sizes, the LAR method was the best-performing method, while the EN method was the least successful. For 20 variables, a correlation coefficient of 0.90, and n=50 or n=100, the LAR method achieved the best model prediction performance. For n=200, the LAR and LASSO methods produced identical model predictions.

TABLE 3: Estimated model when the number of variables is 20 and the correlation coefficient (r) is 0.10, 0.60 and 0.90 respectively

	n=50			n=100			n=200		
	LASSO	LAR	EN	LASSO	LAR	EN	LASSO	LAR	EN
r=0.10									
$β_1$	-	-	-	-	-	-	-	-	-
$β_2$	-	-	-	-	-	-	-	0.059	-
$β_3$	-	-	-	0.067	0.067	0.058	0.056	0.058	0.052
$β_4$	-	-	-	0.071	0.071	0.057	-	0.063	-
$β_5$	-	-	0.070	-	-	-	0.059	-	-
$β_6$	-	-	-	0.071	0.071	-	0.058	0.061	0.053
$β_7$	-	-	-	-	-	0.056	0.061	-	0.050
$β_8$	-	-	-	-	-	-	0.060	0.061	0.050
$β_9$	-	-	-	-	-	-	0.057	0.058	0.047
$β_{10}$	0.090	0.090	0.072	0.067	0.067	0.056	0.058	0.059	0.049
$β_{11}$	-	-	-	-	-	-	0.061	0.059	0.048
$β_{12}$	-	-	-	0.070	0.070	0.060	-	-	0.050
$β_{13}$	-	-	-	-	-	0.059	0.059	0.060	0.049
$β_{14}$	-	-	-	0.078	0.078	0.064	-	-	0.045
$β_{15}$	-	-	-	-	-	-	0.062	0.060	0.047
$β_{16}$	-	-	-	-	-	-	-	0.061	0.052
$β_{17}$	-	-	-	0.075	0.075	-	0.062	0.054	-
$β_{18}$	-	-	-	-	-	-	0.058	-	0.048
$β_{19}$	0.098	0.095	0.078	-	-	-	0.060	0.066	-
$β_{20}$	0.072	0.070	-	-	-	-	-	-	-
r=0.60									
	n=50			n=100			n=200		
	LASSO	LAR	EN	LASSO	LAR	EN	LASSO	LAR	EN
$β_1$	-	-	-	-	-	-	-	-	-
$β_2$	-	-	-	-	-	-	0.075	0.074	0.067
$β_3$	0.108	0.104	0.089	0.086	0.085	0.075	0.073	0.072	0.065
$β_4$	-	-	-	0.085	0.083	0.077	0.073	0.073	0.068
$β_5$	-	-	0.088	-	-	-	0.071	0.071	0.065
$β_6$	-	-	-	-	-	-	0.069	0.069	0.064
$β_7$	-	-	-	0.089	0.089	0.078	-	-	-
$β_8$	-	-	-	-	-	-	-	-	-
$β_9$	-	-	-	0.083	0.082	-	0.071	0.070	0.064
$β_{10}$	-	-	-	-	-	-	0.074	0.074	0.069
$β_{11}$	-	-	-	-	-	-	0.073	0.072	0.067

β_{12}	-	-	-	0.093	0.091	0.080	-	-	-
β_{13}	-	-	-	0.091	0.090	0.079	0.070	0.070	0.065
β_{14}	-	-	-	-	-	0.080	-	-	-
β_{15}	-	0.098	-	-	-	-	-	-	-
β_{16}	0.101	0.110	-	-	-	-	0.072	0.072	0.066
β_{17}	0.113	-	0.093	-	-	-	0.076	0.076	0.070
β_{18}	-	-	-	-	-	-	0.076	0.076	0.069
β_{19}	-	-	-	-	0.088	0.078	0.072	0.071	0.065
β_{20}	-	-	-	0.090	-	-	-	-	-
r=0.90									
	n=50			**n=100**			**n=200**		
	LASSO	**LAR**	**EN**	**LASSO**	**LAR**	**EN**	**LASSO**	**LAR**	**EN**
β_{1}	-	-	-	-	-	-	-	-	-
β_{2}	-	-	0.083	-	-	-	0.076	0.075	0.067
β_{3}	-	-	-	-	-	-	0.075	0.074	0.066
β_{4}	-	-	-	0.096	0.094	-	0.071	0.071	0.065
β_{5}	-	-	-	-	-	-	-	-	-
β_{6}	-	-	-	-	-	-	0.074	0.074	0.066
β_{7}	-	-	-	0.092	0.092	0.078	0.071	0.071	0.064
β_{8}	-	-	-	-	-	-	-	-	-
β_{9}	0.108	0.105	-	-	-	-	0.075	0.074	0.067
β_{10}	-	-	-	0.086	0.085	-	-	-	-
β_{11}	-	-	-	-	-	-	0.074	0.073	0.065
β_{12}	-	0.106	0.085	0.091	0.087	0.077	-	-	0.065
β_{13}	0.113	-	-	0.084	0.085	0.072	0.073	-	-
β_{14}	-	-	-	-	-	0.075	-	-	-
β_{15}	-	-	-	-	-	0.070	0.074	0.073	0.066
β_{16}	-	-	-	-	-	-	0.072	0.072	0.065
β_{17}	-	-	-	-	-	-	0.072	0.071	0.064
β_{18}	-	-	0.094	0.086	0.084	0.073	-	0.071	0.064
β_{19}	0.110	0.107	-	0.089	0.087	0.075	0.073	0.072	0.065
β_{20}	-	-	-	-	-	-	0.074	0.073	-

LASSO: Least absolute shrinkage and selection operator; LAR: Least angle regression; EN: Elastic-net regression; β: Coefficient of regression; r: Coefficient of correlation; n: Sample size

TABLE 4: R^2 and MSE estimates of models estimated with different numbers of variables, different correlation coefficients and different sample sizes depending on the method used

			n=50			n=100			n=200		
			LASSO	**LAR**	**EN**	**LASSO**	**LAR**	**EN**	**LASSO**	**LAR**	**EN**
Number of Variables= 16	r=0.10	R^2	0.360±0.003	0.360±0.003	0.113±0.004	0.214±0.001	0.214±0.001	0.094±0.002	0.137±0.001	0.137±0.001	0.088±0.001
		MSE	0.622±0.004	0.621±0.004	0.861±0.006	0.783±0.003	0.783±0.003	0.902±0.004	0.857±0.003	0.857±0.003	0.906±0.003
	r=0.60	R^2	0.705±0.002	0.705±0.002	0.650±0.002	0.640±0.001	0.642±0.001	0.620±0.002	0.606±0.001	0.606±0.001	0.600±0.001
		MSE	0.282±0.002	0.281±0.002	0.333±0.002	0.353±0.001	0.352±0.001	0.372±0.001	0.389±0.001	0.389±0.001	0.396±0.001
	r=0.90	R^2	0.925±0.001	0.925±0.001	0.912±0.001	0.909±0.001	0.909±0.001	0.904±0.001	0.900±0.001	0.901±0.001	0.899±0.001
		MSE	0.070±0.001	0.070±0.001	0.083±0.001	0.088±0.001	0.088±0.001	0.093±0.001	0.097±0.001	0.097±0.001	0.098±0.001
Number of Variables= 18	r=0.10	R^2	0.412±0.003	0.412±0.003	0.131±0.005	0.239±0.002	0.239±0.002	0.107±0.002	0.148±0.001	0.148±0.001	0.093±0.001
		MSE	0.574±0.004	0.573±0.004	0.847±0.007	0.756±0.003	0.756±0.003	0.887±0.004	0.850±0.002	0.850±0.002	0.905±0.003
	r=0.60	R^2	0.733±0.002	0.733±0.002	0.666±0.002	0.652±0.001	0.651±0.001	0.627±0.001	0.617±0.001	0.617±0.001	0.608±0.001
		MSE	0.261±0.002	0.259±0.002	0.323±0.002	0.343±0.001	0.343±0.001	0.367±0.001	0.380±0.001	0.380±0.001	0.388±0.001
	r=0.90	R^2	0.930±0.001	0.931±0.001	0.914±0.001	0.912±0.001	0.912±0.001	0.905±0.001	0.903±0.001	0.903±0.001	0.901±0.001
		MSE	0.067±0.001	0.066±0.001	0.082±0.001	0.086±0.001	0.086±0.001	0.092±0.001	0.096±0.001	0.096±0.001	0.097±0.001
Number of Variables= 20	r=0.10	R^2	0.445±0.003	0.444±0.003	0.135±0.005	0.255±0.002	0.255±0.001	0.105±0.003	0.160±0.002	0.160±0.002	0.095±0.002
		MSE	0.546±0.004	0.544±0.004	0.846±0.007	0.741±0.002	0.741±0.002	0.889±0.004	0.840±0.002	0.840±0.002	0.905±0.003
	r=0.60	R^2	0.745±0.002	0.745±0.002	0.666±0.002	0.661±0.002	0.661±0.001	0.631±0.001	0.621±0.002	0.621±0.001	0.610±0.001
		MSE	0.246±0.002	0.243±0.002	0.317±0.002	0.331±0.001	0.330±0.001	0.360±0.001	0.375±0.001	0.374±0.001	0.384±0.001
	r=0.90	R^2	0.935±0.001	0.935±0.001	0.915±0.001	0.914±0.001	0.915±0.001	0.907±0.001	0.905±0.001	0.905±0.001	0.902±0.001
		MSE	0.062±0.001	0.061±0.001	0.080±0.001	0.083±0.002	0.082±0.001	0.089±0.001	0.094±0.001	0.094±0.001	0.096±0.001

LASSO: Least absolute shrinkage and selection operator; LAR: Least angle regression; EN: Elastic-net regression; R^2: Coefficients of determination; MSE: Mean square error; n: Sample size

DISCUSSION

In cases where multicollinearity, one of the assumptions of multiple linear regression, is present, the consistency and reliability of results obtained with OLS decrease. In other words, for data with multicollinearity issues, although OLS is an unbiased method, it calculates higher variances for estimators, leading to errors in the predicted model. Consequently, in such cases, biased methods have become popular alternatives to the unbiased OLS method.

Penalized regression methods have earned their place in the literature as biased methods. However, they minimize errors in model prediction without introducing significant bias. As the dimensions of today's datasets increase, identifying the variables that best explain the dependent variable among potentially many independent variables has become critically important.[17]

In clinical datasets, many variables exhibit a high degree of correlation with one another, which can lead to errors in treatment and diagnosis processes. Variable selection is one of the solutions used to address multicollinearity issues. Variable selection is crucial for improving model performance. Among the most commonly preferred methods for variable selection are forward selection, backward selection, and stepwise methods. However, the major drawback of these methods is that they do not account for the relationships among variables.[18] To address this problem, penalized regression methods such as LAR, LASSO, and EN are utilized. LAR and LASSO methods, in cases where the number of variables exceeds the sample size (p>n), include only as many variables in the model as the sample size (n variables). This limitation arises from the nature of convex optimization problems, preventing these methods from constructing accurate models under such conditions.[12,13,19]

The present study indicates that when the level of correlation among independent variables is low (r=0.10), all three methods exhibit lower model prediction performance. As the level of correlation increases, model prediction performance improves. Similarly, when comparing the methods with a fixed number of variables and correlation coefficient while increasing the sample size (n=50, 100, 200), the R^2 values decrease, and MSE values increase.

In the literature, similar studies highlight that LAR is the most successful method in predicting models when compared with LASSO and OLS in time series analyses involving variable selection.[20] In another study using a cholesterol dataset, LASSO and EN methods were compared, and LASSO was found to outperform EN in model prediction.[21] In another study using a diabetes dataset, LAR and LASSO methods were compared with the stepwise selection method. While LAR and LASSO methods provided the same R^2 values, the MSE value for LAR was found to be lower.[6] In a simulation study with 20 independent variables, a sample size of 100, and correlation coefficients of 0.7, 0.8, and 0.9, Ridge, LASSO, and EN methods were compared. It was observed that LASSO and EN selected different variables under certain conditions to construct the model, with EN providing better model predictions.[22] In another study with time series in the field of economics involving LAR, LASSO, and EN methods, these penalized regression methods were compared and it was concluded that the LAR method yielded more successful results than the other methods.[23] In a simulation study using 1000 iterations for each scenario, it was revealed that LASSO outperformed EN in scenarios involving small sample sizes and low correlation coefficients, medium sample sizes and moderate correlation coefficients, and large sample sizes and high correlation coefficients.[24] In another study, a comparison of LASSO and EN methods was conducted on a data set related to egg selection and the results revealed that although the same variables were included in the model, the LASSO method predicted a better model when the model when comparing R^2 values.[25]

CONCLUSION

This study demonstrates that the LAR method performs best under almost all conditions, including varying numbers of variables, correlation values, and sample sizes. In addition, it was determined that the LASSO method predicts the same model as the LAR method in some cases, while it predicts a model that falls behind the LAR method in other cases. The EN method, on the other hand, predicts a model that lags behind the other 2 methods in model prediction success. However, although the EN method lags behind the other 2 methods regarding model prediction success, its standalone model prediction success is high.

In light of these results, it can be concluded that penalized regression methods, LAR, LASSO, and EN, achieve model predictions with minimal error in datasets with multicollinearity problems and that these methods are particularly recommended for datasets with small sample sizes, where model prediction performance tends to be higher.

Source of Finance

During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.

Conflict of Interest

No conflicts of interest between the authors and/or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.

Authorship Contributions

Idea/Concept: *Ali Türker Çiftçi, Didem Derici Yıldırım, Damla Hazal Sucu;* ***Design:*** *Ali Türker Çiftçi, Didem Derici Yıldırım;* ***Control/Supervision:*** *Ali Türker Çiftçi, Didem Derici Yıldırım;* ***Data Collection and/or Processing:*** *Ali Türker Çiftçi;* ***Analysis and/or Interpretation:*** *Ali Türker Çiftçi, Didem Derici Yıldırım, Damla Hazal Sucu;* ***Literature Review:*** *Ali Türker Çiftçi;* ***Writing the Article:*** *Ali Türker Çiftçi;* ***Critical Review:*** *Ali Türker Çiftçi, Didem Derici Yıldırım, Damla Hazal Sucu.*

REFERENCES

1. Topal M, Eyduran E, Yağanoğlu AM, Sönmez AY, Keskin S. Çoklu doğrusal bağlantı durumunda ridge ve temel bileşenler regresyon analiz yöntemlerinin kullanımı [Use of ridge and principal component regression analysis methods in multicollinearity]. Atatürk Üniversitesi Ziraat Fakültesi Dergisi. 2010;41(1):53-7. [Link]
2. Albayrak AS. Çoklu doğrusal bağlantı halinde en küçük kareler tekniğinin alternatifi yanlı tahmin teknikleri ve bir uygulama. Uluslararası Yönetim İktisat İşletme Dergisi. 2005;1(1):105-26. [Link]
3. Büyükuysal MÇ, Öz İİ. Çoklu doğrusal bağıntı varlığında en küçük karelere alternatif yaklaşım: ridge regresyon. DÜ Sağlık bil Enst Derg. 2016;6(2):110-4. [Link]
4. Küçük A. Doğrusal regresyonda ridge, liu ve lasso tahmin edicileri üzerine bir çalışma [Yüksek lisans tezi]. Ankara: Hacettepe Üniversitesi Fen Bilimleri Enstitüsü; 2019. [Link]
5. Wu Y. Can't ridge regression perform variable selection? Technometrics. 2020;63(2):263-71. [Crossref]
6. Efron B, Hastie T, Johnstone I, Tibshrani R. Least angle regression. Institute Of Mathematical Statistics. 2004;32(2):407-99. [Link]
7. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning Data Mining. Inference and Prediction. 2nd ed. Berlin: Springer; 2008. p.73-4.
8. Kumarage PM, Yogarajah B, Ratnarajah N. Efficient Feature Selection for Prediction of Diabetic Using LASSO. 19th International Conference on Advances in ICT for Emerging Regions (ICTer); Colombo, Sri Lanka. 2019. p.1-7. [Crossref]
9. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of The Royal Statistical Society. 1994;58(1):266-88. [Link]
10. Reid S, Tibshirani R, Friedman J. A study of error variance estimation in lasso regression. Institute Of Statistical Science. 2016;26(1):35-67. [Link]
11. Hans C. Elastic net regression modeling with the orthant normal prior. Journal of the American Statistical Association. 2011;106(496);1383-93. [Link]
12. Zou H, Hastie T. Regularization and variable selection via the elastic net. Journal Of The Royal Statistical Society B. 2005;67(2):301-20. [Link]
13. Jia J, Yu B. On model selection consistency of the en when p>n. Institute Of Statistical Science. 2010;20(2):565-611. [Link]
14. Ogutu JO, Schulz-Streeck T, Piepho HP. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. BMC Proceedings. 2012;6(Suppl 2):1-6. [Crossref]
15. De Mol C, De Vito E, Rosasco L. Elastic-net regularization in learning theory. Journal Of Complexity. 2009;25(2):201-30. [Link]
16. Alpar R. Uygulamalı Çok Değişkenli İstatistiksel Yöntemler. 5. Baskı. Ankara: Detay Yayıncılık; 2017. p.400-1.
17. Farbahari A, Dehesh T, Gozashti MH. The usage of lasso, ridge and linear regression to explore the most influential metabolic variables that affect fasting blood sugar in type 2 diabetes patients. Rom. J Diabetes Nutr Metab Dis. 2020;26(4):371-9. [Crossref]
18. Yıldırım DD, Çiftçi AT. Etkili değişkenlerin cezalı regresyon yöntemleri ile belirlenmesi: diyabet veri kümesi üzerine bir uygulama [Determining the effective variables by penalized regression methods: An application on diabetes data set]. Mersin Univ Saglık Bilim Derg. 2020;14(1);105-12. [Link]
19. Zou H, Hastie T. Regression shrinkage and selection via the elastic net, with applications to microarrays. Journal of the Royals Statistical Society. 2003;67:301-20. [Link]
20. Itubide E, Cerda J, Graff M. A comparasion between LARS and LASSO for initialising the time-series forecasting auto-regressive equations. Procedia Technology. 2013;7:282-8. [Link]
21. Gregory DS, Jackson HM. Regulation techniques for multicollinearity: lasso, ridge and elastic nets. 2018;131. [Link]
22. Kayanan M, Wijekoon P. Performance of LASSO and elastic net estimators in misspecified regression model. Ceylon Journal of Science. 2019;48(3):293-9. [Crossref]
23. Bai J, Ng S. Forecasting economic time series using targated predictors. Journal of Econometrics. 2008;146:304-17. [Link]
24. Oyeyemi GM, Ogunjobi EO, Folorunsho AI. On performance of shrinkage methods-a monte carlo study. International Journal of Statistics and Applications. 2015;5(2):72-6. [Link]
25. Çiftsüren MN, Akkol S. Prediction of internal egg quality characterictics and variable selection using regularizations methods: ridge, LASSO and elastic net. Archives Animal. Breeding. 2018;61(3):279-84. [Link]