QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

A Comparison between Some Penalized Methods for Estimating Parameters (Simulation Study)

Dr. Dlshad Mahmood Saleh
Department of Statistics and Informatics, College of Administration and Economics, University of Salahaddin, Erbil, Kurdistan Region, Iraq
dlshad.saleh@su.edu.krd

Dr. Dler Hussein Kadir
Department of Statistics and Informatics, College of Administration and Economics, University of Salahaddin, Erbil, Kurdistan Region, Iraq
Department of Business of Administration, Cihan University-Erbil, Kurdistan region, Iraq.
dler.kadir@su.edu.krd

Lect. Dashty Ismil Jamil
Department of Marketing, College of Administration and Economics, Lebanese French University, Erbil, Kurdistan Region, Iraq
Dashty@lfu.edu.krd

ARTICLE INFO

Article History:
Received: 1/10/2022
Accepted: 8/11/2022
Published: Spring 2023

Keywords:
Penalized Method; Ridge Regression; Elastic-Net Regression; Elastic-Net; Bridge.

Doi:
10.25212/lfu.qzj.8.1.44

ABSTRACT

Regression analysis is one of the most popular statistical methods in various biological and economic studies where, frequently, the number of explanatory variables becomes large. Penalized methods have been adapted and have gained popularity as a key for simultaneously performing variable selection and model estimation. This paper proposes contamination procedure from the viewpoint of different types of penalized regression, aiming at identifying any types of penalized methods that are best to deal with contamination data. This paper demonstrates that the Lasso regression is the best method for contamination data depending on the heavy tail distribution behavior of the response variables and using simulation for (15%) data with contamination. The comparison between types of penalized methods based on the statistical criterion (MAE and MSE) and

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

results shows that the Lasso regression is better than another type of penalized method.

1. Introduction

The penalized least squares method has been repeatedly shown to be an appealing regression shrinkage and selection method. This process differs from standard approaches to variable selection in that it identifies significant variables while also estimating regression coefficients. The estimators produced are as efficient as the Oracle estimator. Furthermore, non-significant variables are eliminated by estimating their coefficients as recent related research includes (Van der Kooij, A.J., 2007).

Multiple regression is often used to estimate a model for predicting future responses, or to investigate the relationship between the response variable and the predictor variables. For the first goal the prediction accuracy of the model is important, for the second goal the complexity of the model is of more interest. Ordinary least squares (OLS) regression is known for often not performing well with respect to both prediction accuracy and model complexity. Several regularized regression methods were developed the last few decades to overcome these flaws of OLS regression, starting with Ridge regression (Hoerl and Kennard 1970a,b), followed by Bridge regression (Frank and Friedman 1993), the Garotte (Breiman 1995), and the Lasso (Tibshirani 1996), and more recently LARS (Efron, Hastie, Johnstone, and Tibshirani 2004), Pathseeker (Friedman and Popescu 2004), and the Elastic Net (Zou and Hastie 2005).

OLS regression may result in highly variable estimates of the regression coefficients in the presence of collinearity or when the number of predictors (k) is large relative to the number of observations (N). Ridge regression reduces this variability by shrinking the coefficients, resulting in more prediction accuracy at the cost of usually only a small increase of bias. In Ridge regression, the coefficients are shrunken towards zero, but will never become exactly zero. So, when the number of predictors is large, Ridge regression will not provide a sparse model that is easy to interpret. Subset selection, on the other hand, does provide interpretable models but does not reduce the variability of the estimates of the coefficients. While not reducing the variability of the coefficient estimates of the selected variables, subset selection can reduce the variability of the prediction estimates, but not as much as Ridge regression or the Lasso. The Lasso was developed by Tibshirani (1996) to improve both predition accuracy and model interpretability by combining the nice features In this study; penalized methods with wavelet shrinkage are proposed for effectively handling of these issues.

2. Penalized Methods:

Penal methods have appeared in recent years and have gained wide popularity among statisticians, as these methods are an important key to performing the selection of variables and estimating parameters simultaneously; so many penalty methods have been proposed through which a penalty constraint is added to the regression models (Tutz, G. and Ulbricht, J., 2009). The goal of adding the penalty restriction is to control the complexity of the model and provide a criterion for the selection of variables, by introducing some restrictions on the transactions that impose on some transactions that their value is equal to zero (Helwig, N.E., 2017).
The penalty constraint quantity works to balance the variance and bias in the chosen model. When the penalty amount is small, a larger number of explanatory variables are selected with a small bias, but the variance will be large, on the contrary, a large penalty amount causes few explanatory variables to be selected with a large bias but the variance will be lower. Therefore, a good choice of penalty amount leads to improving the prediction accuracy and ease of understanding and interpretation of the model,
In general, it is known as Penalized Linear Regression (PLR); as follows:

$$\mathrm{PLR}(\beta;\lambda) = (Y - X\beta)^{T}(Y - X\beta) + \lambda \sum_{j=1}^{p} P_{\lambda}(|\beta_j|) \qquad (1)$$

where the amount $P_{\lambda}(|\beta_j|)$ represents the penalty term, which is a function of coefficients, and (λ) represents the tuning parameter, since $(\lambda \geq 0)$, and that the penalty limit depends entirely on the value of (λ) as it controls the amount of shrinkage of parameter values. When the value is $(\lambda = 0)$ then we get the estimations of the Ordinary Least Squares method (OLS). Conversely, as the value of

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

(λ) increases, the number of variables excluded from the model will increase (Wood, Simon., 2006).

In partial linear regression, estimates of the model parameters are found using this equation:

$$\hat{\beta}_{PLR} = \operatorname{argmin}(Y - X\beta)^T(Y - X\beta) + \lambda \sum_{j=1}^{p} P_\lambda(|\beta_j|) \qquad (2)$$

The two researchers (2001) (Jianging Fan and Li) suggested that a good penalty term should produce an estimator that has three properties, first, (un-biasedness) when the variable is unbiased for large real parameters. Second, (sparsity) makes small estimators exactly zero. Finally, the estimated continuity is (continuous) in the data to avoid instability in the model prediction.

There are many penalized methods that have been proposed and their characteristics studied, including Ridge, Least Absolute Shrinkage and Selection Operator (LASSO), Elastic-Net, Bridge and other methods.

2.1. Ridge Regression:

Regression modeling with associated explanatory variables presents a challenging problem when selecting variables and estimating parameters. The reason for this is, in the case of multicollinearity, the data matrix does not have enough information to distinguish the effect of a correlated variable versus a variable another related. In choosing a variable, selection methods tend to choose arbitrarily for one of the variables associated and does not take into account the significance of the specified variable. In addition, the existence of plurality linearity impairs prediction accuracy by amplifying the variance of parameter estimates, which may lead to removing significant coefficients from the model (Van der Kooij, A.J., 2007).

The ridge regression method was proposed by (Hoerl and Kennard) (1970) and it is considered one of the oldest penalty methods, as it received great attention because of its ability to overcome the problem of multicollinearity without removing the explanatory variables from the regression model. The Ridge Regression method reduces the variance in the coefficient estimates by adding a penalty quantity that

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

follows the rule (L2 - norm) to the sum of the squares of the residuals, as the penalty quantity reduces the regression coefficients.
Penalty linear regression is defined using the ridge term as follows:

$$PLR(\beta;\lambda)^{Ridge} = (Y - X\beta)^T (Y - X\beta) + \lambda \sum_{j=1}^{p} \beta_j^2 \qquad (3)$$

So that the penalty term $(\sum_{j=1}^{p} \beta_j^2)$ represents the (L2 - norm) rule the estimates of the parameters in the penalty regression model can be obtained from equation (1.7) as follows:

$$\hat{\beta}_{PLR}^{Ridge} = (X^T X + \lambda I)^{-1} X^T y \qquad (4)$$

Since I is the identity matrix with capacity P and λ is the positive shrinkage parameter, adding λI to the main diameter elements in the $(X^T X)$ information matrix reduces the variance of the OLS estimates with the addition of an amount of bias to it.
In ridge regression, the coefficients are gradually reduced towards zero, but they do not make them equal to zero at all, and then all the variables remain in the model, as a result, it is not possible in the ridge regression method to choose the variables and therefore the resulting linear regression model cannot be easily explained, especially if the number of Large explanatory variables.

2.2. Lasso Regression (Least Absolute Shrinkage and Selection Operator):
The loss functions for the Lasso can be viewed as constrained versions of the ordinary least squares (OLS) regression loss function. In Lasso Regression constrains the sum of the absolute values of the coefficients as follows (Van der Kooij., 2007):

$$L^{lasso}(\beta_1, \dots, \beta_P) = \left\| y - \sum_{j=1}^{P} \beta_j X_j \right\|^2, \text{ subject to} \sum_{j=1}^{P} |\beta_j| \leq t_1 \qquad (5)$$

with N the number of observations, P the number of predictor variables,

β_j, (j = 1, . . ., P), the regression coefficients, and t_1 the Lasso tuning parameter, and where $\|.\|^2$ denotes the squared Euclidean norm.
This constrains loss functions can also be written as penalized loss functions:

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

$$L^{lasso}(\beta_1, \dots, \beta_P) = \left\|y - \sum_{j=1}^{P} \beta_j X_j\right\|^2 + \lambda_1 \sum_{j=1}^{P} \text{sign}(\beta_j)\beta_j \quad (6)$$

the with λ_1 the Lasso penalty, penalizing the sum of the absolute values of the regression coefficients. In matrix notation, the penalized loss functions are written as:

$$L^{lasso}(\beta_1, \dots, \beta_P) = \|y - X\beta\|^2 + \lambda_1 w^T \beta \quad (7)$$

Where:

The elements wj of (w) are either +1 or −1, depending on the sign of the corresponding regression coefficientβ_j.

Minimizati on of the constrained loss function is more complicated. The regression coefficients are estimated as

$$\beta^{lasso}(\beta_1, \dots, \beta_P) = (X^T X)^{-1}(X^T y + \frac{\lambda_1}{2} w) \quad (8)$$

2.3. Elastic Net Regression:

Elastic net regression combines the penalty terms of ridge and lasso regression. When fitting models with elastic net, we minimize the function.

Zou and Hastie (2005) have proposed the Elastic Net and developed an algorithm, called LARS-EN, based on the efficient LARS algorithm, to overcome the Lasso limitations of selecting at most N predictors and of selecting only one predictor from a group of highly correlated predictors. For the Elastic-Net the regression coefficients are estimated as

$$\hat{\beta}_{PLR}^{Enet} = (X^T X + \lambda_2 I)^{-1} \left(X^T y - \frac{\lambda_1}{2} \text{sign}(B_j^{OLS})\right) \quad (9)$$

2.4. Bridge Regression:

Bridge regression is a broad class of the penalized regression method proposed by Frank and Friedman (1993). The bridge estimate can be obtained by minimizing.

$$\hat{\beta}_{PLR} = \text{argmin}\left\{\sum_{i=1}^{n}\left(Y_i - X_i^T \beta\right)^2 + \lambda \sum_{j=1}^{p} \left|\beta_j\right|^q\right\} \quad (10)$$

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

It does variable selection when $0 < q \leq 1$, and shrinks the coefficients when $q > 1$. Frank and Friedman (1993) did not solve for the estimator of bridge regression for any given $q > 0$, but they pointed out that it is desirable to optimize the parameter q.

Fu (1998) studied the structure of bridge estimators and proposed a general algorithm to solve for $q \geq 1$. The shrinkage parameter q and the tuning parameter λ are selected via generalized cross-validation. Knight and Fu (2000) showed asymptotic properties of bridge estimators with $q>0$ when p is fixed. Huang et al. (2008) studied the asymptotic properties of bridge estimators in sparse, high-dimensional, linear regression models when the number of covariates p may increase along with the sample size n. introduced an Lq support vector machine algorithm that selects q from the data. The effect of the Lq penalty with different q's, and we briefly mention some parts of it here along with the effect of the elastic net.

Despite the flexibility of bridge estimators, the non-convexity of the penalty function may reduce the practical use of the estimators. In order to avoid the non-convex optimization problem, we introduce two algorithms for solving bridge regression. The first method applies the local quadratic approximation (LQA) suggested by Fan and Li (2001) and the second applies the local linear approximation (LLA) suggested by Zou and Li (2008).

For the LLA, we use one-step estimates proposed by Zou and Li (2008), which automatically adopts a sparse representation. The one-step bridge estimator for $0 < q < 1$ is obtained as follows.

Define $x_{ij}^{*} = (\sqrt{2}\left|\beta_{0j}\right|^{1-q} x_{ij})/q$ and $Y_i^{*} = \sqrt{2} x_i^{T} \beta_0$. Using$(x_{ij}^{*}, Y_i^{*})$, we apply the LARS algorithm (Efron et al., 2004) to solve

$$\beta_{PLR}^{\wedge\, Bridge} = \operatorname{argmin}\left\{\sum_{i=1}^{n}\left(Y_i^{*} - X_i^{*T}\beta\right)^2 + \lambda \sum_{j=1}^{p}\left|\beta_j\right|\right\} \qquad (11)$$

Then

$$\beta_{1j} = (\beta_j^{\wedge *}\left|\beta_{0j}\right|^{1-q})/q.$$

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

Zou and Li (2008) showed that the LLA is the best convex MM algorithm, which proves the convergence of the LLA algorithm. The LLA naturally produces sparse estimates without iterations, which also reduces the computational burden.

3. Heavy-Tailed Distributions

In probability theory, a heavy-tailed distribution is one that has heavier tails than the exponential distribution and whose tails are not exponentially constrained. Although a distribution may have a heavy left tail, a heavy right tail, or both tails, it is often the right tail of the distribution that is of importance in applications. The long-tailed distributions and the sub exponential distributions are two significant subclasses of heavy-tailed distributions. Practically speaking, any heavy-tailed distribution that is often utilized belongs to the sub exponential class.

3.1 Student's t-distribution

Statistical scientist Gosset was the first to introduce in 1908 under the name Student. Later, fisher made in 1920 additions to this distribution. The probability density function of the Student's t-distribution can be written as follows: (Härdle & Simar, 2007)

$$f(x\;;\nu)=\left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\;\Gamma\left(\frac{\nu}{2}\right)\left(1+\frac{x^{2}}{\nu}\right)^{\frac{\nu+1}{2}}}\right) \qquad (12)$$

Since ν is a parameter of the distribution, which represents the number of degrees of freedom and $-\infty < x < \infty$.

4. Application Part:

This Part included a practical comparison of the methodologies employed in the estimation process represented by Types of penalized methods. The relative efficiency, which is represented by the mean square of error and mean absolute error,

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

was determined to present with a review of the most essential strategy of regularization for coefficients regression.

4.1. Simulation Study:

To implement the simulation experiments, different levels of the following factors were used sample sizes n, Where three sample sizes were used, namely, the simulation experiment included many cases, as three sizes of samples were used, which are (50, 150, 300) when the number of parameters (k) is equal to(15 and 40), and we contaminate (15) of (e_i) vector without modifying explanatory variables such that this contaminated values can cause outliers. Here original (e_i) values are taken from a standard normal distribution with (zero mean and standard deviation equal to 2 and 6) and generated (15%) values from the Student t distribution. These values produce contaminate the data by using this formula $f(x) = (1 - p) * f_1(x) + p * f_2(x)$. The explanatory variables are independent of a normal distribution (with a mean equal to zero and a standard deviation equal to one). When the number of parameters (k) is equal to (7 2 0 -0.5 0 0 0 3 5 0 0 0 0 0 0) where b =5 are numbers of non- zero parameters, and the second case (K) equal to (2 4 0 -6 0 3 0 1 0 0.5 0 -8 5 0 3 -0.5 -2 6 9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0) where b =13 are numbers of non-zero parameters. For the frequency of (1000) iterations of the assumed regression model and each of the cases shown in tables (1, 2, 3, 4) a comparison was made between the methods used in the estimation process represented by Penalized methods (Ridge, Lasso, Elastic-Net, and Bridge) and parameters can now be defined for the default model. The comparison was made to calculate the relative efficiency, which represents the average mean square of error (AMSE) and average mean absolute error (AMAE).

Case 1: When $\lambda = 0.5$

Table 1: The average (MAE and MSE) values for types of penalized methods.

k= 15	Penalized Method	(15% Contaminate)					
		n =50		n =150		n =300	
		AMAE	AMSE	AMAE	AMSE	AMAE	AMSE
	Ridge	2.5439	16.7695	2.7536	16.6204	2.7941	16.7946

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

σ = 2	Lasso	2.5237	16.5363	2.7524	16.5994	2.7742	16.7897
	Elastic-Net	2.5426	16.7734	2.7532	16.6207	2.7939	16.7946
	Bridge	2.5313	16.6063	2.7527	16.6035	2.7781	16.7905
σ = 6	Ridge	4.6490	49.4381	5.2249	49.2199	5.3382	48.6983
	Lasso	4.6350	49.1467	5.2238	49.1985	5.3380	48.6934
	Elastic-Net	4.6490	49.4440	5.2249	49.2203	5.3382	48.6984
	Bridge	4.6392	49.2316	5.2241	49.2030	5.3381	48.6944

Table 2: The average (MAE and MSE) values for types of penalized methods.

k=40	Penalized Method	(15% Contaminate)					
		n =50		n =150		n =300	
		AMAE	AMSE	AMAE	AMSE	AMAE	AMSE
σ = 2	Ridge	1.5124	21.2100	2.5988	16.9480	2.7357	16.8297
	Lasso	1.3318	16.7809	2.7597	16.7570	2.7318	16.7892
	Elastic-Net	1.5165	21.3507	2.5981	16.9497	2.7353	16.8301
	Bridge	1.7844	29.0506	2.5858	16.8084	2.7328	16.7977
σ = 6	Ridge	2.4409	54.2406	4.7096	49.7400	5.1204	49.0876
	Lasso	2.3030	48.6127	4.7001	49.5415	5.1184	49.0469
	Elastic-Net	2.4451	54.4177	4.7095	49.7418	5.1204	49.0879
	Bridge	2.6496	62.9910	4.7030	49.5981	5.1188	49.0560

Case 2: When $\lambda = 3$

Table 3: The average (MAE and MSE) values for types of penalized methods.

k=15	Penalized Method	(15% Contaminate)					
		n =50		n =150		n =300	
		AMAE	AMSE	AMAE	AMSE	AMAE	AMSE
	Ridge	3.0676	23.0318	2.8866	17.8551	2.8315	17.0891

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

σ = 2	Lasso	2.5437	16.8662	2.7508	16.6458	2.7907	16.7301
	Elastic-Net	3.0702	19.5921	2.8861	17.8729	2.8310	17.0393
	Bridge	2.7494	19.9446	2.7691	16.7994	2.9520	16.7628
σ = 6	Ridge	4.9252	55.6865	5.2509	49.6576	5.3723	49.2280
	Lasso	4.6079	49.1007	5.1834	48.4282	5.3517	48.2280
	Elastic-Net	4.9297	55.8234	5.2516	49.6777	5.3725	49.2332
	Bridge	4.7605	52.1637	5.7674	48.5926	5.3538	48.9019

Table 4: The average (MAE and MSE) values for types of penalized methods.

k=40	Penalized Method	(15% Contaminate)					
		n =50		n =150		n =300	
		AMAE	AMSE	AMAE	AMSE	AMAE	AMSE
σ = 2	Ridge	1.6423	24.7739	2.7778	18.9710	2.8250	17.6759
	Lasso	1.4693	20.0937	2.5757	16.8293	2.7377	16.8923
	Elastic-Net	1.7398	27.7240	2.7754	18.9965	2.8235	17.6833
	Bridge	1.8838	27.8878	2.7565	18.7007	2.7733	17.2022
σ = 6	Ridge	2.5822	60.7328	4.8017	50.8248	5.1444	49.4874
	Lasso	2.4076	52.86564	4.6909	48.5935	5.0998	48.6862
	Elastic-Net	2.6518	63.8661	4.8022	50.8509	5.1445	49.4950
	Bridge	4.1129	67.9239	4.7950	50.6633	5.1188	49.0183

5. Conclusion:

1- Through Simulation study reached results the Lasso regression is better than another type of penalized method according to the criterion of (MAE) and (MSE).

2- Increased values of (MAE) when increasing sample size and standard deviation.

References

1- Breiman, L., 1995. Better subset regression using the nonnegative garrote. Technometrics, 37(4), pp.373-384.

2- Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R., 2004. Least angle regression. The Annals of statistics, 32(2), pp.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(456), pp.1348-1360.
4- Frank, L.E. and Friedman, J.H., 1993. A statistical view of some chemometrics regression tools. Technometrics, 35(2), pp.109-135.
5- Friedman, J. and Popescu, B.E., 2004. Gradient directed regularization. Unpublished manuscript, http://www-stat. Stanford.
6- Fu, W. and Knight, K., 2000. Asymptotic for lasso-type estimators. The Annals of statistics, 28(5), pp.1356-1378.
7- Fu, W.J., 1998. Penalized regressions: the bridge versus the lasso. Journal of computational and graphical statistics, 7(3), pp.397-416.
8- Helwig, N.E., 2017. Regression with ordered predictors via ordinal smoothing splines. Frontiers in Applied Mathematics and Statistics, 3, p.15.
9- Hoerl, A.E. and Kennard, R.W., 1970. Ridge regression: applications to non-orthogonal problems. Technometrics, 12(1), pp.69-82.
10- Hoerl, A.E. and Kennard, R.W., 1970. Ridge regression: Biased estimation for non-orthogonal problems. Technometrics, 12(1), pp.55-67.
11- Huang, J., Ma, S. and Zhang, C.H., 2008. Adaptive Lasso for sparse high-dimensional regression models. Statistica Sinica, pp.1603-1618.
12- Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), pp.267-288.
13- Tutz, G. and Ulbricht, J., 2009. Penalized regression with correlation-based penalty. Statistics and Computing, 19(3), pp.239-253.
14- Van der Kooij, A.J., 2007. Prediction accuracy and stability of regression with optimal scaling transformations. Leiden University.
15- Wood, S.N., 2006. Generalized additive models: an introduction with R. chapman and hall/CRC.
16- Zou, H. and Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology), 67(2), pp.301-320.
17- Zou, H. and Li, R., 2008. One-step sparse estimates in non-concave penalized likelihood models. The Annals of statistics, 36(4), pp.1509-1533.

QALAAI ZANISTSCIENTIFIC JOURNAL
A Scientific Quarterly Refereed Journal Issued by Lebanese French University – Erbil, Kurdistan, Iraq
Vol. (8), No (1), Spring 2023
ISSN 2518-6566 (Online) - ISSN 2518-6558 (Print)

بەراوردێک لە نێوان هەندێک شێوازی سزادراو بۆ خەملاندنی پارامێتەرەکان (داتاهاوشێوەکردن)

پـوخـتـە:

چاودێریکردن تەکنیکێکە بۆ پشکنینی سەقامگیری پەیوەندییە کارایییەکانی نێوان گۆڕاوێکی وەڵامدانەوە و یەکێک یان چەند گۆڕاوە ڕوونکردنەوەیەک بە تێپەڕبوونی کات. بوونی پیسبوون کاریگەری نەرێنی جددی لەسەر مۆدێلکردن، چاودێریکردن و دەستنیشانکردنی داتاکان هەیە. ئەم توێژینەوەیە ڕێکارێکی نوێی پیسبوون لە ڕوانگەی جۆرە جیاوازەکانی پاشەکشەی سزادراوەوە پێشنیار دەکات، بە ئامانجی دیاریکردنی هەر جۆرە شێوازێکی سزادراو کە باشترینە بۆ مامەڵەکردن لەگەڵ زانیارییەکانی پیسبوون. ئەم توێژینەوەیە نیشان دەدات کە پاشەکشەی لاسۆ باشترین ڕێگایە بۆ داتاکانی پیسبوون بەپێی هەڵسوکەوتی دابەشکردنی کلکی قورس لە گۆڕاوە وەڵامدەرەکان و بەکارهێنانی تاقیکردنەوەکانی هاوشێوەکردن بۆ (15%) داتاکان لەگەڵ پیسبوون. بەراوردکردنی نێوان جۆرەکانی شێوازی سزادراو لەسەر بنەمای پێوەرە ئامارییەکان (MAE و MSE) و ئەنجامەکان دەریدەخات کە پاشەکشەی لاسۆ باشترە لە جۆرێکی تری شێوازی سزادراو.

مقارنة بين بعض طرائق الجزائي لتقدير المعاملات (دراسة محاكاة)

الملخص:

المراقبة هي تقنية للتحقق من استقرار العلاقات الوظيفية بين متغير استجابة ومتغير أو أكثر من المتغيرات التوضيحية بمرور الوقت. إن وجود التلوث له آثار ضارة خطيرة على نمذجة البيانات ومراقبتها وتشخيصها. تقترح هذه الورقة إجراءً جديدًا للتلوث من وجهة نظر أنواع مختلفة من الانحدار الجزائي ، بهدف تحديد أي أنواع من طرائق الجزائي الأفضل للتعامل مع بيانات التلوث. يوضح هذا البحث أن الانحدار الجزاء هو أفضل طريقة لبيانات التلوث اعتمادًا على سلوك توزيع الذيل الثقيل لمتغيرات الاستجابة واستخدام تجارب المحاكاة لبيانات (15٪) مع التلوث. تُظهر المقارنة بين أنواع طرائق جزائي عليها بناءً على المعيار الإحصائي (MAE و MSE) والنتائج أن انحدار Lasso أفضل من أي نوع آخر من اطرائق اخرى.