Studying the Performance of Two Ridge Estimators Using Least Absolute Deviation

Tamarah Wathib Mohammad[1], Dr.Awatif Rezzoky Al-Dubaicy[2]

[1] Department Of Mathematics, Education College, Al-Mustansiriyah University, Baghdad ,Iraq

1 wathibtamarah@uomustansiriyah.edu.iq
[2]draraldubaicy@yahoo.com

A B S T R A C T

Multicollinearity is one of the essential and implicit problems in the regression analysis due to its influence on the model estimators. The problem is the independent variables are highly correlated, and the regression results are unclear. The purpose of this paper is to solve this problem using one of the solutions available, one of these solutions is the ridge regression of Least Absolute Deviation (LAD) estimators through adding a suggested ridge parameter as modify ridge parameter of (Hoerl et al. (1975)) say ($\widehat{K}_{HKB}$). A simulation study was performed to compare ($\widehat{K}_{HKB}$) and the suggested ridge parameter using Mean Square Error (MSE) to determine the best one.

Keywords: Multicollinearity problem , Linear Regression , Ridge Regression , Least Absolute Deviation, Variance Inflation Factor.

دراسة أداء اثنين من مقدرات الحرف باستخدام أقل انحراف مطلق

تمارة واثب محمد د. عواطف رزوقي الدبيسي

الجامعة المستنصرية – كلية التربية– قسم الرياضيات

الخلاصه

تعد العلاقة الخطية المتعددة إحدى المشكلات الأساسية والضمنية في تحليل الانحدار نظرًا لتأثيرها على مقدرات النموذج. المشكلة هي أن المتغيرات المستقلة مترابطة بشكل كبير، ونتائج الانحدار غير واضحة. الغرض من هذا البحث هو حل هذه المشكلة باستخدام أحد الحلول المتاحة، وأحد هذه الحلول هو الانحدار الحرف لمقدرات الاقل انحراف مطلق (LAD) من خلال إضافة معلمة الحرف المقترح كتعديل معلمة الحرف لـ (Hoerl et al. (١٩٧٥)) تدعى ($\widehat{K}_{HKB}$). تم إجراء دراسة محاكاة لمقارنة ($\widehat{K}_{HKB}$) ومعلمة الحرف المقترحة باستخدام متوسط الخطأ التربيعي (MSE) لتحديد أفضلها.

الكلمات المفتاحيه: مشكلة التعدد الخطي ,الانحدار الخطي, أنحدار الحرف, أقل انحراف مطلق, عامل تضخم.

1. Introduction

The multiple linear regression model, which is contains several independent variables and one dependent variable, where the explanatory variables are normally assumed to be independent. In actuality, however, the explanatory variables may have strong or almost strong linear connections. The independence assumptions are no longer applicable in this instance, resulting in the multicollinearity issue. [12], one of the regression issues is that the independent variables are highly correlated, making the regression results unclear [6], as a result, it is impossible to estimate the singular impacts of various variables in the regression equation since multicollinearity can be defined as a situation in which two or more independent variables move simultaneously. As a result, it's difficult to determine which independent variables are responsible for the observed change in the dependent variable [1], the definition of the regression coefficient is the change in the independent variable leading to a change in the dependent variable. The issue may be very difficult to discover, it is not a specifications error that can be discovered through the check from the regression residual that it is actually a modeling error that it is a case of imperfect data [6].

There are several approaches to solving this problem, the most popular of which being Ridge regression, which has several real-world benefits, which depends on the ridge parameter (k), where many scientists have worked to estimate the ridge parameter and the first to estimate the ridge parameter by Ordinary Least Squares Estimators (OLS) method It is Hoerl and Kennard (1970) [9] where Ordinary Ridge Regression Estimators (OLS-ridge) by allowing some bias to be introduced into the estimations of the regression coefficients.

Pfaffenber and Dielman (1989) [17] Investigated the performance of (LAD) method and ridge regression separate, both are robust; however, each is better appropriate for a particular type of problem when combined in a single estimation procedure. They attempted to estimate the ridge parameter (k) using estimators (LAD) of each one of the parameters and variance of the errors terms, that is, they used the same biasing parameters as Hoerl et al. (1975) [10]. Still, as an alternative to the variance of error terms and estimating parameters in the formula using (OLS) method, they used the (LAD) estimations. When the independent variables are highly collinear, and the error terms are asymmetric or own a large tail, Least Absolute Deviation Ridge method extracts additional bias space to reduce the Mean Squared Error (MSE) of the model (LAD) estimators [20].

Several researchers have suggested different ridge parameters from them: Hoerl and Kennard (1970) [9], Hoerl et al. (1975) [10], Lawless and Wang (1976) [13], Kibria (2003) [12], Dorugade and Kashid (2010) [5], Muniz et al. (2012) [15] and others.

In this paper, get a new ridge parameter was proposed say ($\widehat{K}_D$), represents ridge parameter Hoerl et al. ($\widehat{K}_{HKB}$) [10] with subtracting the variance inflation factor, then using simulation, compare two estimators to determine which is the best by obtaining the least Mean Square Error (MSE). The following is how the paper is structured: In section (2), we addressed the concept of multicollinearity. The primary ideas for the OLS and the OLS ridge were explained in section (3). Primary ideas for the (LAD) and the (LAD-ridge) have been explained in Section (4), as well as the ridge parameters. Section (5) proposes a ridge parameter. The efficiency of the ridge parameters was checked in section (6), and the results were assessed using the simulation approach to compare these parameters, the result has been analysis.

2. Multicollinearity Problem

The multiple linear regression model is one of the most commonly used in the regression models and can be stated as follows [16]:-

$$y_i = \beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots\ldots+\beta_p x_{ip}+\varepsilon_i \qquad ,i=1,2,\ldots,n \qquad (1)$$

where

p: is the number of explanatory (independent) variables.
n: is the number of observations.

β : are the regression coefficients.
$\mathcal{E}_i$: is the errors model.
x: are the explanatory (independent) variables.
y_i: are response (dependent) variables.

The problem with multiplicity of linearity is the independent variables are highly related so that the regression results are unclear; where regression coefficient is interpreted as the change in the response (independent) variable as a result of change in the independent variable [6].

Because near multi - collinearity can be thought of as a situation where two or more explanatory variables move together, it is impossible to determine which of explanatory variables is causing observed change in response variables, the unique influence of each variable in the regression equation cannot be estimated [1].

Multicollinearity problem might be extremely difficult to identify. It is not a specification error that the check regression residual can discover. It is, in fact, a modeling mistake. It's an instance of faulty data [6].

The number of independent variables in the model is also critical, as the negative effect of collinearity may be amplified, when the model has more independent variables [2].

And several different methods have been suggested to deal with the problem of multiplicity of linearity.

3. Ordinary Least Squares Estimators (OLS)

The Ordinary Least Squares Estimators method is one of the earliest and most widely used estimation methods for parameter estimation in multiple linear regression models, when the data are independent, identical, and normally distributed; the (OLS) method produces unbiased parameter estimates with minimum variance [7] The (OLS) estimator method minimizes the summation of the squares deviation as it is unbiased of the β as E(b_{OLS})=β, has a minimal variability as V(b_{OLS})= $\sigma^2(X'X)^{-1}$,that meets the L.b. of Rao-Gramere inq., the residuals vector given by ε=y-y all independent variables are orthogonal, so (b_{OLS}) is the best linear unbiased estimators (BLUE) for β as follows:

$$\varepsilon'\varepsilon = (Y - X\beta)'(Y - X\beta)$$
$$= Y'Y - \beta'X'Y - Y'X\beta - Y'X\beta + \beta'X'X\beta$$
$$= Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

The Least Squares estimate of the value (b_{OLS}) which when substituted in equation (1) minimizes $\varepsilon'\varepsilon$ with respect to β and setting the resultant matrix equation equal to zero, at the same time replacing β by (b_{OLS}) the solution of this equations is:-

$$b_{OLS} = (X'X)^{-1}X'Y \qquad (2)$$

And irrespective with the distribution of error terms, the fitted values are obtained from $\hat{Y}=Xb_{OLS}$ [9].

Many scientists wanted to discover different solutions to solve the multicollinearity problem, the proposal of ridge parameters was one these method.

3.1. Ordinary Ridge Regression Estimators

In the classic regression model, regression is assumed to be largely independent. However, in a variety the situations of the real world (e.g., engineering (Hoerl & Kennard (1970)), the regressors are frequently found to be nearly dependent. The matrix X'X is then has ill condition (i.e. det(X'X)≈0). As a result, if X'X is ill condition, (b_{OLS}) is highly susceptible to various errors, making effective statistical inference extremely difficult for practitioners to overcome [2].

Different methods for dealing with multiple linear data have been proposed by modifying the Least Squares method to allow the introduction of bias in the regression coefficient estimators. The ridge regression method is one of the most frequently used techniques. For any K ≥ 0, we define the ridge regression estimator (b_{RR}) as follows:-

$$b_{RR}=(X'X + KI)^{-1}X'Y \qquad (3)$$

The analyzer selects the K value on some acceptable criteria stated according to Hoerl and Kennard [9].

The ordinary ridge estimator produces a series of solutions rather than a single solution to the multicollinearity problem. These solutions are dependent on the value of K (ridge parameter); no explicit optimal value for K has been discovered; several random choices for this parameter have been proposed [16].

4. Least Absolute Deviation (LAD)

Least Absolute Deviation approach is robust alternate to the (OLD) method, mainly when the data followed the non-normal distribution with outliers, because data sets display certain characteristics that may not always meet the (OLS) requirements, there are two requirements related to kurtosis and skewness. Regarding kurtosis, (OLS) assumes that the distribution of the residuals is normal. However, the datasets may be messy; for example, the (OLS) regression could be lesser effective than some robust regression techniques due to heavy tails, i.e., high kurtosis or skewed residual distribution, when the distribution of the residuals has a large kurtosis, The (LAD) regression method might be more effective than the (OLS) method. Regarding skewness, the distribution of residual of real datasets is not always completely symmetrical. If the distribution is skewed (for example by having a few large outliers on one side), outliers affect the (LAD) less than the (OLS), because (LAD) is way less sensitive to outliers than (OLS) [3].

The (LAD) method is classified as a robust regression technique as well may be considered to be a regression techniques family and is not a single particular regression technique, in which the robust regression seeks to also be robust to outlier observations [18].

The (LAD) approach has become one of the most widely used strategies for robust regression analysis. In comparison to (OLS) estimates, (LAD) estimates are less impacted by extreme values. However, the behavior of (LAD) estimates is less well known, especially for small samples, and the inference procedure is more difficult [7]. Inference in the (LAD) estimation is a popular study topic, in employing the (LAD) estimate, Koenker and Bassett [11] suggested the wald, Lagrange multiplier (LM) tests, and likelihood ratio (LR). These methods can be used to test the coefficient significance in a regression model. When data are independent but not necessarily normal, Dielman and Pfaffenberger [4] investigate regression inference using (LAD) estimation. However, while (LAD) estimator has been suggested as a substitute for least squares regression, it is a technique that is not widely used and thus may be regarded as unconventional.

In the (LAD) method, parameter estimates are the parameter values that minimize to:

$$\sum |Yi - (a + bXi)|$$

In this simple regression instance, a and b are the parameters, and this may be generalized to multiple regression. The estimate concept (LAD) is not more complicated than the estimation concept (OLS). In actuality, the absolute value of a residual is a more basic measure of its size than the squared residual, but it is more involved in the actual computation of the estimations (LAD). There are no formulae for estimating LAD. Fortunately, many techniques exist to calculate them [19].

4.1. Least Absolute Deviations Ridge regressions

Despite the fact that the method for calculating the LAD is powerful, However, there is still the possibility of having a highly multi-linear relationship between the explanatory variables in the linear regression analysis, where Least Absolute Deviation Ridge (LAD-ridge) represents the ridge parameter estimation using (LAD) method instead of estimating it using (OLS) method, when the independent variables are extremely collinear and error terms are asymmetric or heavy-tailed, the goal of using LAD ridge is to decrease the (MSE) of the LAD estimators by allowing greater space for bias.

Multicollinearity problem exists naturally in most real-world data sets. The question is: to what extent must the problem of medium and strong multicollinearity be somehow addressed. One of the easiest ways ,if one needs to keep all the independent variables1in the model, is to use ridge estimates method, that was first used by Hoerl and Kennard (1970) [9] in regression analysis.

An extensive study has been conducted on the performance of (LAD) method and ridge regression individually; both are robust, but each is suited to a different kind of problem. In an early effort to combine the two approaches into a single estimation procedure, Pfaffenberger and Dielman (1989) [17] attempted to use (LAD) estimation to estimation the ridge biasing parameter (K), for both parameters and the variance of the error terms. In other words, they applied the same biasing parameter as Hoerl et al.(1975) utilized (LAD) estimates rather than (OLS) estimates for the parameters and variance of error terms in the calculation they used (LAD) estimator.

Generally, the OLS ridge estimators are less robust than the LAD ridge estimators at the presence of asymmetry of error terms. For every increase in the asymmetry of the error terms, the relative efficiency of LAD ridge estimators to OLS ridge estimators grows.

The linear regression model is:-

$$y_i = x_i'\beta + \varepsilon_i \qquad , i=1,2,\ldots,n \qquad (4)$$

where ε_i are iid random errors, x_i is a known (p× 1) × 1 vector of predictors and β is a (p× 1) × 1 unknown parameter vector. Then, let us take a full rank augmented linear model like:-

$$\begin{bmatrix} Y \\ 0 \end{bmatrix} = \begin{bmatrix} X \\ \sqrt{K} I_p \end{bmatrix} \beta + \begin{bmatrix} \varepsilon \\ \upsilon \end{bmatrix} \qquad (5)$$

E(ε)=0, E($\varepsilon\varepsilon'$)=Σ, E(υ)=0, E(υυ′)= Ω, E(ευ′)=0, where υ~ N(o,σ^2), and K>0. The LAD estimation of β=(β_1, β_2,…, β_p)' of (5) is

$\hat{\beta}= argmin_{\beta=(\beta_1,\ldots,\beta_p)'}(\sum_{i=1}^{n}|y_i - x_i'\beta| + \sqrt{K}\ \sum_{j=1}^{p}|\beta_j|)$

Some of the formulas for finding (K) depend on the estimated parameter of the standardized variables in the regression models (4)

$$Y=X\beta+\mathcal{E} = S'\Lambda^{\frac{1}{2}}\alpha+\mathcal{E}$$

The singular value decomposition of the standardized design matrix is X=S' $\Lambda^{\frac{1}{2}}$ U', with S a p× n orthogonal matrix, Λ a p×p diagonal matrix of the eigenvalue of X'X and U a p×p orthogonal matrix of the eigenvectors of X'X, and $\alpha = U'\beta$. The (LAD) estimation of α, α^*=($\hat{\alpha}_0$, $\hat{\alpha}'$)', is the (LAD) parameter estimation of the model [20],

$$\begin{bmatrix} Y \\ 0 \end{bmatrix} = \begin{bmatrix} 1_n & S'\Lambda^{\frac{1}{2}} \\ 1_p & \sqrt{K}\ I_p \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha \end{bmatrix} + \begin{bmatrix} \varepsilon \\ \upsilon \end{bmatrix} \qquad (6)$$

There are several solutions to solve the multicollinearity problem; one of these solutions is ridge regression, which is regarded as one of the methods of the bias estimates, where the ridge parameters help give bigger space for bias to minimize MSE. Below are several previously suggested ridge estimates for the (LAD) estimator.

4.2. Ridge Parameters

The Least Absolute Deviation ridge estimator does not give a single solution to the linearity problem but instead a set of options dependent on (K) (ridge parameter). Although it cannot figured, an exact ideal value for (K), numerous random values for this ridge parameter have been proposed. Several of these alternatives can be summed up as follows:-

- Hoerl and Kennard (1970) [9] that (K_i) minimizes MSE($\hat{\alpha}(K)$) which is defined as

$$\text{MSE}\ (\hat{\alpha}(\hat{K})) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+\hat{K}_i)^2} + \sum_{i=1}^{p} \frac{\hat{K}_i^2\ \hat{\alpha}_i^2}{(\lambda_i+\hat{K}_i)^2}$$

Where the λ_i are eigenvalue of the matrix X'X, $\hat{\alpha}_i$ is the i^{th} element of $\hat{\alpha}$ and estimated by (LAD) where α=U'β and U is the orthogonal matrix, $\hat{K}_i$ is represent the ridge parameters which is estimation by unbiased estimation (LAD) where

$$\hat{K}_i=\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \qquad (7)$$

Where $\hat{\sigma}^2=\frac{\hat{\mathcal{E}}'\hat{\mathcal{E}}}{(n-p-1)}$ is the residual mean square estimate, which is unbiased (LAD) estimator of σ^2.

- Hoerl and Kennard (1970) [9] suggested (K) to be (denoted by $\hat{K}_{HK}$)

$$\hat{K}_{HK}=\frac{\hat{\sigma}^2}{\hat{\alpha}_{max}^2} \qquad (8)$$

Where $\hat{\alpha}_{max}$ is represented the maximum element of $\hat{\alpha}$. Hoerl and Kennard claimed that (8) given smaller (MSE) than (LAD) method.

- Hoerl et al. (1975) [10] proposed (K) to be (denoted by $\hat{K}_{HKB}$)

$$\hat{K}_{HKB}=\frac{p\ \hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}} \qquad (9)$$

- Lawless and Wang (1976) [13] proposed (K) to be (denoted by $\hat{K}_{LW}$)

$$\hat{K}_{LW}=\frac{p\,\hat{\sigma}^2}{\hat{\alpha}'\,X'X\,\hat{\alpha}} \qquad (10)$$

- Kibria (2003) [12] suggested (K) based on arithmetic mean is denoted by ($\hat{K}_{AM}$), and geometric is denoted by ($\hat{K}_{GM}$) of ($\hat{\sigma}^2/\hat{\alpha}_i^2$) defined as the follows:-

$$\hat{K}_{AM}=\frac{1}{p}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \qquad (11)$$

$$\hat{K}_{GM}=\frac{\hat{\sigma}^2}{(\prod_{i=1}^{p}\hat{\alpha}_i^2)^{\frac{1}{2}}} \qquad (12)$$

- Muniz et al. (2012) [15] proposed estimators of (K):-

$$\hat{K}_{KM12}=\text{Median}\,(\frac{1}{q_i}) \qquad (13)$$

Where $q_i=\frac{\lambda_{max}\,\hat{\sigma}^2}{(n-p)\hat{\sigma}^2+\lambda_{max}\,\hat{\alpha}_i^2}$, λ_{max} represented maximum eigenvalue of the matrix X'X.

5. Proposed Ridge Estimator

Our contribution in this research is the use of the Variance Inflation Factor (VIF), which is an indicator of the existence of a multiple linear relationship between the independent variables, where the (VIF) is calculated for each independent variable in the model. If (VIF) is large, this is an indication that there is a strong multiple linear relationships between the independent variables in the model.

Dorugade and Kashid (2010) [5], based on (9), suggested (K) to be denoted by ($\hat{K}_D$)

$$\hat{K}_D=\max\left(0\,,\hat{K}_{HKB}\;-\;\frac{1}{n(VIF_j)_{max}}\right) \qquad (14)$$

Where ($\hat{K}_D$) is represent modification ($\hat{K}_{HKB}$) in eq. (9) by subtracted $\frac{1}{n(VIF_j)_{max}}$ from ($\hat{K}_{HKB}$). This amount, however varies with the size of the sample (n) used and strength of the multicollinearity in the model. Dorugade and Kashid (2010) [5], worked on estimation ($\hat{K}_{HKB}$) by (OLS) method, and we will estimation ($\hat{K}_{HKB}$) by (LAD) method.

Where $\text{VIF}_j=\frac{1}{1-R_j^2}$, j= 1, 2,…,p is Variance Inflation Factor of j^{th} regressor and R_j^2 is the coefficient of determination of X_i on other covariates, $X_1, X_2, X_3, \ldots, X_i, X_{i+1}, \ldots, X_p$.

The coefficient of determination R_j^2 can be defined as the ratio as the proportion of total variability in the response (dependent) variable y, which can be calculated by the set of predictors $X_1, X_2, \ldots, X_p$. It is clear that when the value of R^2 is close to one, it means that the model fits the data well, in such a cases the observed and predicted values became close to each other, in this case the error will be small [16].

6. Simulation study

This chapter, describes the simulation research to make a comparison between the performance of the estimated ridge parameters to determine which ridge parameter is the most efficient and has the least number of parameters (MSE). In this paper, comparison between the proposed parameter ($\hat{K}_D$) and the ridge parameter $\widehat{(K_{HKB})}$, where the (MSE) is used as a criterion for quality throughout study to get the best estimator.

Six independent variables were used p=6 in the linear regression model, which represents the following model in (1):-

McDonald and Galarneau (1975) [14] and Gibbons (1981) [8] the explanatory variables were generated by using the following:-

$$x_{ij} = (1-\rho^2)^{\frac{1}{2}}\, \varepsilon_{ij} + \rho\varepsilon_{ij} \qquad , i=1,2,\ldots,n \ , j=1,2,\ldots,p \qquad (15)$$

where $\varepsilon_{ij} \sim N(0,\sigma^2)$ is independent and ρ represents the correlation between two independent variables which is inactivated by ρ^2, in this paper it is assumed that it is $\beta_0=\beta_1=\beta_2=\beta_3=\beta_4=\beta_5=\beta_6=1$, and correlation values (0.2 , 0.7 , 0.9) with different sample sizes (n=25, 60, 100) and (σ^2 = 0.8, 2.5, 5). The experiment was repeated 500 times. Using MATLAB (R2019b) program was utilized to make a comparison between two ridge parameters to determine which one is the best to solve the multicollinearity problems.

The MSE values for two ridge estimators are shown in the tables below.

Table 1. The values of MSE at **ρ=0.2**.

n	**Estimator**	σ^2			**Best**
		0.8	**2.5**	**٥**	
25	$\hat{K}_{HKB}$	٠.٧٤٢٧	٠.٧٩٥١	٠.٩٦٢٢	$\hat{K}_D$
	$\hat{K}_D$	٠.٥٦٨٠	٠.٥٣٤٨	٠.٥١٩٥	
٦٠	$\hat{K}_{HKB}$	٠.٦٨٣٦	٠.٧٠٨٧	٠.٧٩٧٦	$\hat{K}_D$
	$\hat{K}_D$	٠.٥٥٠٣	٠.٥٢٨٩	٠.٤٩٥٦	
١٠٠	$\hat{K}_{HKB}$	٠.٦٧٢٢	٠.٦٨٩٠	٠.٧٣٨٧	$\hat{K}_D$
	$\hat{K}_D$	٠.٥٣٨٩	٠.٥٢٢٨	٠.٤٩٠٩	

Table 2. The values of MSE at **ρ=0.7**.

n	**Estimator**	σ^2			**Best**
		0.8	**2.5**	**٥**	
25	$\hat{K}_{HKB}$	٠.٧٣٥٥	٠.٨١٦٦	٠.٩٧٠٤	$\hat{K}_{HKB}$
	$\hat{K}_D$	0.7359	0.8170	٠.٥٠٠٥	
٦٠	$\hat{K}_{HKB}$	٠.٦٨٣٥	٠.٧١٩٩	٠.٨١١٣	$\hat{K}_D$
	$\hat{K}_D$	٠.٥٤١٠	٠.٥١٤٢	٠.٤٩٠٩	
١٠٠	$\hat{K}_{HKB}$	٠.٦٧١٩	٠.٦٩٠٨	٠.٧٧١٦	$\hat{K}_D$
	$\hat{K}_D$	٠.٥٤٠٢	٠.٥١٤٠	٠.٤٨٧٧	

Table 3. The values of MSE at ρ=**0.9**.

n	Estimator	σ^2			Best
		0.8	**2.5**	**٥**	
25	$\widehat{K}_{HKB}$	٠.٧٣٤١	٠.٧٩٠٦	٠.٩٩٢٥	$\widehat{\boldsymbol{K}}_{\boldsymbol{HKB}}$
	$\widehat{K}_D$	0.7346	0.7911	٠.٥٠٢٢	
٦٠	$\widehat{K}_{HKB}$	٠.٦٨٧١	٠.٧١٤٧	٠.٨١٦٤	$\widehat{\boldsymbol{K}}_{\boldsymbol{D}}$
	$\widehat{K}_D$	٠.٥٤٠٦	٠.٥٢٤٠	٠.٤٨٧٥	
١٠٠	$\widehat{K}_{HKB}$	٠.٦٧١٧	٠.٦٨٥٣	٠.٧٧٤٢	$\widehat{\boldsymbol{K}}_{\boldsymbol{D}}$
	$\widehat{K}_D$	٠.٥٣٩٠	٠.٥٢١١	٠.٤٨٥٢	

6.1. Analysis of simulation results

Tables 1, 2, and 3 give the simulation results of MSE of ridge parameter (K), and we can see that:

In all experiments and all sample sizes $\widehat{K}_D$ is the best, accept in experiment 2, 3 and in (n=25) with strong correlation and variance the best $\widehat{K}_{HKB}$, these results confirm the effectiveness of the suggested estimator $\widehat{K}_D$.

References

١- محبوب، عادل عبد الغني (١٩٩٨)، " أصول الاقتصاد القياسي- النظرية و التطبيق "، شركة الاعتدال للطباعه بغداد

[2] B. M .G. Kibria and S. Banik, (2016) "Some Ridge Regression Estimators and Their Performances" Journal of Modern Applied Statistical Methods, Vol.15, No.1, PP.206-238.

[3] Briand, L. C. and Wieczorek, I., (2000) "Resource Estimation in Software Engineering", In: Encyclopedia of Software Engineering, Wiley.

[4] Dielman, T. E. and Pfaffenberger, R. (2009) "Tests of Linear Hypothesis in LAV Regression". Communications in Statistics-Simulation and Computation, 19, PP. 1179-1199.

[5] Dorugade, A. V. and Kashid, D. N. (2010). "Alternative Method for Choosing Ridge Parameter for Regression", Applied Mathematical Sciences, 4(9), 447-456.

[6] Dorugade, A. V. (2010); "Alternative Method for Choosing Ridge Parameter for Regression", Applied Mathematical sciences, Vol.4, No.9, PP 447-456.

[7] Dadamos, T. R. L, Damaceno, A. J, Fertonani, F. L. and Bettencourt Silva, R.J.N. (2018) "Use of Monte Carlo Simulation for The Calculation of The Uncertainty Estimate in The Electrochemical Detection of Uric Acid in Plasma Samples". Periódico Tchê Química, 15(29), PP.196-203.

[8] Gibbons, D. G. (1981). "A Simulation Study of Some Ridge estimators". Journal of the American Statistical Association 76:131-139.

[9] Hoerl, A. E. and Kennard, R. W. (1970); "Ridge Regression Biased Estimation for Nonorthogonal Problems", Technometrics, Vol.12, PP55-67.

[10] Hoerl, A. W., and Baldwin, K. F. (1975). " Ridge Regression: Some Simulations ". Communications in Statistics, 4(2), 105-123.

[11] Koenker, R. and Bassett, G. (2002) "Tests of Linear Hypothesis and L1 Estimation". Econometrica, 50, PP. 1577-1583.

[12] Kibria, B. M. G. (2003). " Performance of Some New Ridge Regression Estimators". Communications in Statistics-Simulation and Computation, 32(2), 419-435.

[13] Lawless, J. F. and Wang, P. (1976). " A Simulation Study of Ridge and Other Regression Estimators". Communications in Statistics-Theory and Methods, 5(4), 307-323.

[14] McDonald, G. C., Galarneau, D. I. (1975). "A Monte Carlo Evaluation of Some Ridge-Type Estimators". Journal of the American Statistical Association 70:407-416.
[15] Muniz, G., Kibria, B. M. G., Mansson, K., and Shukur, G. (2012). " on Developing Ridge Regression Parameters: A Graphical Investigation". Statistics and Operations Research Transactions, 36(2), 115-138.
[16] Mahdi, Fatima Assim (2017); "A Comparison Different Ridge Regression Method with Application", University of Baghdad-College of Education for Pure Science, Ibn Al- Haitham, Baghdad, Iraq.
[17] Pfaffenberger, R.C., and Dielman, T. E., (1989),"A Comparison of Regression Estimators When Both Multicollinearity and Outliers are Present". In: Lawrence, K. D. and Arthur, J. L., editors, Robust Regression: Analysis and Applications, New York: Marcel Dekker, Inc. PP. 243-270.
[18] Pickard, L., Kitchenham, B. and Linkman, S., (1999) "An Investigation of Analysis Techniques for Software Datasets", Proc. Metrics 99 Symposium, IEE Computer Society, PP.130-142.
[19] Tronfass, Ingunn Myrtveit, Erikstensrud, (2001) "A Comparison of LAD and OLS Regression for Effort Prediction of Software Projects".
[20] Zangin Zeebari (2012). "A Simulation Study on the Least Absolute Deviations Method for Ridge Regression". Department of Economics and Statistics, *JÖnkÖping* International Business School , Sweden, Article in Communication in statistics-theory and method.