IBN AL- HAITHAM J. FOR PURE & APPL. SCI. VOL.23 (3) 2010

# Bayesian Analyses of Ridge Regression Prooblems

**H. M. Gorgees**

**Department of Mathematics, College of Education Ibn-Al-Haitham,University of Baghdad**

## Abstract:

A Bayesian formulation of the ridge regression problem is considerd, which derives from a direct specification of prior informations about parameters of general linear regression model when data suffer from a high degree of multicollinearity.A new approach for deriving the conventional estimator for the ridge parameter proposed by Hoerl and Kennard (1970) as well as Bayesian estimator are presented. A numerical example is studied in order to compare the performance of these estimators.

**Introduction:**

The problem of multicollinearity exists when there exists a linear relationship or an approximate linear relationship among two or more explanatory variables.

Multicollinearity can be thought of as a situation where two or more explanatory variables in the data set move together. As a consequence it is impossible to use this data set to decide which of the explanatory variables is producing the observed change in the response variable. No treatment of the data or transformation of the model will cure this deficiency. Consequently,the best way to deal with multicollinearity may be to find a different data set, or additional data to break the association between the related variables.

However some multicollinearity is nearly always present, but the important point is whether the multicollinearity is serious enough to cause appreciable damage to

the regression. Indicators of multicollinearity include a low determinant of the information matrix, a very high correlation among two or more explanatory variables , very high correlations among two or more estimated coefficients,

and significant regression of one explanatory variable on one or more explanatory variables.

**Key words**: linear regressioin model , multicollinearity , ridge regression , generalized shrinkage estimators , Bayesian estimator , singular value decomposition.

This paper deals with multicollinearity in the classical linear regression model

y=xβ+u …..(1)

Where y is an $(n \times 1)$ $vector$ of observations on the response variable, x = $(x_{ij}), i = 1,2,\ldots,n, j = 1,2,\ldots,p$ is an $(n \times p)$ matrix and of full column rank, β is a $(p \times 1)$ parameter vector (vector of unknown regression coefficients) and u is an $(n \times 1)$ vector of random disturbances, E(u)=0 and var(u)=$\sigma^2$ I, and both β and $\sigma^2$ are unknown. The least squares estimator of β is: *(see[1])* $b_{LS}=(x'x)^{-1}$x'y …..(2)

Where $b_{LS}$ denote the least squares estimator of β. The two key properties of $b_{LS}$ are that it is unbiased,E($b_{LS}$)=β, and that it has minimum variance among all linear unbiased estimators.The mean square error of $b_{LS}$ is:

*MSE($b_{LS}$)=* $\sigma^2 \sum_{i=1}^{p} \frac{1}{d_i}$ *……(3) (see[2])*

Where $d_i$"s are the eigenvalues of x'x and $d_1 > d_2 > \ldots \geq d_p > 0$ .. If the smallest eigenvalue of x'x is very much smaller than 1, then a seriously ill_conditioned (or multicollinearity) problem arises. Thus,for ill_conditioned data, the least squares solution yields coefficients whose absolute values are too large and whose signs may actually reverse with negligible changes in the data. That is in the case of multicollinearity the least squares estimator $b_{LS}$ can be poor in terms of various mean squared error criterion. Consequently, a great deal of work has been done to construct alternatives to the least squares estimator when multicollinearity is present. In the seventies Hoerl and Kennard introduced a class of biased estimators for parameters in general linear regression model labeled ridge estimators as a rival to the least squares estimator when sample data are affected by a high degree of multicollinearity.

The outhers show that in any given problem there is at least one member of this class which has total mean square error smaller than the total variance of the corresponding least squares estimator. The ridge estimator depends crucially

upon an exogeneous parameter, say k. For any k $> 0$ the corresponding ridge

estimator denoted by $b_{RR}$ is defined to be:

$b_{RR}=(x'x + kI)^{-1}$x'y…….(4) ( see[1])

We argue that it is typically true that there is available prior informations about the parameters, and this may be exploited to find improved estimators. In this paper attention is focuced upon Bayesian formulation for generalized shrinkage estimators and ordinary ridge regression estimator,moreover a Bayes estimator as well as a conventional estimator for the ridge parameter is derived.

**Generalized Shrinkage Estimators:**

Given an ($n \times p$) matrix of regressors x and an ($n \times 1$) vector of the corresponding response y. Assume that sample means have been removed from the data (so that 1'x=0'and 1'y=0, where 1 is an n_vector of ones.) and write the standard linear

regression model as E(y|x)=xβ and var(y|x)=$\sigma^2$ (I-11'/n) where β is a vector of unknown regression coefficients and $\sigma^2$ is the unknown error

variance. The singular value decomposition of x will be denoted by: *(see[3]).*

X=H$D^{1/2}$G' …….(5)

Where H is an ($n \times p$) semi orthogonal matrix satisfy H'H=$I_p$ , $D^{1/2}$ is a ($p \times p$) diagonal matrix of orderd singular values of x, $d_1^{1/2} \geq d_2^{1/2} \geq \ldots \geq d_p^{1/2}$> 0, G is a ($p \times p$) orthogonal matrix whose columns represent the eigenvectors of the information matrix x'x. Assume that $d_p$>0 so that β is estimable, then as in" Obenchain (1978)" *(see [4])* we get $b_{LS}$=GC where C= $D^{-1/2}$H'y contains the uncorrelated components of $b_{LS}$, where E(C)=E(G'$b_{LS}$) = G'β =γ say, and:

Var(C)=var(G'$b_{LS}$)=G'var($b_{LS}$)G=$\sigma^2$G'$(GDG')^{-1}$G=$\sigma^2 D^{-1}$.

Notice that the elements of C are uncorrelated since their variance matrix is diagonal. The vector of generalized shrinkage estimator (or generalized ridge regression estimator) will be denoted here by $b_{SH}$ and will be of the general form : *(see[4])*

$b_{SH}$=GΔC=$\sum_{j=1}^{p} g_j' \delta_j c_j$ …….. (6)

Where $\overrightarrow{g_j}$ is the j_th column of the matrix G, $\delta_j$ is the j_th diagonal element of

the shrinkage factors matrix Δ, we will usually restrict the range of shrinkage factors to 0≤ $\delta_j \leq 1$, j= 1,2, ….,p. $c_j$ is the j_th element of the uncorrelated components vector C. In the remaining of this section we discuss Bayesian methods for defining the form of shrinkage of sample estimates towards a subjective prior distribution.Lindely and Smith (1972) describe a Bayesian formalizim for hierarchical (multi_stage) analysis of linear models using conjugate multivariate_normal prior distribution.*(see [5])* This formalism expresses unknown parameters at each stage of an analysis in terms of a linear model at the previous lower stage . But, although dispersion matrices at each stage can be arbitrary , they must be known. And, at the final stage, both the mean vector and the dispersion matrix must be known.The fundamental lemma of Lindely and Smith (1972) states that:

*Lemma:* If the sampeling distribution of the response,y, is

$y|\theta_1 \sim N(A_1\theta_1, V_1)$ where $\theta_1$ is ( $p_1 \times 1$) parameter vector and the prior distribution is

$\theta_1 \mid \theta_2 \sim N(A_2\theta_2, V_2)$ where $\theta_2$ is ( $p_2 \times 1$) parameter vector, then the marginal (unconditional) distribution of y is:

$y \sim N(A_1 A_2\theta_2 , V_1 + A_1 V_2 A_1')$ ……(7)

and the posterior (conditional) distribution of $\theta_1$ given y is:

$\theta_1 |y \sim N(Bb , B)$ ………(8)

Where : $B^{-1} = A_1'V_1^{-1}A_1 + V_2^{-1}$ and $b = A_1'V_1^{-1}y + V_2^{-1}A_2\theta_2$

To apply this lemma and demonstrate that a simple 2_stage Bayesian formalism produces generalized shrinkage estimators, we first make the identification:

$A_1\theta_1 = x\beta$ , $V_1 = \sigma^2 I$ and $B^{-1} = \sigma^{-2}x'x + V_2^{-1}$. Next, we set the prior mean value for β to zero by taking $\theta_2 = 0$ and assume that $V_2^{-1}$ (and $V_2$) will be simultaneously diagonalizable with x'x by restricting attention to prior variance_covariance matrices of the general form $V_2 = \sigma^2.GK^{-1}G'$, where K is

a diagonal p×p matrix and G is defined as in (5) . Now the Bayes estimate is the mean Bb of the posterior distribution of β given y and this mean vector is of the

general form:

$E(\beta|y) = Bb = (A_1'V_1^{-1}A_1 + V_2^{-1})^{-1}(A_1'V_1^{-1}Y + V_2^{-1}A_2\theta_2)$

$= (\sigma^{-2}.x'x + \sigma^{-2}GKG')^{-1}\sigma^{-2}x'y$

$= (x'x + GKG')^{-1}x'y$

$= (GDG' + GKG')^{-1}GD^{1/2}H'y$

$= G(D + K)^{-1}D^{1/2}H'y$

$=G(D+K)^{-1}DD^{-1/2}H'y = G\Delta C = b_{SH}$ …….(9)

Where $\Delta = (D+K)^{-1}D$ is the diagonal matrix of generalized shrinkage factors and C is the vector of uncorrelated components of the least squares estimator.

**The Ordinary Ridge Regression Estimator:**

In the previous section we have demonstrated that all generalized shrinkage estimators are 2_stage Bayes estimators. This include of course the case of

ordinary ridge regression estimator proposed by Hoerl and Kennard (1970). To demonstrate this fact let us assume that an orthogonal matrix P is given such that $P'(x'x)^{-1}P = D = \text{diag}(d_1,\ldots,d_p)$, $d_1 \geq \ldots \geq d_p > 0$. Morever, suppose that

$Z = P'b_{LS} = (z_1,\ldots,z_p)'$ and $W=P'\beta = (w_1,\ldots,w_p)'$ then: $Z \sim N_p(W,\sigma^2 D)$. Let us assume that W has a prior distribution given by $W \sim N_p(0,\sigma^2\lambda I)$ for some positive constant $\lambda$, thus, according to the lemma stated in section 2, the posterior distribution of W given Z is:

$W|Z \sim N_p[\sigma^2\lambda(\sigma^2 D+\sigma^2\lambda I)^{-1}Z,\ \sigma^2\lambda I-\sigma^4\lambda^2(\sigma^2 D+\sigma^2\lambda I)^{-1}]$ …….(10)

then the

Bayes estimator of W is:

$\hat{W} = (\lambda^{-1}D+I)^{-1}Z$ for $Z=P'b_{LS}$. Consequently, the Bayes estimator of $\beta$ is:

$P\hat{W} = P(\lambda^{-1}D+I)^{-1}P'b_{LS}$

$=(\lambda^{-1}PDP'+I)^{-1}b_{LS}$

$=[\lambda^{-1}(x'x)^{-1}+I]^{-1}b_{LS} = [k(x'x)^{-1}+I]^{-1}(x'x)^{-1}x'y$ for $k=\lambda^{-1}$

$= (x'x+kI)^{-1}x'y = b_{RR}$ ……(11).

It is obvious that the estimator in (11) coincide with the ordinary ridge regression estimator given in (4).

**Estimating The Ridge Parameter_Bayesian Approach:**

There are many different methods for selecting the value of the ridge parameter k. The method we try here is based upon the lemma stated in section (2). Accordingly,the marginal distribution of Z is:

$Z \sim N_p(0,\sigma^2 D+\sigma^2\lambda I)$

$E(Z'D^{-1}Z) = trD^{-1}var(Z) = trD^{-1}(\sigma^2 D+\sigma^2\lambda I) = \sigma^2 trI_p+\lambda\sigma^2 trD^{-1}$ .....(12)

where "tr" denotees the trace of the matrix.To find an unbiased estimator for λ, the following result is necessary. *(see[6])*

$E(\frac{\sigma^2}{s^2}) = \frac{n-p-1}{n-p-3}$ ........(13)

Where $s^2$ is an estimator of $\sigma^2$. The result in (13) can be easily proved by setting (n-p-1)$s^2/\sigma^2$=r then r ~ $\chi^2_{(n-p-1)}$

$E(\frac{\sigma^2}{s^2}) = E(\frac{n-p-1}{r}) = \frac{n-p-1}{n-p-3}$ *(see [7] page 176)*

From (12) and (13) an unbiased estimator for λ can be obtained as:

$\hat{\lambda} = [\frac{n-p-3}{n-p-1}(\frac{Z'D^{-1}Z}{s^2}) - p]/trD^{-1}$ or equivalently:

$\hat{\lambda} = \hat{k}^{-1} = [\frac{n-p-3}{n-p-1}(\frac{b_{LS}'x'x\,b_{LS}}{s^2}) - P]/tr(x'x)$ .........(14)

**Estimating the ridge parameter_New Approach:**

Substituting from Δ in formula (6) by $(D+kI)^{-1}$D it can easily be shown that the ordinary ridge regression estimator is a member of generalized shrinkage estimators class, in such a case the shrinkage factors $\delta_i$ of ordinary ridge regression estimator will have the form :

$\delta_i = \frac{d_i}{d_i+k}$ , $1 \leq i \leq p$ .........(15)

In (1970) Hoerl and Kennard proposed an estimator for the ridge parameter given as: $\hat{k}_{HK} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}$ ........(16) [ see[ 8 ])

In this section, a new approach is used to derive $\hat{k}_{HK}$ given in (16) , this approach is represented by minimizing MSE($b_{GH}$) as follows:

MSE($b_{GH}$) = MSE(GΔC) = GMSE(ΔC)G'

Where MSE(ΔC) = $\sigma^2\Delta^2 D^{-1}$+(I-Δ)γγ'(I-Δ) Is the mean squared error matrix of ΔC with ith diagonal element given as: *{ see[3]}*

$MSE(\hat{\delta}_i c_i) = \sigma^2 \delta_i^2 / d_i + (1 - \delta_i)^2 \gamma_i^2$ ……….(17)

Differintiating $MSE(\hat{\delta}_i c_i)$ with respect to $\hat{\delta}_i$ we obtain:

$$\frac{\partial MSE(\delta_i c_i)}{\partial \delta_i} = 2\sigma^2 \delta_i / d_i - 2(1-\delta_i)\gamma_i^2 \quad \text{……….} \quad (18)$$

While the second partial derivative is nonnegative constant it follows that equating the derivative in (18) to zero will yield a minimum value for $MSE(\hat{\delta}_i c_i)$ this optimal amount of shrinkage for the i_th uncorrelated component $c_i$ is:

$$\delta_i^{MSE} = \frac{\gamma_i^2}{\gamma_i^2 + (\frac{\sigma^2}{d_i})} = \frac{d_i}{d_i + \frac{\sigma^2}{\gamma_i^2}} \quad \text{……..} \quad (19)$$

We derive the formula in (19) as follows:

$2\sigma^2 \delta_i / d_i - 2(1-\delta_i)\gamma_i^2 = 0$ then $\sigma^2 \delta_i / d_i = \gamma_i^2 - \delta_i \gamma_i^2$ hence $\frac{\sigma^2}{d_i} = \frac{\gamma_i^2}{\delta_i} - \gamma_i^2$ which implies that $\frac{\sigma^2}{d_i} + \gamma_i^2 = \frac{\gamma_i^2}{\delta_i}$ then solving for $\delta_i$ we obtain the required result.

Now we suggest comparing the shrinkage factor of ridge regression estimator given in equation (15) with $\delta_i^{MSE}$ given in equation (19), hence we conclude that the value of the ridge parameter k must be equal to $\sigma^2/\gamma_i^2$. Since each of $\sigma^2$ and $\gamma_i^2$ is unknown, we can use their estimated values, thus :

$$\hat{k} = \frac{s^2}{b_{LS}' G' G b_{LS} / P} = \frac{p s^2}{b_{LS}' b_{LS}} \quad \text{……….(20)}$$

Where $s^2$ is the residual mean square in the analysis of variance table obtained from the standard least squares fit.

**Numerical Example:**

In this section a data set suffer from a high degree of multicollinearity is used. The data are from Aljibori (2004) *(see [11] for details)*. It was designed to measure the effect of five explanatory variables $x_1, x_2, \ldots, x_5$ on the response variable y. Where the explanatory variables represent the number of , managerials , technicians, skilled workers, unskilled workers and service workers respectively, while the response variable y represents the productivity of the industrial sector in Iraq measured by the surplus value method for the period 21 years from 1970 to 1990.Our purpose is only to compare the performance of Bayesian and conventional estimators for the ridge parameter. The original data are presented in table (1).

For this data we found that the estimated value of the ridge parameter was $\hat{K}$ = 0.125260 and $\hat{K}$= 0.494107 obtained by applying the formula in (14) and in (20) respectively. In order to make the ridge regression analysis we used the numeric calculation statistical system (NCSS) and the results were given in table (2) through table (6).

Pearson correlations were given for all variables in table (2). These correlation coefficients show which explanatory variables are highly correlated with response variable and with each other. Explanatory variables that are highly correlated with one another may cause multicollinearity problem.

Table (3) gives an eigenvalue analysis of the explanatory variables after they have been centerd and scaled. Notice that incremental percent is the percent this eigenvalue is of the total. Percents near zero indicate a multicollinearity problem. The condition number is the largest eigenvalue divided by each corresponding eigenvalue. Condition numbers more than 100 indicates multicollinearity problem. *(See[9])*.

**Conclusions:**

1_A new approach for estimating the ridge parameter was introduced by using the singular value decomposition technique.

2_In their development of ridge regression, Hoerl and Kennard focus attention on the eigenvalues of the information matrix x'x. A seriously non orthogonal problem is characterized by the fact that the smallest eigenvalue is very much smaller than unity. In our problem the smallest eigenvalue is 0.0093 this indicates that our data set suffer from a high degree of multicollinearity. 3_It should also be noted that the variance inflation factor (VIF) is an additional measure of multicollinearity. It is the reciprocal of (1-$r_i^2$) where $r_i^2$ is the square value of the multiple correlation coefficient between the explanatory variable $x_i$ and other explanatory variables. A VIF of 10 or more indicates a multicollinearity problem. *(See[10])*. In our problem the largest VIF value is 65.16 which is an additional indicator that our data set suffer from a high degree of multicollinearity.

4_ Since one of the objects of ridge regression is to reduce the standard error of the regression coefficients, it is of interest to see how much reduction has taken place. For our problem it should be noted from table (5) and table (6) that the standard errors for the ridge regression coefficients are less than the corresponding standard errors for least squares coefficients. Also we note that the standard errors for the ridge regression coefficients obtained by using the conventional mehod for estimating the ridge parameter are less than the corresponding standard errors for the ridge regression coefficients obtained by using the Bayesian method. From this comparison we conclude that the conventional mehod is performed better than the Bayesian method for this data set.

**References:**

1. Draper, N.R. and Smith, H. (1981),"Applied regression analysis" second edition, John Wiley and Sons, New York.

2.Goldstein,M. and Smith,A.F.M.(1974),"Ridge type estimator for regression analysis."Jornal Royal Statistical Society. B36,284_291.

3.Rao,C.R.(1973),"Linear Statistical Inference and Its Applications."second edition,John Wiley and Sons. New York.

4.Obenchain,R.L. (1978),"Good and optimal ridge estimators." Ann. Statist.6,1111_1121.

5.Lindley,D.V. and Smith,A.F.M.(1972),"Bayes estimators for the linear model."Jornal Royal Statistical Society. B34, 1_41.

6.Srivastava,M.I. (2002),"Methods Of Multivariate Statistics."Wiley, New York

7.Hogg,R.V. and Craig,A.T.(1978),"Introduction to mathematical statistics." Fourth edition,Macmillan pub.Co.New York.

8.Hoerl,A.E. and Kennard,R.W.(1970),"Ridge regression: Biased estimation for non orthogonal problems."Technometrics. 12, 55_67.

9.Chatterjee,S. and Price,P.(1977),"Regression analysis by example."John Wiley and Sons, New York.

10.Marquardit,D.W.(1970),"Generalized inverse,Ridge regression, and nonlinear estimation."Technometrics.12, 591_612.

11.الجبوري ، خالد ابراهيم سلمان.( 2004 ) : تخطيط الموارد البشرية ودورها في تنمية القطاع الصناعي في العراق للفترة 1970_1990. رسالة ماجستير مقدمة الى مجلس المعهد العالي للدراسات السياسية والدولية في الجامعة المستنصرية.

**Table(1)Effects of five explanatory variables x1,x2,….,x5**

**on the response variable y**

| y | *X1* | *X2* | *X3* | *X4* | *X5* |
|---|---|---|---|---|---|
| 1271 | 1516 | 473 | 24524 | 50096 | 14216 |
| 1440 | 1514 | 1923 | 26012 | 57133 | 16467 |
| 1355 | 2037 | 2490 | 30360 | 66795 | 18697 |
| 1509 | 2019 | 2734 | 29385 | 66263 | 18716 |
| 1254 | 2210 | 3322 | 36619 | 61718 | 20696 |
| 1220 | 2234 | 3563 | 37393 | 61028 | 23474 |
| 1546 | 2525 | 4205 | 40034 | 69344 | 25283 |
| 1916 | 2381 | 4293 | 10310 | 66181 | 26172 |
| 2381 | 2944 | 5278 | 49163 | 62039 | 28108 |
| 2585 | 3528 | 8529 | 57299 | 69339 | 30707 |
| 2810 | 4308 | 9661 | 60965 | 69671 | 32561 |
| 1440 | 4444 | 9840 | 56600 | 69047 | 33609 |
| 2493 | 4588 | 10830 | 55789 | 65441 | 32789 |
| 3285 | 4075 | 12968 | 44125 | 44175 | 52340 |
| 3062 | 4563 | 12875 | 46492 | 52476 | 47633 |
| 3403 | 4031 | 12163 | 46599 | 48322 | 63398 |
| 2875 | 4479 | 12712 | 49070 | 49672 | 62205 |
| 2861 | 4343 | 12610 | 49035 | 47330 | 47530 |
| 2596 | 4299 | 12615 | 48013 | 46160 | 46260 |
| 1710 | 4345 | 12630 | 40860 | 45123 | 4510 |
| 1777 | 4865 | 12955 | 49095 | 44925 | 44915 |

**Table(2) Correlation matrix section**

| | X1 | X2 | X3 | X4 | X5 | Y |
|---|---|---|---|---|---|---|
| X1 | 1.000000 | 0.967806 | 0.780634 | -0.380475 | 0.821688 | 0.749380 |
| X2 | 0.967806 | 1.000000 | 0.688120 | -0.541955 | 0.912351 | 0.790267 |
| X3 | 0.780634 | 0.688120 | 1.000000 | -0.028945 | 0.513239 | 0.631133 |
| X4 | -0.380475 | -0.541955 | -0.028945 | 1.000000 | -0.645230 | -0.295671 |
| X5 | 0.821688 | 0.912351 | 0.513239 | -0.645230 | 1.000000 | 0.806230 |
| y | 0.749380 | 0.790267 | 0.631133 | -0.295671 | 0.806230 | 1.000000 |

**Table(3) Eigen values of correlations**

| No. | eigenvalue | Incremental percent | Cumulative percent | Condition number |
|---|---|---|---|---|
| 1 | 3.623598 | 72.47 | 72.47 | 1.00 |
| 2 | 1.051811 | 21.04 | 93.51 | 3.45 |
| 3 | 0.211248 | 4.22 | 97.73 | 17.15 |
| 4 | 0.104043 | 2.08 | 99.81 | 34.83 |
| 5 | 0.009300 | 0.19 | 100.00 | 389.65 |

**Table(4)Eigenvectors of correlations**

| No. | eigenvalue | X1 | X2 | X3 | X4 | X5 |
|---|---|---|---|---|---|---|
| 1 | 3.623598 | -0.502625 | -0.518719 | -0.284390 | 0.304255 | -0.487824 |
| 2 | 1.051811 | -0.199033 | -0.007497 | -0.592083 | -0.751940 | 0.210802 |
| 3 | 0.211248 | 0.190854 | 0.222736 | -0.539120 | 0.569859 | 0.425416 |
| 4 | 0.104043 | 0.552851 | 0.292886 | -0.303237 | -0.109447 | -0.710381 |
| 5 | 0.009300 | 0.604718 | -0.771674 | -0.035429 | -0.074329 | 0.179037 |

**Table(5)Ridge vs. Least squares comparison section for k = 0.125260**

| Explanatory variables | Standardized Ridge coefficients | Standarded LS coefficients | Ridge VIF | LS VIF | Ridge standard error | LS standad error |
|---|---|---|---|---|---|---|
| X1 | 0.0126 | -0.9859 | 0.955 | 42.54 | 0.0884 | 0.4965 |
| X2 | 0.2329 | 1.1247 | 0.637 | 65.16 | 0.0184 | 0.1562 |
| X3 | 0.1695 | 0.2634 | 1.248 | 3.326 | 0.0093 | 0.0127 |
| X4 | 0.1653 | 0.4680 | 1.085 | 2.808 | 0.0114 | 0.0154 |
| X5 | 0.5359 | 0.7006 | 1.447 | 9.261 | 0.0085 | 0.0181 |

**Table(6)Ridge vs. Least squares comparison section for k = 0.494107**

| Explanatory variables | Standardized Ridge Coefficients | Standardized LS Coefficients | Ridge VIF | LS VIF | Ridge standard error | LS standard error |
|---|---|---|---|---|---|---|
| X1 | 0.1319 | -0.9859 | 0.1892 | 42.54 | 0.0447 | 0.4965 |
| X2 | 0.1970 | 1.1247 | 0.1254 | 65.16 | 0.0092 | 0.1562 |
| X3 | 0.1563 | 0.2634 | 0.3861 | 3.326 | 0.0059 | 0.0127 |
| X4 | 0.0452 | 0.4680 | 0.4101 | 2.808 | 0.0079 | 0.0154 |
| X5 | 0.3126 | 0.7006 | 0.2951 | 9.261 | 0.0044 | 0.0181 |

مجلة ابن الهيثم للعلوم الصرفة والتطبيقية المجلد23 (3) 2010

# توظيف اسلوب بيز في تحليل انحدار الحرف

***حازم منصور كوركيس***

**قسم الرياضيات ،كلية التربية-ابن الهيثم ،جامعة بغداد**

**الخلاصه**

في هذا البحث وظف اسلوب بيز في تحليل انحدار الحرف وتقدير معلمة الحرف على فرض توافر معلومات مسبقة عن معلمات انموذج الانحدار الخطي العام، وان الانموذج يعاني من مشكلة التعدد الخطي غير التام بدرجة كبيرة ، كما تم استخدام اسلوب جديد في ايجاد مقدر معلمة الحرف الذي اقترحه كل من Hoerl and Kennard في عام 1970 ومن خلال دراسة مثال عددي اجريت مقارنة لأفضلية اداء هذه المقدرات .