American Journal of Applied Sciences 8 (1): 97-102, 2011
ISSN 1546-9239

A Modification of the Ridge Type Regression Estimators

Moawad El-Fallah Abd El-Salam
Department of Statistics and Mathematics and Insurance,
Faculty of Commerce, Zagazig, University, Zagazig, Egypt

Abstract: Problem statement: Many regression estimators have been used to remedy multicollinearity problem. The ridge estimator has been the most popular one. However, the obtained estimate is biased. **Approach:** In this stuyd, we introduce an alternative shrinkage estimator, called modified unbiased ridge (MUR) estimator for coping with multicollinearity problem. This estimator is obtained from Unbiased Ridge Regression (URR) in the same way that Ordinary Ridge Regression (ORR) is obtained from Ordinary Least Squares (OLS). Properties of MUR estimator are derived. **Results:** The empirical study indicated that the MUR estimator is more efficient and more reliable than other estimators based on Matrix Mean Squared Error (MMSE).**Conclusion:** In order to solve the multicollinearity problem, the MUR estimator was recommended.

Key words: Multicollinearity, Ordinary Least Squares (OLS), Ordinary Ridge Regression (ORR), Unbiased Ridge Regression (URR), Modified Unbiased Ridge (MUR), Matrix Mean Squared Error (MMSE), Cumulative Density Function (CDF), ridge parameter, alternative shrinkage estimator, harmonic mean

INTRODUCTION

Consider the following linear regression model:

$$Y = X\beta + \epsilon \tag{1}$$

with the usual notation. The Ordinary Least Squares (OLS) estimator:

$$\hat{\beta}_{LS} = (X'X)^{-1}X'Y \tag{2}$$

follows $N(\beta, \sigma^2(X'X)^{-1})$. If X'X is singular or near singular, we say that there is multicollinearity in the data. As a consequence, the variances of elements of $\hat{\beta}_{LS}$ are inflated. Hence, alternative estimation methods have been proposed to eliminate inflation in the variances of $\hat{\beta}_{LS}$. Hoerl and Kennard (1970) proposed Ordinary Ridge Regression (ORR) as:

$$\begin{aligned}\hat{\beta}(k) &= [I - K(X'X + KI_p)^{-1}]\hat{\beta}_{LS} \\ &= (X'X + KI_p)^{-1}X'Y,\ K \geq 0\end{aligned} \tag{3}$$

Usually 0 < K < 1. This estimator is biased but reduces the variances of the regression coefficients. Subsequently, several other biased estimators of β have been proposed (Swindel, 1976; Sarkar, 1996; Batah and Gore, 2008; Batah *et al*., 2009; Arayesh and Hosseini, 2010; Asekunowo *et al*., 2010; Hirun and Sirisoponsilp, 2010; Rana *et al*., 2009). Swinded (1976) defined Modified Ridge Regression (MRR) estimator as follows:

$$\hat{\beta}(k, b) = (X'X + KI_p)^{-1}(X'Y + Kb), \quad K \geq 0 \tag{4}$$

where, b is a prior estimate of β. As K increases indefinitely, the MRR estimator approaches b. Crouse *et al*. (1995) defined the Unbiased Ridge Regression (URR) estimator as follows:

$$\hat{\beta}(k, j) = (X'X + KI_p)^{-1}(X'Y + Kj), \quad K \geq 0 \tag{5}$$

where, $J \sim N\left(\beta, \frac{\sigma^2}{K}I_p\right)$ for K>0. They also proposed the following estimator of the ridge parameter K:

$$\hat{K}_{CJH} = \begin{cases} \dfrac{P\hat{\sigma}^2}{(\hat{\beta}_{LS} - J)(\hat{\beta}_{LS} - J) - \hat{\sigma}^2 tr(X'X)^{-1}} & \text{if } (\hat{\beta}_{LS} - J)'(\hat{\beta}_{LS} - J) > \hat{\sigma}^2 tr(X'X)^{-1} \\ \dfrac{P\hat{\sigma}^2}{(\hat{\beta}_{LS} - J)'(\hat{\beta}_{LS} - J)} & \text{Otherwise} \end{cases}$$

where, $\hat{\sigma}^2 = \frac{(Y - X\hat{\beta}_{LS})'(Y - X\hat{\beta}_{LS})}{(n - P)}$ is an unbiased estimator of $\hat{\sigma}^2$. They further noted that $\hat{K}_{CJH}$ is a generalization of $\hat{K}_{HKB} = \frac{P\hat{\sigma}^2}{\hat{\beta}_{LS}\hat{\beta}_{LS}}$ of Hoerl *et al*. (1981). Consider spectral decomposition of X'X, namely X'X

= TAT', where TT' = T'T = 1. In this case, Eq. 1 can be written as:

$$Y = XTT'\beta + \in = ZL + \in \tag{6}$$

with Z = XT, L = T'β where Z'Z = T'X'XT = A= diag (λ_1 , λ_2 , ….., λ_P). The diagonal elements of A are the eigenvalues of X'X and T consists of corresponding the eigenvalues of X'X. Hence OLS, ORR and URR of L are written as $\hat{L}_{OLS} = A^{-1}Z'Y, \hat{L}_{(K)} = (A + KI_p)_{-1}Z'Y$ and $\hat{L}_{(K,J)} = (A + KI_p)^{-1}(Z'Y + KJ)$, respectively.

In this study, we introduce an alternative shrinkage estimator , called Modified Unbiased Ridge (MUR) estimator. This estimator is obtained from URR in the same way that ORR is obtained from OLS. It is observed that OLS is unbiased but has inflated variances under multicollnearity. Similarly, URR suffers from inflated variances while eliminating the bias. The construction of MUR is based on the logic that just as ORR avoids inflating the variances at the cost of bias, MUR would have similar properties. With pre-multiple the matrix $[I - K (X'X + KI_p)^{-1}]$ to reduce the inflated variances in OLS, so that we expect the same effect with URR. so that we expect the same effect with URR. This is our motivating the alternative modified estimator. In this respect, the main object of this paper is that the MUR estimator performs well under the conditions of multicollinearity. The properties of this alternative modified estimator are studied , and some conditions for this estimator to have smaller MMSE than ORR and URR are derived also. In addition, as the value of K must be specified for K in MUR in the same way as in ORR and URR, so three different ways for determining K are compared using simulated data.

MATERIALS AND METHODS

The proposed estimator: We propose the following estimator of β:

$$\begin{aligned}\hat{\beta}_J(k) &= [I - k(X'X + KI_p)^{-1}]\hat{\beta}(k,J) \\ &= [I - k(X'X + KI_p)^{-1}](X'X + KI_p)^{-1}(X'Y + KJ)\end{aligned} \tag{7}$$

where, J~ $N(\hat{\beta},(\frac{\sigma^2}{k}I_p))$ and k>0. This estimator is called Modified Unbiased Ridge Regression (MUR) because it is developed from URR. The MUR in model (6) becomes:

$$\hat{L}_J(k) = [I - k(A + KI_p)^{-1}]\hat{L}(k,J) \tag{8}$$

The MUR estimator has the following properties.

Bias:

$$\text{Bias } (\hat{\beta}_J(k)) = E(\hat{\beta}_J(k)) - \beta = -kS_k^{-1}\beta \tag{9}$$

where, S = X'X and $S_{k=}$ (S + KI).

Variance:

$$\begin{aligned}\text{Var } (\hat{\beta}_J(k)) &= E[(\hat{\beta}_J(k) - E(\hat{\beta}_J(k)))(\hat{\beta}_J(k))] \\ &= \sigma^2 WS_k^{-1}W'\end{aligned} \tag{10}$$

where, $W = [I = KS_k^{-1}]$.

Matrix Mean Squared Error (MMSE):

$$\begin{aligned}&MMSE(\hat{\beta}_J(k)) = Var(\hat{\beta}_J(k)) + [bias(\hat{\beta}_J(k))] \\ &[bias(\hat{\beta}_J(k))] \quad = \sigma^2 WS_k^{-1}W' + k^2 S_k^{-1}\beta\beta' S_k^{-1}\end{aligned} \tag{11}$$

Scalar Mean squared Error (SMSE):

$$\begin{aligned}SMSE(\hat{\beta}_J(k)) &= E[(\hat{\beta}_J(k) - \beta)'(\hat{\beta}_J(k) - \beta)] \\ &= tr(MMSE(\hat{\beta}_J(k)))\end{aligned}$$

where, tr denotes the trace. Then:

$$SMSE(\hat{L}_J(k)) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i^2}{(\lambda_i + k)^3} + k^2 \sum_{i=1}^{p} \frac{(\lambda_i + k)L_i^2}{(\lambda_i + k)^3} \tag{12}$$

where, $\{\lambda_i\}$ are eigenvalues of X'X.

$\hat{\beta}_J(k = 0) = \hat{\beta}_{LS} = (X'X)^{-1}X'Y$ is the OLS estimator:

$\lim_{k \to 0} \hat{\beta}_J(k) = \hat{\beta}_{LS}$

Comparison with other estimators: MUR is biased and it is therefore compared with other estimators in terms of MMSE. We obtain conditions for MUR to have smaller MMSE than another estimator.

Comparison with ORR: The MMSE of ORR is (Ozkale and Kaçiranlar, 2007):

$$MMSE(\hat{\beta}_J(k)) = \sigma^2 WS_k^{-1}W' + k^2 S_k^{-1}\beta\beta' S_k^{-1} \tag{13}$$

so that:

$$SMSE(\hat{L}(k)) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \sum_{i=1}^{p} \frac{L_i^2}{(\lambda_i + k)^2} \tag{14}$$

Consider:

$$\begin{aligned}\Delta &= \text{MMSE}(\hat{\beta}(k)) - \text{MMSE}(\hat{\beta}_J(k)) \\ &= \sigma^2 W(S^{-1} - S_k^{-1})W' = \sigma^2 H\end{aligned} \tag{15}$$

Since $S_k - S = KI_p$ is positive definite (p.d.), it is easy to show that $S^{-1} = S_k^{-1}$ is p.d. whenever $k > 0$. Hence we have the following result.

Result 1: MUR has smaller MMSE than ORR when $k > 0$.

Comparison with URR: The MMSE of the URR estimator is (Ozkale and Kaçiranlar, 2007):

$$\text{MMSE}(\hat{\beta}(k,J)) = \sigma^2 S_k^{-1} \tag{16}$$

and hence:

$$\text{SMSE}(\hat{\beta}(k,J)) = \text{tr}\ (\text{MMSE}(\hat{\beta}(k,J))) \tag{17}$$

Then:

$$\text{SMSE}(\hat{L}(k,J)) = \sigma^2 \sum_{i=1}^{p} \frac{1}{(\lambda_i + k)} \tag{18}$$

From (11):

$$\begin{aligned}\Delta &= \text{MMSE}(\hat{\beta}(k,J)) - \text{MMSE}(\hat{\beta}_J(k)) \\ &= \sigma^2[S^{-1} - WS_k^{-1}W'] - k^2 S_k^{-1}\beta\beta' S_k^{-1} \\ &= S_k^{-1}[k^2\sigma^2(\frac{2}{k}I_p - S_k^{-1}) = k^2\beta\beta']S_k^{-1}\end{aligned}$$

Now, Δ is non-negative definite (n.n.d.) (assuming $k > 0$) if and only if $\Phi = \frac{1}{k^2}S_k \Delta S_k$ is n.n.d. Further:

$$\Phi = \sigma^2(\frac{2}{k}I_p - S_k^{-1})\beta\beta' \tag{19}$$

Since the matrix $\frac{2}{k}I_p - S_k^{-1}$ is positive definite (Farebrother, 1976), Φ is n.n.d. if and only if:

$$\beta'[\frac{2}{k}I_p - S_k^{-1}]^{-1}\beta \le \sigma^2 \tag{20}$$

Hence we have the following result.

Result 2: MUR has smaller MMSE than URR if:

$$\beta'[\frac{2}{k}I_p - S_k^{-1}]^{-1}\beta \le \sigma^2$$

The condition of result (2) is verified by testing:

$$H_0 : \beta'[\frac{2}{k}I_p - S_k^{-1}]^{-1}\beta \le \sigma^2$$

Against:

$$H_1 : \beta'[\frac{2}{k}I_p - S_k^{-1}]^{-1}\beta > \sigma^2$$

Since $\Lambda\text{-}\Lambda^*$ (k) is positive semi definite, the condition in Result (2) becomes $\beta'T\Lambda^*(k)^{-1}T'\beta \le \sigma^2$ if $\beta'T\Lambda^{-1}T'\beta \le \sigma^2$. Under the assumption of normality:

$$\begin{aligned}&\sigma^{-1}\Lambda^*(k)^{-\frac{1}{2}}T'\hat{\beta}_J(k) \sim N(\sigma^{-1}\Lambda^*(k)^{-\frac{1}{2}} \\ &(I - k\Lambda_k^{-1})T'\hat{\beta}, \Lambda^*(k)^{-1}(I - k\Lambda_k^{-1})^2)\end{aligned}$$

and the test statistics:

$$F = \frac{\hat{\beta}_J(k)'T\Lambda^{-1}T'\hat{\beta}_J(k)/p}{\hat{\epsilon}'\hat{\epsilon}/n-p} \sim F\left(p, n-p, \frac{\beta'T\Lambda^{-1}T'\beta}{2\sigma^2}\right)$$

under H_0. The conclusion is that MUR has a smaller MMSE than URR if H_0 is accepted and hence Result (2) holds.

Optimal ridge parameter: Since the MMSE of MUR depends on the ridge parameter k, the choice of k is crucial for the performance of MUR. Hence we find conditions on the values of k for MUR to be better than other estimators in terms of SMSE.

Result 3: We have:

$$\text{SMSE}_i\,(\hat{L}_J(k)) < \text{SMSE}_i\,(\hat{L}(k,J)), \text{for } 0 < k_i < k_{i1}$$

$$\text{SMSE}_i\,(\hat{L}_J(k)) > \text{SMSE}_i\,(\hat{L}(k,J)), \text{for } k_{i1} < k_i < \infty$$

Where:

$$k_{i1} = \frac{(\sigma^2 - \lambda_i L_i^2)}{2L_i^2} + \left[\frac{(\sigma^2 - \lambda_i L_i^2)^2}{4L_i^4} + \frac{2\sigma^2\lambda_i}{L_i^2}\right]^{1/2} > 0 \tag{21}$$

Proof: Result (3) can be proved by showing that:

$$\begin{aligned}&(\lambda_i + k_i)^3[\text{SMSE}_i\,(\hat{L}_J(k)) - \text{SMSE}_i\,(\hat{L}(k,J))] = \\ &k_i[L_i^2 k_i^2 - (\sigma^2 - \lambda_i L_i^2)k_i - 2\lambda_i\sigma^2]\end{aligned}$$

which is obtained from (12) and (18). This completes the proof.

Next, we compare SMSE of $(\hat{L}_J(k))$ with that of OLS component-wise. Notice that the MUR estimator reduced to OLS when k = 0. The i-th component for SMSE of L of OLS is given by:

$$\text{SMSE}_i(\hat{L}_{LS}) = \frac{\sigma^2}{\lambda_i},\ i = 1, 2, \ldots\ldots, p \tag{22}$$

We state the following result.

Result 4: We have:

- If $\lambda_i L_i^2 - \sigma^2 \le 0$, then the:

$$\text{SMSE}_i(\hat{L}_J(k)) < \text{SMSE}_i(\hat{L}_{LS}), \text{for } 0 < k_i < \infty$$

- If $\lambda_i L_i^2 - \sigma^2 > 0$, then there exists a positive k_{i2}, such that:

$$\text{SMSE}_i(\hat{L}_J(k)) > \text{SMSE}_i(\hat{L}_{LS}), \text{for } 0 < k_i < k_{i2}$$

and:

$$\text{SMSE}_i(\hat{L}_J(k)) < \text{SMSE}_i(\hat{L}_{LS}), \text{for } k_{i2} < k_i < \infty$$

Where:

$$k_{i2} = \left[\frac{(\lambda_i^2 L_i^2 - 3\sigma^2\lambda)^2}{4(\lambda_i L_i^2 - \sigma^2)^2} + \frac{3\lambda_i^2\sigma^2}{(\lambda_i L_i^2 - \sigma^2)}\right]^{\frac{1}{2}} - \frac{(\lambda_i^2 L_i^2 - 3\sigma^2\lambda)}{2(\lambda_i L_i^2 - \sigma^2)} > 0 \tag{23}$$

Proof: Result (4) can be proved by showing that:

$$\lambda_i(\lambda_i + k_i)^3[\text{SMSE}_i(\hat{L}_J(k)) - \text{SMSE}_i(\hat{L}_{LS})] =$$
$$k_i[\lambda_i L_i^2 - \sigma^2)k_i^2] + \lambda_i^2 L_i^2 - 3\sigma^2\lambda_i)k_i - 3\lambda_i^2\sigma^2]$$

which is obtained from (12) and (22). This completes the proof.

Furthermore, differentiating $\text{SMSE}_i(\hat{L}_J(k))$ with respect to k_i and equating to zero, we have the following equation:

$$\frac{\partial \text{SMSE}_i(\hat{L}_J(k))}{\partial k} = \frac{2\lambda_i L_i^2 k_i^2 + 2\lambda_i^2 L_i^2 k_i - 3\sigma\lambda_i^2}{(\lambda_i + k_i)^4} 0$$

Thus, the optimal value of the ridge parameter k_i is:

$$k_{i(FG)} = \frac{\lambda_i}{2}[(1 - (\frac{6\sigma}{L_i^2}))^{\frac{1}{2}} - 1] \tag{24}$$

From (21), (23) and (24), it can be easily verified that $k_{i1} < k_{i(FG)} < k_{i2}$ if $\lambda_i L_i^2 - \sigma^2 > 0$. In case $k = k_1 = k_2 \ldots = k_p$, we can obtain k as the harmonic mean of $k_{i(FG)}$ in (24). It is given by:

$$k_{(FG)} = \frac{p\sigma^2}{\sum_{i=1}^{p}[L_i^2 / [(\frac{L_i^4\lambda_i^2}{4\sigma^4} + \frac{6L_i^2\lambda_i}{\sigma^2})^{\frac{1}{2}} - \frac{\lambda_i L_i^2}{2\sigma^2}]]} \tag{25}$$

Using an argument from Hoerl *et al*. (1981), it is reasonable to adopt the harmonic mean of the regression coefficients. Note that $k_{(FG)}$ in (25) depends on unknown parameters L and σ^2 and hence has to be estimated.

Estimating the ridge parameter k: In this section, we propose to construct MUR by using the operational ridge parameter proposed by Hoerl *et al*. (1981) and Crouse *et al*. (1995). First, since the harmonic mean of optimal ridge parameter values, (see (24)) depends on the unknown parameters L and σ^2, we use their OLS estimates. The operational ridge parameter in (25) is:

$$k_{(FG)} = \frac{p\sigma^2}{\sum_{i=1}^{p}[\hat{L}_i^2 / [(\frac{L_i^4\lambda_i^2}{4\hat{\sigma}^4} + \frac{6\hat{L}_i^2\lambda_i}{\hat{\sigma}^2})^{\frac{1}{2}} - \frac{\lambda_i \hat{L}_i^2}{2\hat{\sigma}^2}]]} \tag{26}$$

This is called the (FG) ridge parameter. Second, the HKB ridge parameter (Hoerl *et al*., 1981) is:

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{L'_{LS}\hat{L}_{LS}} \tag{27}$$

Third, CJH ridge parameter (Crouse *et al*., 1995) is:

$$\hat{k}_{CJH} = \begin{cases} \dfrac{p\hat{\sigma}^2}{(\hat{\beta}_{LS} - J)'(\hat{\beta}_{LS} - J) - \hat{\sigma}^2 tr(X'X)^{-1}} & \text{if } (\hat{\beta}_{LS} - J)' (\hat{\beta}_{LS} - J) > \hat{\sigma}^2 tr(W'X)^{-1} \\ \dfrac{p\hat{\sigma}^2}{(\hat{\beta}_{LS} - J)'((\hat{\beta}_{LS} - J))} & \text{Otherwise} \end{cases}$$

Using these three operational ridge parameters, we compare the following ten estimators:

- OLS
- ORR using the HKB ridge parameter (ORR (HKB))
- ORR using the CJH ridge parameter (ORR (CJH))

Table 1: Values of estimates and SMSE for $\hat{k}_{HKB} = 0.0133, \hat{k}_{CJH} = 0.436$ and $\hat{k}_{FG} = 0.0481$ where SMSE shows the SMSE for estimators

	β_1	β_2	β_3	β_4	β_5	SMSE
Population	9.02690	8.3384	3.0903	3.3411	11.3258	
OLS	7.95670	16.6563	2.6446	-5.9090	12.3692	144.6341
$\hat{\beta}(\hat{k}_{HKB})$	7.49966	12.5610	1.4810	0.0517	11.8267	111.7236
$\hat{\beta}(\hat{k}_{CJH})$	6.94390	10.2121	1.4999	3.4230	11.0095	157.6149
$\hat{\beta}(\hat{k}_{FG})$	6.89220	10.0442	1.5541	3.6480	10.9105	162.3889
$\hat{\beta}(\hat{k}_{HKB}, \bar{J})$	7.50580	12.6224	1.5332	-0.0062	11.8634	88.1827
$\hat{\beta}(\hat{k}_{CJH}, \bar{J})$	6.98120	10.3265	1.6252	3.3621	11.1187	52.3663
$\hat{\beta}(\hat{k}_{FG}, \bar{J})$	6.93420	10.1645	1.6888	3.5901	11.0296	49.6882
$\hat{\beta}_{\bar{J}}(\hat{k}_{HKB})$	7.14330	10.4899	1.1554	3.1145	11.4138	94.5319
$\hat{\beta}_{\bar{J}}(\hat{k}_{CJH})$	6.93420	10.1645	1.6888	3.5901	11.0296	147.5083
$\hat{\beta}_{\bar{J}}(\hat{k}_{FG})$	6.41870	8.4265	2.2479	5.9084	9.9834	152.8586

- ORR using the FG ridge parameter (ORR (FG))
- URR using the HKB ridge parameter (URR (HKB))
- URR using the CJH ridge parameter (URR (CJH))
- URR using the FG ridge parameter (URR (FG))
- MUR using the HKB ridge parameter (MUR (HKB))
- MUR using the CJH ridge parameter (MUR (CJH))
- MUR using the FG ridge parameter (MUR (FG))

RESULTS

We analyze the data generated by Hoerl and Kennard (1981). The data set is generated by taking a factor structure a real data set and choosing β_1 = 9.0269, β_2 = 8.3384, β_3 = 3.0903, β_4 = 3.3411 and β_5 = 11.3258 at random with constraint $\beta'\beta$ = 300 and a standard normal error $\in$ is added to form the observed response variable β_1, β_2, β_3, β_4 and β_5 are random with the constraint $\beta'\ \beta$ = 300 and normal error e has zero mean and $\sigma^2 = 1$. The resulting model is $Y=X\beta+\epsilon$ and ϵ is normally distributed as $N(0,\sigma^2 I)$.

The data was then used by Course *et al*. (1995) to compare SMSE performance of URR, ORR and OLS. Recently, Batah *et al*. (2009) used the same data to illustrate the comparisons among OLS and various ridge type estimators. We now use this data to illustrate the performance of the MUR estimator to the OLS, ORR and URR estimators to compare the MMSE performance of these estimators.

DISCUSSION

Table 1 shows the estimates and the SMSE values of these estimators. The eigenvalues of X'X matrix are 4.5792, 0.1940, 0.1549, 0.0584, 0.0138. the ratio of the largest to the smallest eigenvalue is 331.1251 which implies the existence of multicollinearity in the data set. The comparison between SMSE ($\hat{\beta}_{LS}$) and SMSE ($\hat{\beta}(\hat{k}_{HKB})$) show that the magnitude of shrinkage is not enough.

When biased and unbiased estimators are available, we prefer unbiased estimator. Crouse *et al*. (1995) suggested $\bar{J} = [\sum_{i=1}^{5}\hat{\beta}_{iLS}/5]1_{5\times1}$ as a realistic empirical prior information where 1 is the vector of ones. URR with $\hat{k}_{FG}$ leads to smaller SMSE than with $\hat{k}_{CJH}$ and $\hat{k}_{HKB}$ and correct the wrong sign. We thus find that $\hat{k}_{FG}$ is sufficient. MUR has smaller SMSE than ORR. Table 1 summarizes the performance of estimators for special values of k. we observe that MUR estimator with $\bar{J}$ = (6.7437, 6.7437, 6.7437, 6.7437, 6.7437) is not always better than other estimators in terms of having smaller SMSE. Also we can see that MUR is better than ORR for all $\hat{k}_{HKB}$, $\hat{k}_{CJH}$ and $\hat{k}_{FG}$ under the MMSE criterion, which is result (1).

The value of $\hat{\beta}'_{LS}[\frac{2}{k}I_p - S_k^{-1}]^{-1}\hat{\beta}_{LS}$ given in result (2) is obtained as 4.0791 for $\hat{k}_{HKB}$, 14.9195 for $\hat{k}_{CJH}$ and 16.6142 for $\hat{k}_{FG}$ which are not smaller than the OLS estimate of $\sigma^2 = 1.4281$. Therefore, URR estimator is better than the MUR estimator for $\hat{k}_{HKB}$, $\hat{k}_{CJH}$ and $\hat{k}_{FG}$ in terms of MMSE as in Table 1. The value of the F test in Result (2) is F_{Cul}=39.1003, the non-central F parameter value calculated is 392.888 with numerator degrees of freedom 5, denominator degrees of freedom 10 by using the Cumulative Density Function (CDF) Calculator for the Non-central-F Distribution (see website http://wwwdanielsoper.com/statcalc/calc06.aspx). Here, the non-central F_{CDF} is equal to 0.03118. Then H_0 is accepted and the condition in Result (2) holds. That is, MUR has smaller MMSE than URR.

CONCLUSION

In this study article we have introduced Modified Unbiased Ridge (MUR) estimator. Comparison of this estimator to that ORR and URR has been studied using the MMSE. Conditions for this estimator to have smaller MMSE than other estimators are established. The theortical results indicate that MUR is not always better than other estimators in terms of MMSE. MUR is best and depends on the unknown parameters β, σ^2 and also using the ridge parameter k. for suitable estimates of these parameters, MUR estimator might be considered as one of the good estimators using MMSE.

REFERENCES

Arayesh, B. and S.J. Hosseini, 2010. Regression analysis of effective factor on people participation in protecting, revitalizing, developing and using renewable natural resources in Ilam province from the view of users. Am. J. Agric. Bio. Sci., 5: 228-234. DOI: 10.3844/ajabssp.2010.228.234

Asekunowo, V.O., G. Olutunla and A.G. Daramola, 2010. Seeking an alternative modality to the management of Nigeria's fertilizer subsidy scheme-an empirical approach to the case study of ondo state (1976-1996). J. Soc. Sci., 6: 498-507. DOI: 10.3844/jssp.2010.498.507

Batah, F.S.B, M.R. Özkale and S.D. Gore, 2009. Combining unbiased ridge and principal component regression estimators. Communi. Statistics-Theory Methods, 38: 2201-2209. DOI: 10.1080/03610920802503396

Batah, F.S.M , S.D. Gore and M.R.Verma, 2008. Effect of jackknifing on various ridge type estimators. Model Assis. Stat. Appl., 1: 201-210.

Crouse, R.H., C. Jin and R.C. Hanumara, 1995. Unbiased ridge estimation with prior information and ridge trace. Communi. Statistics-Theory Methods, 24: 2341-2354. DOI: 10.1080/03610929508831620

Farebrother, R.W. ,1976. Further results on the mean square error of ridge regression. J. Royal Stat. Soci. B., 38: 248-250.

Hirun, W. and S. Sirisoponsilp, 2010. Analysis of interregional commodity flows. Am. J. Eng. Applied Sci., 3: 728-733. DOI: 10.3844/ajeassp.2010.728.733

Hoerl, A. and R. Kennard, 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12: 55-67.

Hoerl, A. and R. Kennard, 1981. Ridge regression-1980: Advances, algorithms and applications. Am. J. Math. Manage. Sci., 1: 5-83.

Rana, M.S., M. Habshah and A.H.M.R. Imon, 2009. A robust rescaled moment test for normality in regression. J. Math. Stat., 5: 54-62. DOI: 10.3844/jmssp.2009.54.62

Sarkar, N., 1996. Mean square error matrix comparison of some estimators in linear regressions with multicollinearity. Stat. Probabil. Lett., 30: 133-138. DOI: 10.1016/0167-7152(95)00211-1

Swindel, B.F., 1976. Good ridge estimators based on prior information. Communi. Statistics-Theory Methods, 5: 1065-1075. DOI: 10.1080/03610927608827423