GLOBAL RESEARCH NETWORK
ONLINE RESEARCH HUB

ISSN: 2576-5973
Vol. 5, No. 3, 2022

Estimations of Kibria-Lukman and Lui to identify the most important variables influencing COVID-19

Namariq Qassim Hussain [1] ***Zainab Kadhum Mezher*** [2] ***Mahdi Wahab Neama*** [3]
University of Karbala - College of Administration and Economics - Department of Statistics [1]
University of Karbala - College of Administration and Economics - Department of Statistics [2]
University of Karbala - College of Administration and Economics - Department of Statistics [3]
namariq.q@uokerbala.edu.iq [1] zainabalqurishi46@gmail.com [2] mehdi.wahab@uokerbala.edu.iq [3]

Abstract: The problem of Multicollinearity between explanatory variables is one of the problems that have attracted the attention of many researchers and that the ordinary least squares method is unable to solve it , so the researchers found solutions to this problem by using the principal components method, the ridge regression method, or the lasso regression method, but these methods are not accurate enough to obtain on efficient estimates of the multiple regression model, especially in the presence of the problem of Multicollinearity. In this research, the capabilities of Kibria-Lukman and Lui will be used to determine the most important factors affecting infection with the Covid-19 virus among the variables that suffer from the problem of Multicollinearity to solve this problem and then obtain the best variables that represent the most important factors affecting infection with the virus. It was concluded that the Kibria-Lukman estimator is better than the rest of the estimators because it achieved the least mean squares error (MSE), and was able to solve the problem of Multicollinearity among the explanatory variables, followed by the Liu estimator and then the Ridge regression estimator, and finally the ordinary least squares estimator that was unable to solve the problem of linear multiplicity.

Keywords: Multicolinearity, OLS, Kibria-Lukman estimator, Lui estimator, multiple regression models, COVID-19.

1. Introduction

Multicollinearity according to the concept (Ranger Frisch, 1934) means a linear or almost perfect relationship, resulting in no sign of linearity. Show the values of observations of the explanatory variables in the regression model to be estimated. Table of interlocking variables leading to an inaccurate evaluation of the regression model. The multiple linear rate ranges from multiple, multiple, and collinearity in a linear sense. (Gujarati, 436). Polylinearity can be defined by the concept of orthogonal it That is, when there is a perfect linear relationship between two or more explanatory variables, that is, the information matrix (S=X'X) is perfectly ordered and all the distinct values (Eigen values) have equal one, and this indicates that orthogonality exists and with which it is impossible to find The inverse of the information matrix, since the determinant of this matrix will be zero. (Amore and Bassem, 2002, 182) It faces the problem of linear multiplicity when some or all of the independent variables are linked in a linear relationship, it becomes difficult to separate the effect of each variable on the dependent variable, or when the value of one of the independent variables is equal for all observations, or when the value of one of the independent variables depends on The value of one or more of the other independent variables in the model.

Or when using time-return variables (time-deceleration variables) as explanatory variables in the model. The multilinearity problem can lead to biased and inaccurate estimates of the regression coefficients, amplify the standard errors of the regression coefficients and underestimate partial t-tests of the regression coefficients, as well as reduce the predictability of the model. (Al-Quraishi, 2018, 7) Linear relationships vary according to the nature of the relationship between the explanatory variables, so that they are divided into the complete linear relationship.

Which is achieved when there is a linear relationship between the values of two or more explanatory variables, so the result of the determinant of the matrix is zero |X´X|=0 and this leads to a violation of the Rank Condition, as the matrix is incomplete, meaning that Rank(X <P) The rank is less than the number of variables, so it is not possible to find the inverse of the information matrix, and therefore it is not possible to estimate the parameters of the model, and here the complete polylinearity appears.

So if we have the explanatory variables $x_1, x_2, \dots \dots, x_p$and it is found that p of the constants $C_1, C_2, \dots \dots, C_p$are not all zeros and the following equation is achieved

$$\sum_{j=1}^{p} C_j X_j \qquad (2)($$

And the relationship (2) shows that the vectors are not linearly independent, and if $C_1 = C_1 = \cdots = C_p$also fulfills the relationship (2), while if $C_j \neq 0$then it can be solved with respect to

$$X_j = \frac{1}{C_1}\left[C_1X_1, C_2X_2, \dots \dots, C_pX_p\right]$$

From equation (3) it becomes clear that if p of the vectors is linearly related, this indicates that one of them is a linear combination of the other vectors, and any vector in terms of the other vectors can be obtained. (Gabriel, 2014, 62). The other type of multicollinearity is the semi multicollinearity, which appears when the determinant of the information matrix is not equal to zero and is very small |X´X|≈0 then the estimated parameters are highly stigmatized which occurs when some variables are Explanatory function in the combination of other variables with random values if the following relationship is true: (Amore and Bassem, 2002, 190(

$$\sum_{j=1}^{p} \alpha_j X_j + \varepsilon_i = 0 \quad ; \text{ i=1,2,…,p} \qquad (4)$$

And relationship (4) shows that the vectors are not linearly independent, and if at least $\alpha_1, \alpha_1, \ldots, \alpha_p$ is not equal to zero, as well as $\varepsilon_i \neq 0$, in this case the determinant of the information matrix is not equal to zero, but it is close to zero, and thus the variance of the estimated parameters increases and therefore the lack of The accuracy of the ordinary least squares method in estimation. (Gabriel, 2014, 64)

Multicollinearity Diagnosis

There are several proposed techniques to detect the problem of multilinearity as follows:

2.1.1Variance Inflation Factor (VIF): This test was presented by (Ferrar & Glauber m 1967) as the VIF. VIF is one of the basic and broad methods for detecting the existence of a polylinear problem, and it measures the extent of the inflation of the variances of the estimated regression coefficients when there is a linear relationship between the explanatory variables, and the diagonal elements of the inverse of the information matrix are useful in revealing the problem of multilinearity, and the diagonal elements J^th can be written for the C matrix

$$C_{jj} = \left(1 - \mathrm{R}^2{}_{\mathrm{j}}\right)^{-1}$$

It is noted that if X_j is close to orthogonality with other explanatory variables (p-1), then the coefficient of determination $(\mathrm{R}^2{}_{\mathrm{j}})$ is small and the matrix C_{ij} is close to one, while if X_j is dependent on some explanatory variables, then $(\mathrm{R}^2{}_{\mathrm{j}})$ is close to one and the matrix C_{jj} is large, and since the variance of the parameter j is σ^2can be seen as a factor that increases the C_{jj} variance of $\hat{\beta}_j$ as a result of its close linear dependence on some explanatory variables included in the model And the equation (5) measures the so-called variance inflation factor: (Net, 2015, 58-57)

$$VIF = C_{jj} = \left(1 - \mathrm{R}^2{}_{\mathrm{j}}\right)^{-1} \qquad (5)$$

Where: VIF_j is the variance inflation factor for the explanatory variable, $\mathrm{R}^2{}_{\mathrm{j}}$ is the coefficient for determining the regression model of the explanatory variable J on the remaining explanatory variables p-1.

The variance amplification factor VIF measures the effect of multilinearity among the explanatory variables in the regression model, which is always greater than or equal to one, but there is no limit value for VIF to determine the severity of the multilinearity effect on the model. The largest value of VIF is often used as an indicator of non-linearity. Desirable and often if it exceeds (10) it is considered an indication of the possibility of an unacceptable effect of high polylinearity on the estimations of the ordinary least squares.

Published by "Global Research Network LLC"
https://www.globalresearchnetwork.us

If there is a perfect correlation between the explanatory variable $X_j$and the other explanatory variables so that =1, then the variance inflation factor goes to infinity $R^2{}_j$ =0, $R^2{}_j$is orthogonal to the other variables so that X_j, If the variable $VIF = C_{jj} = (1 - R^2{}_j)^{-1}$ then the value of the variance inflation factor is equal to the correct one.

The average VIF values are polylinear risk information based on the spacing between the estimated standard regression parameters $\hat{b}_j$ and their true values $\hat{\beta}_j$, which is calculated according to the following formula: : (Roman et al , 2020, 4(

$$\overline{VIF} = \frac{\sum_{j=1}^{p-1}(VIF)_K}{p-1} \qquad (6)$$

The large values of VIF produce, on average, large differences from the estimated standard regression parameters and the true standard regression parameters, and if the average values of the VIF are much greater than one, it is an indication of a significant impact of multilinearity on the model. (M. Marcoulides&

. (The inverse of the variance inflation factor, which is known as the Tolerance Measure, is also used to detect the presence of linear multiplicity between the explanatory variable X_j and the other explanatory variables, according to the following formula: (Gabriel,69 ,2014)

$$Tolerance = \frac{1}{VIF_j} = 1 - R^2{}_j \qquad (7)$$

It is noted that the variance of the least squares estimator for the parameter b_j in terms of the variance inflation factor, which is calculated according to the following formula:

) Roman et al , 2020, 5) (El-Dereny & N. I. Rashwan, 12(

$$var(b_j) = \frac{S^2 VIF_j}{(n-1)S_j^2} = \frac{1}{1-R^2{}_j} \cdot \frac{S^2}{(n-1)S_j^2} \qquad (8)$$

Where :

$S2$ is estimator of the variance of the regression model of the dependent variable Y on the rest of the explanatory variables

S_j^2 Covariance of the explanatory variable X , which will equal $S_j^2 = \frac{\sum_{j=1}^{n}(X_j - \overline{X}_j)^2}{n-1}$

$VIFj$ Variance inflation factor for the independent variable Xj

Also, the variance inflation factor for the variable j is the diagonal element number j for the inverse of the matrix of simple correlation coefficients between the explanatory variables as: : (Yong-wei, 2008, 44)

$$VIF_j = diad(R_{XX}^{-1})jj \qquad (9)$$

2Eigen Values: The characteristic values of the correlation matrix are the parameters of the regression model, good indicators of polylinearity, which are calculated by the following steps:

Find the matrix Z so that:

Z=(X´X) (10 (

Find the diagonal matrix S such that:2

$S = dig(X'X)^{-\frac{1}{2}}$ (11)

When one of the eigenvalues equals λ1, λ2, ..., λp of the matrix (X´X) to zero, it means that there is complete polylinearity, and if the value of one of the characteristic roots is close to zero, it indicates the presence of high polylinearity

.

2-6-3Conditional Criterion (CI):

This indicator is used to detect the presence of polylinearity. Through the characteristic values, this indicator is calculated according to the following formula: . . (Shrestha, 2020, 40(

$$CI_j = \sqrt{\frac{\lambda_{max}}{\lambda_j}} \qquad j = 1,2, \ldots . . p \qquad (12)$$

Where: λ_max is the largest characteristic value, λ_j is the characteristic value of the number j, so if the value of C.I. Greater than (30), this is evidence of a polylinear problem.

2.1.3Conditional Number: Condition Number (CN) The conditional number is used to measure the degree of polylinearity according to the following formula:

$$\text{C. N} = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \qquad (13)$$

Johnston suggested that if the value of C.N ranges between 20 to 30, it is an indicator of the presence of high polylinearity, while (Belsley) suggested that if the value of C.N ranges between 30 to 100, it is an indicator of the presence of very high polylinearity. (Shrestha, 2020, 41)

2.1.4Correlation Matrix: Testing the matrix of explanatory variables is one of the simplest ways to detect multilinearity. If the correlation between explanatory variables X_j , X_i is high, then the absolute value of $|r_{ij}|$ It is close to one, which exposes the model to multi-linearity, but the weak correlation between the variables does not mean that there is no linear multiplicity. (Shrestha, 2020, 41(

5Farrar & Glauber test: This test was developed in 1967 and is based mainly on the Chi square test to test the following hypothesis: (Josef & Sanni, 2017, 7-8) based on the determinant of the correlation coefficient matrix |R| The number of independent variables (p) and the number of experimental units (n) and degrees of freedom (n(n-1))/2 and its mathematical formula is as follows:

$$\chi_0^2 = -\left[\left[n - 1 - \frac{1}{6}(2p+5)\right] ln \mid R \mid\right] \qquad (14)$$

R: matrix of correlation coefficients, p: number of variables, n: sample size

Equation (14) is compared with the tabular value of ²χ at a freedom jar (n(n-1))/2, and the alternative hypothesis is rejected if the calculated value is greater. (Dabdoub, 2006, 89(

.3Ridge regression

Ridge regression is a technique for analyzing multicollinearity data. This method has shown effectiveness in getting rid of the problem of multiple linear relationship, as the overlap between the explanatory variables leads to the large size of the variance of estimators and confusion in the difference in the relative relationships between the predictive variables and the response variable when using the least squares method. (Al-Hassan, Yazid M., 2010). The first to point out the danger of multiple linear relationships and their impact on the results of regression analysis was the scientist Fisher, and that was in 1934. It was followed by many researchers who elaborated the different aspects of the problem and methods for solving it, until Hoerl & Kennard (1970) added To solve this problem, they called this constant amount the Biasing parameter, and the method of ridge regression. (Hoerl & Kennard, 1970(

The character estimator can be obtained by minimizing the following objective function:

$$(y - X\beta)'(y - X\beta) + K(\beta'\beta - C) \qquad (15)$$

$$(X'X + KI_P)\beta = X'y \qquad (16)$$

Where K is a positive constant,

and by solving equation No. (16), we get the estimator of the letter as follows:

$$\hat{\beta}(k) = (S + KI_P)^{-1}X'y = w(k)\hat{\beta} \qquad (17)$$

Where is bias parameter$W(K) == [I_P + KS^{-1}]^{-1}, k$, $S = X'X$,

(Hoerl et al) defined a harmonic mean of bias parameter k as follows:

$$\hat{k} = \frac{P\hat{\sigma}^2}{\sum_{i=1}^{p}\alpha_i^2} \qquad (18)$$

Where :

$\hat{\sigma}^2 = \frac{(Y'Y - \beta\prime X\prime Y)}{n-p}$ The mean squared error estimated from the OLS method

α_i^2 is the coefficient of α=Q'β and is defined in the equation y=Zα+ε

4. liu estimator

The Liu estimator for the parameter β can be obtained by integrating dβ̂=β+ε' with equation (1) and with this, applying the OLS method estimator to estimate the parameter, the Liu estimator can be obtained as follows:

$d\hat{\beta} = \left(S + I_P\right)^{-1}\left(X'y + d\hat{\beta}\right) = F(d)\hat{\beta}$ (23)

where d of the liu estimator is defined as:

$\hat{d} = 1 - \hat{\sigma}^2 \left[\frac{\sum_{i=1}^{p}(1/\lambda_i(\lambda_i+1)))}{\sum_{i=1}^{p}\alpha_i^2/(\lambda_i+1)^2}\right]$ (24)

4. Kibria-Lukman estimator (Kibria & Lukman, 2020)

It is estimated with a single parameter that can be obtained from minimizing the following objective function:

$(y - X\beta)'(y - X\beta) + K\left[\left(\beta - \hat{\beta}\right)'\left(\beta - \hat{\beta}\right) - c\right]$ (26)

With respect to β which leads to the normal equation

$(X'X + KI_P)\beta = X'y - k\hat{\beta}$ (27)

Where K is a positive constant, and by solving equation No. (27), we get the estimator of the letter as follows: (B. M. Golam Kibria & F. Lukman, 2020)

$\hat{\beta}_{KL} = (S + KI_P)^{-1}(S - KI_P)\beta = W(K)M(K)\hat{\beta}$ (28)

Where $,S = X'X$ $W(K) == [I_P + KS^{-1}]^{-1}$ and K is bias parameter, and $M(K) = [I_P - KS^{-1}]^{-1}$

- ❖ 4.1 Characteristics of the Kibria–Lukman Estimator
- ❖ Expected of Kibria–Lukman

$E\left(\hat{\beta}_{KL}\right) = W(K)M(K)E\left(\hat{\beta}\right) = W(K)M(K)\beta$ (29)

❖ The estimator of Kibria–Lukman is unbiases if K=0

$\beta\left(\hat{\beta}_{KL}\right) = [W(K)M(K) - I_P]\,\beta$ (30)

$D\left(\hat{\beta}_{KL}\right) = \sigma^2 W(K)M(K)S^{-1}M'(K)W'(K)$ (31)

❖ MSEM defined as following:

$MSEM\left(\hat{\beta}_{KL}\right) = \sigma^2 W(K)M(K)S^{-1}M'^{(K)}W'^{(K)} + [W(K)M(K) - I_P]\beta\beta'[W(K)M(K) - I_P]'$
(32)

5. Comparing the OLS, RR ,Liu and LK

We will use the following approach to comparing of the performance OLS, RR, Liu and KL ,

$y = Z\alpha + \varepsilon$ (33)

Where $\alpha = Q'\beta$ و $Z = XQ$ and Q: is orthogonal matrix

$Z'Z = QX'XQ = \Lambda = diag(\lambda_1, \lambda_2 \ldots, \lambda_p)$

The OLS estimator for α is :

$\hat{\alpha} = \Lambda^{-1}Z'y$ (34)

$MSE(\hat{\alpha}) = \sigma^2\Lambda^{-1}$ (35)

The ridge regression of α is

$\hat{\alpha}(K) = W(K)\,\hat{\alpha}$ (36)

Where $W(K) = \left[I_P - K\Lambda^{-1}\right]^{-1}$ and K bias parameter

$$MSE(\hat{\alpha}(K)) = \sigma^2 W(K)\Lambda^{-1}W(K) + (W(K) - I_P)\alpha\alpha'(W(K) - I_P)' \qquad (37)$$

Where $(W(K) - I_P) = -K(\Lambda + KI_P)^{-1}$

The Liu estimator for α

$$\hat{\alpha}(d) = (\Lambda + I_P)^{-1}(Z'Y - d\hat{\alpha}) = F(d)\hat{\alpha} \qquad (38)$$

Where $F(d) = [\Lambda + I_P]^{-1}[\Lambda + dI_P]$

$$MSE\left(\hat{\alpha}(d)\right) = \sigma^2 F_d\, \Lambda^{-1}\, F_d + (1-d)^2(\Lambda + I)^{-1}\alpha\alpha'(\Lambda + I)^{-1} \qquad (39)$$

Where $F_d = (\Lambda + I)^{-1}(\Lambda + dI)$

The Kibria–Lukman estimator for α

$$\hat{\alpha}(KL) = (\Lambda + KI_P)^{-1}(\Lambda + KI_P)^{-1} = W(K)M(K)\hat{\alpha} \qquad (40)$$

Where $W(K) = \left(I_P + K\Lambda^{-1}\right)^{-1}$, $(K) = \left[I_P + K\Lambda^{-1}\right]$ و

There are several characteristics of an estimateis: $\hat{\alpha}(KL)$

- If we have matrices nxn>0 , M>0 و N فان ,N>M if and only if $\lambda_1(NM^{-1}) < 1$ $\lambda_1(NM^{-1})$is largest Eigen value NM^{-1}
- اذا وفقط اذا كان M-$\alpha\alpha'$≥0 متجه , فان α و M وهي محددة موجبة وكان nxn>0 مصفوفة Mليتكن $\alpha' M^{-1}\alpha$≤1
- ليكن $\hat{\alpha} = A_i y; i = 1,2$ مقدرين خطيين للـ α $D = Cov(\hat{\alpha}_1) - Cov(\hat{\alpha}_2) > 0$ حيث $Cov(\hat{\alpha}_i)\ i = 1,2$

The covariance and covariance matrix of α ̂_i is denoted by:

$$b_i = Bias(\hat{\alpha}_i) = (A_i X - I)\alpha \ ; i = 1,2$$

$$\Delta(\hat{\alpha}_1 - \hat{\alpha}_2) = MSEM(\hat{\alpha}_1) - MSEM(\hat{\alpha}_2) = \sigma^2 D + b_1 b'_1 - b_2 b'_2 > 0 \qquad (41)$$

If and onle if $MSEM(\hat{\alpha}_i) = Cov(\hat{\alpha}_i) + b_i b'_i$ حيث أن $b'_2[\sigma^2 D + b_1 b'_1]^{-1} b'_2 < 1$

6. Applied side

It was relied on a data set representing the factors affecting infection with the Covid-19 virus for period of (100) days, which was obtained from the Wasit Health Department / Al-Zahra Teaching Hospital. The dependent variable Y represented the number of infections with the Covid-19 virus, Explanatory variables X1 Average age X2 Blood sugar level X3 Smoking (1 smoker 0 non-smoker) X4 Washing hands with soap and water (1 running 0 not washing) X5 Using sanitizers (1 use 0 to use) X6 Staying at home (1 stay 0 no staying) X7 Incidence of bronchial asthma disease (1 infected 0 unaffected) X8 Gender of the patient (1 male 2 female) X9 Blood type X10 Weight X11 Percentage of white blood cells in the blood, For the estimating the following regression model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{11} x_{i11} + u_i \qquad (42)$$

Using the Ordinary Least Squares (OLS), the letter regression method (Ride), the Liu estimator, the KL estimator, and the comparison between the estimators based on the Mean Squared Error Standard (MSE.(

Table (1) as follows:
Table (1), the matrix of correlations between explanatory variables

X_1	1.00										
X_2	0.87	1.00									
X_{13}	0.59	0.33	1.00								
X_4	0.66	0.25	0.71	1.00							
X_5	0.47	0.02	-0.11	0.00	1.00						
X_6	0.78	0.16	0.03	0.19	-0.21	1.00					
X_7	0.66	-0.21	-0.35	-0.67	-0.19	0.34	1.00				
X_8	0.78	0.06	0.23	-0.18	0.67	0.56	0.33	1.00			
X_9	0.35	0.41	0.37	0.54	0.15	0.03	-0.10	0.16	1.00		
X_{10}	0.84	-0.32	-0.78	-0.35	0.60	0.00	0.00	0.85	-0.10	1.00	
X_{11}	0.46	0.12	-0.34	0.66	0.35	0.00	-0.13	0.78	0.11	0.40	1.00

We note from the matrix of correlations between the explanatory variables in Table (1), that some variables have stronger correlations than others, which are X_1X_2) =0.87X_1X_6 ,=0.78X_1X_8 ,=0.-0.78$X_{10}X_8$,=0.84$X_{10}X_3$, =-0.78$X_{10}X_8$, =-0.85$X_{11}X_8$,=-0.78 , X_4X_3=-0.78 ,From the chi-square tables, the tabular value of the chi-square with a degree of freedom P(P-1)/2=55 and at a significant level of 0.05 n=100, and P=11 was (73.31), which is less than the calculated value, so we reject the null hypothesis, meaning that there is a problem of multiplicity This result is confirmed by the extraction of the variance inflation factor VIF, where the results were in Table (2), which showed the presence of variance inflation factors greater than (10) for the variables) .x_1, x_2,x_3,x_4,x_8,,x_{10},x_{11}:(

Table (2) The variance inflation factor for the variables

Variable	VIF	Result of Collinearity
x_1	21.322	Collinearity
x_2	19.575	Collinearity
x_3	18.582	Collinearity
x_4	17.672	Collinearity
x_5	2.826	Non- Collinearity
x_6	3.773	Non- Collinearity
x_7	2.395	Non- Collinearity
x_8	13.550	Collinearity
x_9	1.936	Non - Collinearity
x_{10}	15.936	Collinearity
x_{11}	11.236	Co-Linearity

Table (3) shows the values of the generated regression coefficients and the corresponding mean squared error at n = 100, P = 3.5, d = 0.5 and sigma = 5. If we notice that the KL estimator achieved the least mean squares error, followed by the Liu estimator, then the RR estimator and finally the Ordinary Least Squares estimator.

Variable	$\hat{\alpha}_{OLS}$	Sig.	$\hat{\alpha}_{RR}$	Sig.	$\hat{\alpha}_{Liu}$	Sig.	$\hat{\alpha}_{KL}$	Sig.
α_0	29.786	0.0171	22.7651	0.0042	28.3222	0.0001	28.6773	0.0001
α_1	-0.0589	0.6101	0.3211	0.0121	0.2322	0.0067	0.2118	0.0031
α_2	-0.3156	0.5292	0.8351	0.0059	-0.1355	0.0016	-0.3357	0.0006
α_3	-0.2311	0.0121	0.6759	0.4533	-0.2412	0.3321	-0.3097	0.6322
α_4	0.0616	0.0200	3.7778	0.0051	0.0567	0.0012	0.0115	0.0003
α_5	-0.1282	0.1128	2.5677	0.0173	-0.3422	0.0011	-0.3923	0.0001
α_6	0.0856	0.4677	3.1357	0.0046	0.0894	0.0005	0.0797	0.0002
α_7	0.0955	0.2105	1.5689	0.1223	0.1322	0.0000	0.3429	0.0000
α_8	-0.8979	0.0988	-0.2173	0.1233	-0.7976	0.2311	-0.8275	0.8312
α_9	0.3519	0.1786	0.1191	0.2133	0.3519	0.5678	0.3635	0.7676
α_{10}	1.3072	0.3876	2.5664	0.0054	1.3229	0.0003	1.4236	0.0001
α_{11}	0.5572	0.4173	-1.5134	0.0011	0.7578	0.0003	0.6274	0.0001
MSE	4787.56		2344.19		2089.31		1978.44	

7. Results discussion:

Through the results of the practical application, it was found that the OLS estimator achieved the highest mean squares error, as is evident by estimating the regression coefficients and the Sig value. For each parameter, this method showed the significance of some variables, despite their lack of importance and their non-contribution to the number of infection with the Corona virus, such as the non-significance of the variable x_5 which is the use of sterilizers, despite the importance of using sterilizers in reducing infection with the virus, and this is due to the presence of the problem of linear multiplicity between the variables that The ordinary least squares method was unable to solve this problem, and also showed the letter regression estimator, mean squares less than the least squares method, and showed good significance for the study variables . The significance of the variable x_5, which showed the least squares method, showed its insignificance, but this method showed the insignificance of the variable x_7, which is the patient's bronchial asthma, despite the importance of this variable, and thus it becomes clear that the no-letter regression method did not solve the problem of multilinearity completely. As for the Liu estimator, it showed an average error squared less than the two previous methods and showed the correct significance for all the variables, and this shows that the Liu estimator was able to solve the problem of multilinearity completely, and finally the KL estimator recorded the least mean squared error than the rest of the methods, and its significance is more accurate than the significance of the Liu estimator. This indicates that the KL estimator is better than the rest of the estimators in solving the polylinear problem.

References:

1. Al-Quraishi, Zainab Kazem Mezher, (2018), “Comparing the regression estimators of crafts, major compounds and partial least squares in the regression model”, a master’s thesis, College of Administration and Economics, University of Karbala.
2. Gujarati, Damodar, authored, Odeh, Hind Abdul Ghaffar and Al-Dash, Afaf Ali Hassan, (2015), "Econometrics", Marrakesh Publishing House, Saudi Arabia, Part One.
3. Gabriel, Mohamed Suleiman Mohamed (2014), “Polylinearity, its causes, effects, and treatment with ridge slope and principal component regression with application to hypothetical data”, PhD thesis, Republic of Sudan, Sudan University of Science and Technology, College of Postgraduate Studies, sudan
4. Hoerl, A.E. and Kennard, R.W. (1970). Ridge Regression: Biased Estimation for Non-Orthogonal Problems. Technometrics, 12, 55-67.
5. F. Lukman, K. Ayinde, S. Binuomote, and O. A. Clement, “Modified ridge-type estimator to combat multicollinearity:
6. F. Swindel, “Good ridge estimators based on prior information,” *Communications in Statistics-0eory and Methods*,
7. M. Golam Kibria & Adewale F. Lukman, (2020), " A New Ridge-Type Estimator for the Linear Regression Model:

Published by “Global Research Network LLC"
https://www.globalresearchnetwork.us

8. *Berkley Symposium on Mathematical and Statistics Probability*, J. Neyman, Ed., vol. 1, pp. 197–206, Springer, Berlin, Germany, 1956.
9. C. Stein, "Inadmissibility of the usual estimator for mean of multivariate normal distribution," in *Proceedings of the 0ird*
10. El-Dereny and N. I. Rashwan, " Solving Multicollinearity ProblemUsing Ridge Regression Models" **Int. J. Contemp. Math. Sciences, Vol. 6, 2011, no. 12, 585 – 600**
11. Donald E. Farrar and Robert R. Glauber, 1967, "**Multicollinearity in Regression Analysis: The Problem Revisited**", The Review of Economics and Statistics Vol. 49, No. 1 (Feb., 1967), pp. 92-107 (16 pages) Published By: The MIT Press https://doi.org/10.2307/1937887 https://www.jstor.org/stable/1937887
12. F. Akdeniz and S. Kaçiranlar, "On the almost unbiased generalized liu estimator and unbiased estimation of the bias
13. H. Yang and X. Chang, "A new two-parameter estimator in linear regression," *Communications in Statistics-0eory and*
14. Hoerl, Arthur .E. and Kennard, Robert W., (1970a)."Ridge regression: Biased estimation for non-orthogonal Problems", T Econometrics Journal , Vol.12, No.1,55-67
15. JOSEPH , AKINNIYI, ALABA ; SANNI, ENEJI ADEMOH, (2017), " THE FARRAR-GLAUBAR APPROACH IN TESTING FOR MULTICOLLINEARITY IN ECONOMIC DATA", International Journal for Research in Business, Management and Accounting ISSN: 2455-6114, Volume 02 Issue 01
16. K. Liu, "A new class of biased estimate in linear regression," *Communication in Statistics- 0eory and Methods*, vol. 22,
17. Rom´an Salmer´on-G´omez, Catalina B. Garc´ıa-Garc´ıa and Jos´e Garc´ıa-P´erez, (2020), " Overcoming the inconsistences of the variance inflation factor: a redefined VIF and a test to detect statistical troubling multicollinearity." , ectaart.cls ver. 2006/04/11 file: Manuscript_econometrika_def.tex date: May 6, 2020
18. S. Sakallıo˘glu and S. Kaçıranlar, "A new biased estimator based on ridge estimation," *Statistical Papers*, vol. 49, no. 4,
19. Shrestha, Noora, (2020), " Detecting Multicollinearity in Regression Analysis " , American Journal of Applied Mathematics and Statistics, 2020, Vol. 8, No. 2, 39-42 Available online at http://pubs.sciepub.com/ajams/8/2/1 Published by Science and Education Publishing DOI:10.12691/ajams-8-2-1
20. Simulations and Applications", Hindawi Scientifica Volume 2020, Article ID 9758378, 16 pages https://doi.org/10.1155/2020/9758378 vol. 12, no. 1, pp. 55–67, 1970. vol. 5, no. 11, pp. 1065–1075, 1976.