International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 08
April 2016

A Comparison Study of Ridge Regression and Principle Component Regression with Application

Gariballa Abdelmageed Abdelgadir & Hussein Eledum
Assistant prof of Econometrics
Taibah University , KSA.
drgariballa@gmail.com
Associate prof of Applied Statistics
University of Tabuk, KSA.
heledum@yahoo.com

Abstract

The purpose of this paper is to discuss the multicollinearity problem in regression models and presents some typical ways of handling the collinearity problem. In Addition, the paper attempts to compare RR , and PCR and LS methods using minimum squared error MSE and the accuracy of the prediction. The results of this paper showed that, RR method performs better than PCR and LS methods , because RR had minimum MSE and a higher predicted accuracy than other methods. The results of this paper showed that, based on the criteria of model accuracy PCR performs better than RR, whereas, according to mean squares errors criterion MSE , RR performs slightly better. In general, the two biased estimator RR and PCR perform better than LS.

Keywords: Least Squares; Correlation Matrix; Multicollinearity; Ridge Regression; Principal Component Regression.

1. Introduction

The least squares method LS is one of the oldest techniques in statistical analysis for estimating the parameters of general linear regression model, and to fit the data under some assumptions with a single or multiple explanatory variables in order to choose the best regression line which minimize the sum of the squares of errors. However, if these assumptions are violated, LS method does not assure the desirable results. The influence of the multicollinearity is one of these problems, this occurs when the number of explanatory variable is relatively large in comparison to the sample or if the variables are almost collinear, the Ridge regression method RR and Principal component Regression PCR are used to deal with it.

Many studies has been conducted to compare between methods of estimating regression parameters when data suffered from multicollinearity.

(Eledum H.,2011) used simulation technique to compare between the three biased estimation methods ridge regression (RR), principle component regression (PCR), and Latent Root LR . (Yazid M. & Mowafaq M., 2009) and (Kianoush F., 2013) were used Monte Carlo simulation to estimate the regression coefficients by ridge regression (RR) and principle component regression (PCR). (Moawad E., 2014) presented and compared the partial least squares (PLS) regression as an alternative procedure for handling multicollinearity problem with two ridge regression (RR) and principle component regression (PCR). (Ali G. et al, 2014) studded the prediction of rangeland biomass using different methods including Multiple regression, Principal Component Analysis, Partial Least Square regression and Ridge regression and compared them. (Piyush K. et al, 2013) compared ridge regression (RR) and principle component regression (PCR) estimators with the r-k class estimator, which is composed by combining the RR estimator and the PCR estimator into a single estimator.

The purpose of this paper is to discuss the multicollinearity problem in regression models and presents some typical ways of handling the collinearity problem. In Addition, the paper attempts to compare RR , PCR and LS methods using minimum squared error MSE and the accuracy of the prediction.

This paper is organized as follows. In the next section, we present the multiple linear regression; section 3 discusses the ridge regression RR and its properties. Section 4 pertains to the principal components regression PCR. The application domain explains in section 5. In section 6 we give a brief summary and conclusions.

2. Multiple Linear Regression

Suppose there is a linear relationship between dependent variable$Y$, and explanatory variables $X_1, X_2, \dots, X_p$, and error term ϵ, we can write this relationship as follows (Xin Y. & Xiao Su., 2009: 224)

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \varepsilon_i \qquad (1)$$
$$i = 1, 2, \dots, n$$

Where:

Y_i: is the $\bar{i}th$ observation of response variable.

X_{ij} : is the $\bar{i}th$ observation of explanatory variable j.

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$: are the parameters or regression coefficients.

ε_i: is a random error term or disturbance term.

In matrix form we can write Eq(1) as (Jeffrey M., 2009 : 799)

$$Y = X\beta + \epsilon \qquad (2)$$

Where
$Y: n \times 1$ vector observations of the dependent variable .
$X: n \times p + 1$ matrix of explanatory variables.(assumed to be with full rank)
$\beta: p + 1 \times 1$ vector of regression coefficients.
$\epsilon: n \times 1$ vector of uncorrelated errors . With properties $E(\epsilon) = 0$, and $\mathrm{Cov}(\epsilon) = \sigma^2 I_n$, where $I_n$represents the n dimensional identity matrix .
For the purpose of making statistical inferences, further assumed that the errors are normally distributed.
The LS estimators of β is (Carl F. & Praveen K.,2002)

$$\hat{\beta}_{LS} = (X^T X)^{-1} X^T Y \qquad (3)$$

Both LS estimators and its covariance matrix heavily depend on the characteristics of the matrix $X^T X$. If $X^T X$ is ill-conditioned, i.e. the column vectors of X are linearly dependent, LS estimators are sensitive to a number of errors. for example, some of regression coefficients may be statistically insignificant or have the wrong signs, and they may result in wide confidence intervals for individual parameters. With ill conditioned matrix$X^T X$, it is difficult to make valid statistical inferences about the regression parameters. For instance, see (Eledum H., 2011). More than one methods proposed to deal with such case, the most important of them are the Ridge regression method RR and Principal component Regression PCR.

In most statistical analysis, it is desirable to center the variables. The $X^T X$ matrix contains the variance-covariance data if data are mean centered (Zeaiter M., & Rutledge D., 2009). A detailed description of different centering methods has been given by (Bro R., & Smilde K., 2003). In this paper we use:
$Y_i = Y_i - \bar{Y}$ and $X_i = X_i - \bar{X}$ where, $\bar{Y}$ and $\bar{X}$ are means of Y and X respectively. All of the equations are written under the mean centering assumption. Thus, there is no intercept required in the models, and $X^T X$ is $p \times p$ correlation matrix between explanatory variables.

2.1 Canonical form

Since X^TX is a correlation and symmetric matrix there exist an orthogonal matrix $V = [V_1, V_2, \dots, V_p]$ is a corresponding matrix of orthogonal eigenvectors such that: see for instance (Abdalla H., 2009) :

Since X^TX is a correlation and symmetric matrix there exist an orthogonal matrix $V = [V_1, V_2, \dots, V_p]$ such that: see for instance (Abdalla H., 2009) :

$$V^T(X^TX)V = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$$

Where Λ is a matrix of eigenvalues, λ_i is the i^{th} element of Λ , and the columns of V are normalized eigenvectors associated with eigenvalues. $\Lambda_{p\times p}$ is a diagonal matrix in which $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$

Knowing that $VV^T = 1$ since V is an orthogonal matrix, thus, model (2) can be written in the canonical form as (Myres Rymond H.,1986: 363):

$$Y = Za + \epsilon \qquad (4)$$

Where $Z = XV$ and $a = V^T\beta$. The LS estimator in eq3 is given as :

$$\hat{a}_{LS} = \Lambda^{-1}Z^TY \qquad (5)$$

2.2 Multicollinearity

The term multicollinearity is due to Ranger Frisch (1933). This term defines itself, multi implying many and collinear implying linear dependences (Gujarati N. ,2004:319) .

Originally, multicollinearity meant the existence of a perfect, or exact, linear relationship among some or all independent variables of a regression model. For the p-variable regression involving independent variable $X_1, X_2, \dots, X_p$, an exact linear relationship is said to exist if the following condition is satisfied: (Myres Rymond H.,1986)

$$C_1X_1 + C_2X_2 + \cdots + C_pX_p = 0$$

Where $C_1, C_2, \dots, C_p$ are constants such that not all of them are zero simultaneously.

Or

$$\sum_{i=1}^{p} C_iX_i = 0 \qquad (6)$$

Multicollinearity is also the name we give to the problem of nearly perfect linear relationships among explanatory variables , this is the more common problem , and it said to exist if the following condition is satisfied:

$$C_1X_1 + C_2X_2 + \cdots + C_pX_p + d_i = 0$$

Where d_i is stochastic error term.

Or

$$\sum_{i=1}^{p} C_iX_i \cong 0 \qquad (7)$$

Consider the Eigen values $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_p]$ and Eigen vector $V = [V_1, V_2, \dots, V_p]$ of the correlation matrix X^TX, we can redefine the multicollinearity as follows.

Denoting the i^{th} element of the vector V_j by v_{ij}, if multicollinearity is present at least one $\lambda_i \cong 0$ thus ,

$$V_j^T(X^TX)V_j \cong 0$$

Which implies that for at least one Eigen vector V_j ,

$$\sum_{i=1}^{p} V_{ij}X_i \cong 0$$

Thus the number of small Eigen values of the correlation matrix relate to the number of multicollinearities according to the definition ineq7 and the “weights” $C_i's$ are the individual elements in theassociated eigenvectors.

There are several methods that have been proposed to remedy multicollinearity problem by modifying the method of LS to allow biased estimators of regression coefficients , these methods are ridge regression RR , and principal components regression PCR. When an estimator has only small biased it is substantially more precise than an unbiased estimator, it may be the preferred estimator since it have a large probability of being close to the true parameter value (Myres Rymond H.,1986).

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 08
April 2016

3. Ridge Regression

Ridge regression (RR) has been introduced by Hoerl and Kinnard (Hoerl, A. & Kennard, R,1970,1975), they suggested a small positive number $k \geq 0$ to be added to the diagonal elements of the $X^T X$ matrix from the multiple regression, and resulting estimator is obtained as:

$$\hat{\beta}_{RR} = (X^T X + kI)^{-1} X^T Y \qquad (8)$$

where I is a matrix unit and k is a constant selected by the analyst, $k > 0$. It is to be noted that when $k = 0$ then the ridge estimator is the least-square estimator. The ridge estimator is a linear transformation of the least-squares estimator $\hat{\beta}_{LS}$

$$\hat{\beta}_{RR} = (I_n + k(X^T X)^{-1})^{-1} \hat{\beta}_{LS} \qquad (9)$$

Using canonical form of eq 4 the ridge estimator can be written as

$$\hat{a}_{RR} = (I_n + k\Lambda^{-1})^{-1} \hat{a}_{LS} \qquad (10)$$

Mean squared error for ridge regression is

$$MSE(\hat{a}_{RR}) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)} + k^2 \sum_{i=1}^{p} \frac{a_i^2}{(\lambda_i + k)^2} \qquad (11)$$

Where σ^2 is error variance and a_i is the i-th component of a.

3.1 Selection methods of ridge parameter

The popular method of selecting value of ridge parameter k is ridge trace. Furthermore, there are another tenth methods, the most important of them are:

Horl and Konard's method:

Horl and Kenard (1970) proposed estimating ridge parameter as follows:

$$k_{HK} = \frac{\hat{\sigma}^2}{Max(\hat{a}^2{}_{LS_i})} \qquad (12)$$

Lawless and Wang's method

Lawless and Wang (1976) proposed the following ridge parameter estimation:

$$k_{LW} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p} \lambda_i \hat{a}^2{}_{LS_i}} \qquad (13)$$

4. Principal components Regression

This procedure proposed by Harold Hotelling in 1933 (Massy,1965). In principal components regression method, instead of using regression variables, principal components are used as regression variables. Thus, the replaced regression variables are independent from each other. In principal components regression model, a subset of principal components is used instead of all components. The method varies somewhat in philosophy from ridge regression but like ridge, gives biased estimates, when using successfully this method results in estimation and prediction will be superior to LS.

Assume q first components are used in regression model ($q < p$); then, a is estimated as follows:

$$\hat{a}_q = \left(Z_q^T Z_q\right)^{-1} Z_q^T Y \qquad (14)$$
$$= \Lambda_q^{-1} V_q^T X^T Y$$

So that $Z_q = XV_q$ and Λ_q are diagonal matrix of q first eigenvalues (where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_q$) and V_q is a matrix with q corresponding eigenvector. In section 2.1 a is defined as $a = V^T \beta$. Then, $\beta = Va$ can be written and estimated value of β using principal component method is equal to:

$$\hat{\beta}_{PC} = V\hat{a} \qquad (15)$$

and by replacing $\hat{a}$ with its value in equ 14 with $\hat{\beta}_{PC}$, the following is given for the reduced model

$$\hat{\beta}_{PC} = V_q \Lambda_q^{-1} V_q^T X^T Y \qquad (16)$$

Mean squared error for principal components regression is

$$MSE(\hat{a}_{RR}) = \sigma^2 \sum_{i=1}^{q} \frac{1}{\lambda_i} + k^2 \sum_{i=q+1}^{p} (V_i^T \beta)^2 \qquad (17)$$

where V_i that is the i-th vector of eigenvalues from matrix $X^T X$.

There are different ideas about selecting the number of components for presence in regression model. In reality, the number of components extracted in a principal component analysis is equal to the number of observed variables being analyzed. However, Kasier (1960) proposed leaving the components with special values of greater than 1 in the model. Mansfileld et al (1977) suggested that only the first few components account for meaningful amounts of variance, so only these first few components are retained and used in multiple regression analyses. Frontier (1976) and Legendre (1983) proposed a Broken Stick model and is defined by:

$$b_k = \sum_{i=k}^{p} \frac{1}{i} \qquad k = 1,2, \ldots\ldots, p \qquad (18)$$

so b_k is a criterion for selecting the number of principal components in the model.

Values of λ_i, s are compared with the corresponding values of $b_k$from large to small, respectively. For the first value of k which $b_k > \lambda_i$, the comparison is stopped and eigenvalue before *k*, i.e. $\lambda_{k-1}, \ldots.., \lambda_1$, are remained in principal components regression model.

5. Application

5.1 Data collection

The data source in this study is based on the data used by the researcher (Nawal H., 2011) in her research, which was published in the Journal of Anbar University of Economic Sciences and Administration, Volume 4 - Issue 7 (2011). Iraq, and so entitled : Using the Method of the Mutual Integration analysis to show the effect of the monetary and real variables in inflation through Period (1970 – 2007). While noting that the researcher was not exposed at all to discuss the issue of multicollinearity during her study. The dependent variable in the model is inflation INF and the independent variables are Gross Domestic Product (*GDP*), the governmental expenditure (*Gov*),The money supply (*Mon*) and the Exchange Rate (*Exr*).The inflation model can be expressed as

$$INF = B_0 + B_1 GDP + B_2 Gov + B_3 Mon + B_4 Exr + \varepsilon_i \qquad (19)$$

The NCSS 10 Statistical Software used to analyze the data.

5.2 Method

First, The correlations matrix among variables were calculated to test the linear relationship between these variables. Then least squares LS method was conducted to construct a linear model between an Inflation rate (*INF*) and (*GFP*, *Gov*, *Mon* and *Exr*).

In order to lay the foundation for detection of multicollinearity problem, some classic symptoms are present in our data:

From table 2 we can see that *F* is highly significant (p-value-0.000), implying that the variables chosen are valid explanatory variables, and the regression coefficient for *Mon* is insignificant at 5% level of significance, also value of R^2 is quite large, i.e., **0**.962. Variance inflation factor (VIF,s) for all variables are greater than 4, and almost all the conditional numbers are greater than 1. From Table 1, which represents the correlation matrix among explanatory variables, we can see that how the explanatory variables are correlated. Variable *Gdp* is highly positive correlated with variable *Mon* and Exr , i.e., 0.986 and 0.830 respectively. Similarly the Correlation between Gov and *Mon* is 0.624 and between *Exr* and *Mon* is 0.782.

Table 1: Correlation Matrix of explanatory variables

	Gdp	*Gov*	*Mon*	*Exr*
Gdp	1	0.572	0.986	0.830
Gov		1	0.624	0.325
Mon			1	0.782
Exr				1

Table 2: The Least Square estimated model results and collinearity statistics

Mode1 / Variables	Unstandardized				Standardized	Collinearity statistics	
	coefficient	SE	t	P Value	coefficient	VIF	C
Constant	23.458	4.438	5.29	0.000	0		
Gdp	1E-08	2.5E-09	2.832	0.008	0.7329	58.2122	1.00
Gov	-0.003616	0.001404	-2.575	0.015	-0.12113	1.9223	4.43
Mon	0.000101	0.001744	0.058	0.954	0.0142	51.7195	17.43
Exr	52.045	10.65	4.89	0.000	0.33082	3.9788	337.46
R2=0.962 S=5.941 F= 209 , p-value: 0.0000							

5.3 Application of Ridge Regression

To determine the best model fitted the data using ridge regression, firstly we present methods of choosing *k*.

Table 3 below summarizes the results of Ridge Regression for each method of selecting *k* for Inflation rate Data.

Table3: The results of Ridge Regression for each method of selecting k for Inflation rate Data

Variable	HKB		LW		Ridge trace	
	K=0.132229		K=0.09090848		K=0.489	
	unstandardized	standardized	unstandardized	standardized	unstandardized	standardized
Intercept	23.51985	0	23.26801	0	27.13074	0
Gdp t	3.517584E-09 (8.78)	0.3640	3.696974E-09 (8.14)	0.3826	2.864229E-09 (9.56)	0.2964
Gov t	-0.002459046 (- 1.57)	-0.0824	-0.002865748 (-1.86)	-0.0960	-0.0004179146 (-2.72)	-0.0140
Mon t	0.002097501 (6.56)	0.2934	0.002091695 (5.87)	0.2926	0.001890136 (8.00)	0.2644
Exr t	56.40065 (6.1)	0.3585	57.04361 (5.99)	0.3626	48.87162 (6.48)	0.3107
R^2	0.96185		0.96193		0.96122	
MSE	34.07		34.00		34.64	
Max VIP	1.2567		1.1611		0.4440	

In ridge trace method we start from *k*=0 and then after taking three values 0.001, 0.002, 0.005 for *K*, we give the equal space of 0.01. We plot the regression coefficient against *k* in figure 1. The system has been stabilized at $0.4 < k < 0.45$ i.e. *k*= 0.489 is the ridge parameter.

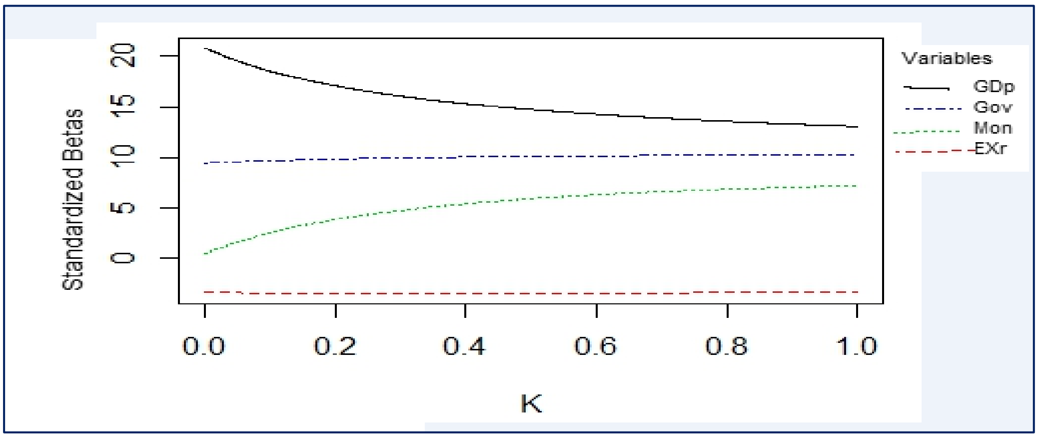

Figure 1: The values of the estimated regression coefficients plotted Against *k* with using ridge trace method

From table 3 above, we see that all methods for selecting k provided estimated models with significant regression coefficients and high values of determination coefficient R^2. Furthermore, the problem of multicollinearity disappeared in all these models because all maximum VIF's were less than 4.

Also noting that, Lawless and Wang's method LW considers as the best method of choosing k i.e. k=0.09091, because it has a minimum value of MSE i.e. 34.00. Therefore the estimated model of inflation by this method is:

$$IN\hat{F} = 23.27 + (3.69E - 09)GDP - 0.0029Gov + 0.0021Mon + 57.04Exr \qquad (20)$$

t	8.14	-1.86	5.87	5.99
$p.v$	0.000	0.038	0.000	0.0000

$R^2 = 0.96193$ Max VIF $= 1.1611$

5.4 Application of The Principal Components Regression

In PCR model, the eigenvalues of the correlation matrix and their corresponding eigenvectors are shown in table 4, Here the first principal components $(0.555, 0.389, 0.544, 0.484)^T$, corresponding to the largest eigenvalue and accounting for 77.8% of the variance is designated as PC1. The first component of this vector denotes the contribution of the first objective function towards this vector; the second component denotes the contribution of the second objective and so on.

Table 4 :The Eigenvalues and Eigenvectors

Eigenvalues (λ_i)	3.1107	0.7016	0.1785	0.0092
Proportion	0.778	0.175	0.045	0.002
Cumulative Proportion	0.778	0.953	0.998	1
Eigenvectors				
Independent Variables	PC_1	PC_2	PC_3	PC_4
Gdp : X_1	0.555	0.125	-0.386	-0.727
Gov :X_2	0.389	-0.850	0.353	-0.037
Mon :X_3	0.544	0.026	-0.478	0.681
Exr :X_4	0.484	0.511	0.706	0.083

After obtaining, the coefficients related to the four principal components to create expressions of four PCs as

$$Z_1 = 0.555X_1 + 0.389X_2 + 0.544X_3 + 0.484X_4 \qquad (21)$$

$$Z_2 = 0.125X_1 - 0.850X_2 + 0.026X_3 + 0.511X_4 \quad (22)$$
$$Z_3 = -0.386X_1 + 0.353X_2 - 0.478X_3 + 0.706X_4 \quad (23)$$
$$Z_4 = -0.727X_1 - 0.037X_2 + 0.681X_3 + 0.083X_4 \quad (24)$$

As shown in table 4, the variance of Z_4 is $\lambda_4 = 0.0092$ which is small and can be taken to be approximately zero. This implies that the variable Z_4 is approximately constant, and hence is equal to its mean. Since Z_4 is a linear function of standardized variables X_i , Z_4 has a mean zero. It follows that the variable Z_4 is itself approximately zero and is the source of multicollinearity. Let us exclude Z_4 and regress Y on Z_1, Z_2, Z_3. The possible regression to be consider are

$$Y = \sum_{i=1}^{k} a_i Z_i + \epsilon \quad , k = 1,2,3 \quad (25)$$

Each of these models will lead to estimates all four of the original coefficients $\beta_1, \beta_2, \dots, \beta_4$. These estimates will be biased since Z_4 has been excluded in all cases. The inclusion of Z_4 would produce exactly the same estimates as were obtained by using the LS regression of Y on all the four independent variables given in table 2.

Based on The matrix $Y = Za + \epsilon$ the regression equation of the dependent variable (y) on the all-principal components Z's is

$$y = 0.530\, Z_1 + 0.364\, Z_2 - 0.101 Z_3 - 0.487\, Z_4 \quad (26)$$

t	27.40	8.99	− 1.26 6	− 1.38
$p.v$	0.000	0.000	0.217	0.177

Setting $a_1 = 0.530$, and $a_2 = a_3 = a_4 = 0$ in equation β=Va we get estimated β's corresponding to the regression on the first principal component only. The estimated β's corresponding to the first two principal components are obtain by setting $a_1 = 0.530$, and $a_2 = 0.364, and\ a_3 = a_4 = 0$, The estimates of β's corresponding to other regression are obtained in similar way. The results of all these regressions on different numbers of principal components are given in table **5**

Table 5:The results of Principal components regression for Inflation rate Data

Variable	Column (1)		Column (2)		Column (3)		Column (4)	
	First P.C K=1		First 2 P.C K=2		First 3 P.C K=3		All 4 P.C K=4	
	unstandardized	standardized	unstandardized	standardized	unstandardized	standardized	unstandardized	standardized
Interceptt	20.07	0	18.46	0	23.29	0	23.46	0
Gdp t	2.83E-09 (14.55)	0.2926	3.27E-09 (27.24)	0.3381	3.63E-09 (11.03)	0.3761	1E-08 (2.832)	0.7329
Gov t	6.13E-03 (14.55)	0.2052	-3.11E-03 (-2.81)	-0.1042	-4.15E-03 (-2.98)	-0.1391	-0.003616 (-2.575)	-0.1211
Mon t	2.09E-03 (14.55)	0.2921	2.16E-03 (26.8)	0.3015	2.49E-03 (8.5)	0.3486	0.000101 (0.058)	0.0142
Exr t	40.13 (14.55)	0.2551	69.39 (18.50)	0.4411	58.43 (5.91)	0.3714	52.05 (4.89)	0.3308
R^2	0.865		0.958		0.959		0.962	
Max VIP	0.099		1.078		3.238		58.212	

The table shows that the difference in results obtained by using different numbers of principal components are quite substantial. As was mentioned before, the estimates in the last principal component equation(column 4) involving all the four possible principal components are the same as the OLS estimates.

Now we decide the number of principal components to be included in to the equation(26). The criteria used here for choosing the best principal components are the first k principal components which explain more than 85% of the total variance and those components statistically significant at 5 percent level. From table 2. The first principal component explains 77.8% of the total variation in the inflation

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 08
April 2016

rate. The first two principal components collectively explain 95.3% of the total variation in the inflation rate. Consequently, sample variation is summarized very well by two principal components. In addition. The coefficients of Z_1 and Z_2 are statistically significant at 5 percent level in equation (26). Thus, we selected the model that based on the only first two principal components. The coefficients estimate using the principal components regression in terms of the original variables gives

$$Inf = 18.46 + 3.27\text{E}-09Gdp - 0.00311Gov + 0.00216Mon + 69.39Exr \qquad (27)$$

t	27.40	− 2.81	26.8	18.50
$S.e$	$1.19E-10$	− 0.0011	0.00008	3.75
VIF	0.121	1.078	0.099	0.447

$R2 = 0.958$

5.5 Comparison among Least Squares, Ridge Regression, and Principal Components Regression

From Table 6, we see that the Multicollinearity problem between the independent variables for the inflation model has been solved by using ridge regression RR and principal components regression PCR. According to table 6, at all three methods the sign of the variable (Gov) is found to be contrary to economic theory. While the parameters of other independent variables for RR and PCR regression methods are compatible with economic theory and statistically significant, this means that the variables that have significant effect on inflation rate in Jordan during the period (1970-2007) are: Gdp, Mon, and Exr.

Table 6: The Results of OLS, RR, PCR

	Estimated of Parameters		
Explanatory Variables	**OLS**	RR	**PCR**
Gdp	1E-08	3.696974E-09	3.27E-09
t	2.832	8.14	27.24
VIF	58.2122	0.9620	0.121
Gov	-0.003616	-0.002865748	-0.00311
t	-2.575	-1.86	-2.81
VIF	1.9223	1.1611	1.078
Mon	0.000101	0.002091695	0.00216
t	0.058	5.87	26.8
VIF	51.7195	1.0824	0.099
Exr	52.045	57.04361	69.39
t	4.89	5.99	18.50
VIF	3.9788	1.5939	0.447
Constant	23.458	23.26801	18.46
R^2	0.962	0.9242	0.958
Max VIF	58.21	1.1611	1.078
C.V	0.085	0.083	0.089
MSE	35.295	34.006	38.978

When we compare the results of PCR method with the results of the RR in table 6, we found that PCR is better than the RR, based on the following criteria:

- The calculated values of the t-test for all parameters according to PCR are larger than those calculated using RR method.
- The value of R^2 in PCR is greater than its value in RR.
- Max VIF value of the PCR is lower than in the RR method.

On the other hand, the RR method is considered better than the PCR method, according to the following criteria:

- The value of the coefficient of variation (C.V) of RR is less than that of PCR.
- The value of the MSE of RR is less than that of PCR method.

Generally, the results of two method RR and PCR are better than that of LS based on criteria above.

6. Conclusions

In this paper, we have given a description of the algorithms of the two most popular biased regression methods RR and PCR. Then we discussed the connections and differences among them. RR and PCR, are proposed regression methods under collinear situation when OLS fails. They make a compromise between accuracy and precision. Their estimators are biased but with a lower mean squared estimation error. Furthermore, RR and PCR are applied to an inflation data set . The solution and performance of RR and PLR, tend to be quite similar in those practical data sets we applied. Based on the criteria of model accuracy PCR performs better than RR, whereas, according to mean squares errors criterion MSE , RR performs slightly better. In general the two biased estimator RR and PCR perform better than LS.

Reference

[1] Abdalla H. (2009).A Simulation Study Of Ridge regression Method With Auto correlated Errors. Shendi University Journal . Issue No 6, 1-19.

[2] Ali G., Azin Z., & Pejman T. (2014). Comparing Multiple Regression ,Principal Components Analysis ,Partial Least Square Regression and Ridge Regression in Predicting Rangeland Biomass in the Semi Steppe Rangeland of Iran. Environment and Natural Resources J .12(1), 1-21.

[3] Bro R. & Smilde K. (2003). Centering and scaling in component analysis. Journal of Chemo metrics. 17(1), 16–3.

[4] Carl F. & Praveen K. (2002). The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations. Applied Econometrics. Vol 34(6), 667-677.

[5] Eledum H. (2011). Biased Estimation Methods with Autocorrelation using Simulation. LAMBERT Academic Publishing. ISBN-133844324761-978 .

[6] Frontier, S. (1976). Étude de la décroissance des valeurs propres dans une analyse en composantes principales: comparison avec le modèle du bâton brisé. J. Exp. Mar. Biol. Ecol. 25: 67-75.

[7] Gujarati N. Damodar (2004). Basic Econometrics. Fourth Edition. McGraw-Hill. p 319.

[8] Hoerl A. E. & Kennard R.W.(1970). Ridge Regression Biased Estimation for Nonorthognal Problems. Technometrics. 12(1), 55-67 .

[9] Hoerl A. E.,Kennard R.W., & Baldwin K.F (1975). Ridge Regression: Some Simulations. Commun, Stat. 4(2), 105-123.

[10] Jeffrey M. (2009).Introductory Econometrics A modern Approach. Fourth edition. South-Western-Cengage Learning. USA. P 799.

[11] Kaiser, H.F. 1960. The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20: 141-151.

[12] Kianoush F. (2013). Comparing Ridge Regression and Principal Components Regression by Monte Carlo Simulation Based on MSE. Journal of Computer Science & Computational Mathematics. 3(2), 25-29.

[13] Lawless J.F. & Wang P. (1976). A simulation study of ridge and other regression estimators. Communications in Statistics –Theory and Methods. 14, 1589-1604.

[14] Legendre, L. & Legendre, P. (1983). Numerical ecology. Elsevier Scientific Publishing Co. New York. NY.

[15] Mansfield E. R., Webster J. T. & Gunst R. F.(1977). An Analytic Variable Selection Technique for Principal Components Regression. Applied Statistics. 26(1), 34-40.

[16] Massy W.(1965). Principal Components Regression in Exploratory Research. Journal of the American Statistical Association. 60(309), 243-265.

[17] Moawad B. (2014) .A Note on Partial Least Squares Regression for Multicollinearity (A Comparative Study). International Journal of Applied Science and Technology Vol. 4(1), 163-170.

[18] Myres Rymond H.. (1986).Classical and Modern Regression with Application. Ist edition. PWS Kent Publishers. ISBN 13.9780871509468 .

[19] Nawal H. (2011). Using the Method of the Mutual Integration analysis to show the effect of the monetary and real variables in inflation through Period (1970 – 2007). Journal of Anbar University of Economic Sciences and Administration. 4(7), 179-189.

[20] Piyush K., Sarla P., & Hemlata J.(2013). Regression Analysis of Collinear Data using r-k Class Estimator:Socio-Economic and Demographic Factors Affecting the Total Fertility Rate (TFR) in India. Journal of Data Science. Vol 11, 323-342.

[21] Xin Y. & Xiao Su.(2009).Linear Regression Analysis Theory and Computing. World Scientific. Singapore. p 224.

[22] Yazid M. & Mowfag M. (2009). A Monte Carlo Comparison between Ridge and Principal Components Regression Methods. Applied Mathematical Sciences. 3(42), 2085 – 2098.

[23] Zeaiter M. & Rutledge D.(2009). Preprocessing Methods. Comprehensive Chemometrics. vol.3, 139.