Surveys in Mathematics and its Applications

ISSN 1842-6298 (electronic), 1843-7265 (print)
Volume **4** (2009), 99 – 109

RIDGE REGRESSION ESTIMATOR: COMBINING UNBIASED AND ORDINARY RIDGE REGRESSION METHODS OF ESTIMATION

Feras Sh. M. Batah and Sharad Damodar Gore

Abstract. Statistical literature has several methods for coping with multicollinearity. This paper introduces a new shrinkage estimator, called modified unbiased ridge (MUR). 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 are derived. Results on its matrix mean squared error (MMSE) are obtained. MUR is compared with ORR and URR in terms of MMSE. These results are illustrated with an example based on data generated by Hoerl and Kennard [8].

1 Introduction

Consider the linear regression model

$$Y \;=\; X\beta+\epsilon, \tag{1.1}$$

with the usual notation. The ordinary least squares (OLS) estimator

$$\hat{\beta}_{LS} \;=\; (X'X)^{-1}X'Y, \tag{1.2}$$

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 techniques have been proposed to eliminate inflation in the variances of $\hat{\beta}_{LS}$. Hoerl and Kennard [7] proposed Ordinary Ridge Regression (ORR) as

$$\begin{aligned}\hat{\beta}(k) &= [I-k(X'X+kI_p)^{-1}]\hat{\beta}_{LS}\\ &= (X'X+kI_p)^{-1}X'Y, \quad k\geq 0.\end{aligned} \tag{1.3}$$

2000 Mathematics Subject Classification: 62J05; 62J07.

Keywords: Multicollinearity; Ordinary Least Squares Estimator; Ordinary Ridge Regression Estimator; Unbiased Ridge Estimator

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 (see, Swindel [12]; Sarkar [11]; Batah and Gore [1]; Batah et al. [2]; Batah et al. [3]; Batah et al. [4]). Swindel [12] defined modified ridge regression (MRR) estimator as follows:

$$\hat{\beta}(k,b) \;=\; (X'X+kI_p)^{-1}(X'Y+kb), \quad k\geq 0, \tag{1.4}$$

where b is a prior estimate of β. As k increases indefinitely, the MRR estimator approaches b. Crouse et al. [5] defined the unbiased ridge regression (URR) estimator as follows:

$$\hat{\beta}(k,J) \;=\; (X'X+kI_p)^{-1}(X'Y+kJ), \quad k\geq 0, \tag{1.5}$$

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

$$\hat{k}_{CJH} = \begin{cases} \frac{p\hat{\sigma}^2}{(\hat{\beta}_{LS}-J)'(\hat{\beta}_{LS}-J)-\hat{\sigma}^2 tr(X'X)^{-1}} & \text{if } \left(\hat{\beta}_{LS}-J\right)'\left(\hat{\beta}_{LS}-J\right) > \hat{\sigma}^2 tr\left(X'X\right)^{-1} \\ \frac{p\hat{\sigma}^2}{(\hat{\beta}_{LS}-J)'(\hat{\beta}_{LS}-J)} & \text{otherwise,} \end{cases}$$

where $\hat{\sigma}^2 = \frac{(Y-X\hat{\beta}_{LS})'(Y-X\hat{\beta}_{LS})}{(n-p)}$ is an unbiased estimator of σ^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. [8]. Consider Spectral decomposition of $X'X$, namely $X'X = T\Lambda T'$, where $TT' = T'T = I$. Model (1.1) can be written as

$$\begin{aligned} Y &= XTT'\beta+\epsilon \\ &= Z\gamma+\epsilon, \end{aligned} \tag{1.6}$$

with $Z = XT$, $\gamma = T'\beta$ where $Z'Z = T'X'XT = \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_p)$. The diagonal elements of Λ are the eigenvalues of $X'X$ and T consists of corresponding the eigenvectors of $X'X$. Hence OLS, ORR and URR of γ are written as $\hat{\gamma}_{OLS} = \Lambda^{-1}Z'Y$, $\hat{\gamma}(k) = (\Lambda+kI_p)^{-1}Z'Y$, and $\hat{\gamma}(k,J) = (\Lambda+kI_p)^{-1}(Z'Y+kJ)$, respectively.

This paper introduces a new shrinkage estimator, called modified unbiased ridge (MUR). 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 multicollinearity. 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. This is our motivating the new estimator. In this paper, it is indeed observed that MUR performs well under the conditions of multicollinearity. The properties of this new estimator are studied in Section 2. Some conditions for the new estimator to have smaller MMSE than ORR and URR are derived in Section 3. The value of k must be specified for k in MUR in the same way as in ORR and URR. Three different ways of determining k are compared using simulated data. Optimal ridge parameters are considered in Section 4. Section 5 contains some estimators of the ridge parameter k. Results of the paper are illustrated with Hoerl and Kennard data in Section 6. The paper ends with concluding remarks in section 7.

2 The proposed Estimator

We propose the following estimator of β

$$\begin{aligned}\hat{\beta}_J(k) &= [I - k(X'X + kI_p)^{-1}]\hat{\beta}(k, J) \\ &= [I - k(X'X + kI_p)^{-1}](X'X + kI_p)^{-1}(X'Y + kJ), \qquad (2.1)\end{aligned}$$

where $J \sim N(\beta, \frac{\sigma^2}{k} I_p)$ and $k > 0$. This estimator is called modified unbiased ridge regression (MUR) because it is developed from URR. The MUR in model (1.6) becomes

$$\hat{\gamma}_J(k) = [I - k(\Lambda + kI_p)^{-1}]\hat{\gamma}(k, J). \qquad (2.2)$$

The MUR estimator has the following properties.

1. **Bias**

$$\begin{aligned}Bias(\hat{\beta}_J(k)) &= E(\hat{\beta}_J(k)) - \beta \\ &= -kS_k^{-1}\beta, \qquad (2.3)\end{aligned}$$

   where $S = X'X$, and $S_k = (S + kI)$.

2. **Variance**

$$\begin{aligned}Var(\hat{\beta}_J(k)) &= E[(\hat{\beta}_J(k) - E(\hat{\beta}_J(k)))(\hat{\beta}_J(k) - E(\hat{\beta}_J(k)))' \\ &= \sigma^2 W S_k^{-1} W', \qquad (2.4)\end{aligned}$$

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

3. **Matrix Mean Squared Error (MMSE)**

$$\begin{aligned}MMSE(\hat{\beta}_J(k)) &= Var(\hat{\beta}_J(k)) + [bias(\hat{\beta}_J(k))][bias(\hat{\beta}_J(k))]' \\ &= \sigma^2 W S_k^{-1} W' + k^2 S_k^{-1}\beta\beta' S_k^{-1}. \qquad (2.5)\end{aligned}$$

4. **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

$$SMSE(\hat{\gamma}_J(k)) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i^2}{(\lambda_i+k)^3} + k^2 \sum_{i=1}^{p} \frac{(\lambda_i+k)\gamma_i^2}{(\lambda_i+k)^3}. \tag{2.6}$$

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

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

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

3 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.

3.1 Comparison with ORR

The MMSE of ORR is (Özkale and Kaçiranlar [10])

$$MMSE(\hat{\beta}(k)) = \sigma^2 W S^{-1} W' + k^2 S_k^{-1}\beta\beta' S_k^{-1}, \tag{3.1}$$

so that

$$SMSE(\hat{\gamma}(k)) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+k)^2} + k^2 \sum_{i=1}^{p} \frac{\gamma_i^2}{(\lambda_i+k)^2}. \tag{3.2}$$

Consider

$$\begin{aligned} \Delta &= MMSE(\hat{\beta}(k)) - MMSE(\hat{\beta}_J(k)) \\ &= \sigma^2 W(S^{-1} - S_k^{-1})W' \\ &= \sigma^2 H. \end{aligned} \tag{3.3}$$

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* $k > 0$.

3.2 Comparison with URR

The MMSE of the URR estimator is (Özkale and Kaçiranlar [10])

$$MMSE(\hat{\beta}(k,J)) \quad = \quad \sigma^2 S_k^{-1}, \tag{3.4}$$

and hence

$$SMSE(\hat{\beta}(k,J)) \quad = \quad tr(MMSE(\hat{\beta}(k,J))). \tag{3.5}$$

Then

$$SMSE(\hat{\gamma}(k,J)) = \sigma^2 \sum_{i=1}^{p} \frac{1}{(\lambda_i + k)}. \tag{3.6}$$

From (2.5),

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

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,

$$\Phi = \sigma^2(\frac{2}{k}I_p - S_k^{-1}) - \beta\beta'. \tag{3.7}$$

Since the matrix $\frac{2}{k}I_p - S_k^{-1}$ is positive definite (Farebrother [6]), Φ is n.n.d. if and only if

$$\beta'[\frac{2}{k}I_p - S_k^{-1}]^{-1}\beta \leq \sigma^2. \tag{3.8}$$

Hence we have the following result.

Result 2. *MUR has smaller MMSE than URR if*

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

The condition of Result (2) is verified by testing

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

against

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

Since $\Lambda - \Lambda^*(k)$ 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 $\sigma^{-1}\Lambda^*(k)^{-1/2}T'\hat{\beta}_J(k) \sim N(\sigma^{-1}\Lambda^*(k)^{-1/2}(I - k\Lambda_k^{-1})T'\beta, \Lambda^*(k)^{-1}(I - k\Lambda_k^{-1})^2)$, and the test statistics

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

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

4 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*

1. $SMSE_i(\hat{\gamma}_J(k)) < SMSE_i(\hat{\gamma}(k, J))$, *for* $0 < k_i < k_{i1}$.
2. $SMSE_i(\hat{\gamma}_J(k)) > SMSE_i(\hat{\gamma}(k, J))$, *for* $k_{i1} < k_i < \infty$

where

$$k_{i1} = \frac{(\sigma^2 - \lambda_i\gamma_i^2)}{2\gamma_i^2} + [\frac{(\sigma^2 - \lambda_i\gamma_i^2)^2}{4\gamma_i^4} + \frac{2\sigma^2\lambda_i}{\gamma_i^2}]^{\frac{1}{2}} > 0. \tag{4.1}$$

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

$$(\lambda_i + k_i)^3[SMSE_i(\hat{\gamma}_J(k)) - SMSE_i(\hat{\gamma}(k, J))] = k_i[\gamma_i^2 k_i^2 - (\sigma^2 - \lambda_i\gamma_i^2)k_i - 2\lambda_i\sigma^2],$$

which is obtained from (2.6) and (3.6). This completes the proof. □

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

$$SMSE_i(\hat{\gamma}_{LS}) = \frac{\sigma^2}{\lambda_i}, \quad i = 1, 2, \cdots, p. \tag{4.2}$$

We state the following result.

Result 4. *We have*

1. *If* $\lambda_i\gamma_i^2 - \sigma^2 \le 0$, *then the*

$$SMSE_i(\hat{\gamma}_J(k)) < SMSE_i(\hat{\gamma}_{LS}), \quad for \;\; 0 < k_i < \infty.$$

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

$$SMSE_i(\hat{\gamma}_J(k)) > SMSE_i(\hat{\gamma}_{LS}), \quad for \;\; 0 < k_i < k_{i2},$$

and

$$SMSE_i(\hat{\gamma}_J(k)) < SMSE_i(\hat{\gamma}_{LS}), \quad for \;\; k_{i2} < k_i < \infty,$$

where

$$k_{i2} \;\; = \;\; [\frac{(\lambda_i^2\gamma_i^2 - 3\sigma^2\lambda)^2}{4(\lambda_i\gamma_i^2 - \sigma^2)^2} + \frac{3\lambda_i^2\sigma^2}{(\lambda_i\gamma_i^2 - \sigma^2)}]^{\frac{1}{2}} - \frac{(\lambda_i^2\gamma_i^2 - 3\sigma^2\lambda)}{2(\lambda_i\gamma_i^2 - \sigma^2)} > 0. \tag{4.3}$$

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

$$\begin{aligned}\lambda_i(\lambda_i + k_i)^3[SMSE_i(\hat{\gamma}_J(k)) - SMSE_i(\hat{\gamma}_{LS})] \;\; = \;\; & k_i[(\lambda_i\gamma_i^2 - \sigma^2)k_i^2 \\ & + (\lambda_i^2\gamma_i^2 - 3\sigma^2\lambda_i)k_i - 3\lambda_i^2\sigma^2],\end{aligned}$$

which is obtained from (2.6) and (4.2). This completes the proof. □

Furthermore, differentiating $SMSE_i(\hat{\gamma}_J(k))$ with respect to k_i and equating to zero, we have the following equation

$$\frac{\partial SMSE_i(\hat{\gamma}_J(k))}{\partial k} \;\; = \;\; \frac{2\lambda_i\gamma_i^2k_i^2 + 2\lambda_i^2\gamma_i^2k_i - 3\sigma\lambda_i^2}{(\lambda_i + k_i)^4} = 0.$$

Thus, the optimal value of the ridge parameter k_i is

$$k_{i(FG)} \;\; = \;\; \frac{\lambda_i}{2}[(1 - (\frac{6\sigma}{\gamma_i^2}))^{\frac{1}{2}} - 1]. \tag{4.4}$$

From (4.1), (4.3), and (4.4), it can be easily verified that $k_{i1} < k_{i(FG)} < k_{i2}$ if $\lambda_i\gamma_i^2 - \sigma^2 > 0$. In case $k = k_1 = k_2 = \ldots = k_p$,we can obtain k as the harmonic mean of $k_{i(FG)}$ in (4.4).It is given by

$$k_{(FG)} \;\; = \;\; \frac{p\sigma^2}{\sum_{i=1}^{p}[\gamma_i^2/[(\frac{\gamma_i^4\lambda_i^2}{4\sigma^4} + \frac{6\gamma_i^2\lambda_i}{\sigma^2})^{1/2} - \frac{\lambda_i\gamma_i^2}{2\sigma^2}]]}. \tag{4.5}$$

Using an argument from Hoerl et al.[8], it is reasonable to adopt the harmonic mean of the regression coefficients. Note that $k_{(FG)}$ in (4.5) depends on unknown parameters γ and σ^2, and hence has to be estimated.

5 Estimating the Ridge Parameter k

In this section, we propose to construct MUR by using the operational ridge parameter proposed by Hoerl et al.[8] and Crouse et al. [5]. First, since the harmonic mean of optimal ridge parameter values, see (4.4) depends on the unknown parameters γ and σ^2, we use their OLS estimates. The operational ridge parameter in (4.5) is

$$\hat{k}_{FG} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}[\hat{\gamma}_i^2/[(\frac{\hat{\gamma}_i^4\lambda_i^2}{4\hat{\sigma}^4} + \frac{6\hat{\gamma}_i^2\lambda_i}{\hat{\sigma}^2})^{1/2} - \frac{\lambda_i\hat{\gamma}_i^2}{2\hat{\sigma}^2}]]}. \tag{5.1}$$

This is called the (FG) ridge parameter. Second, the HKB ridge parameter (Hoerl et al. [8]) is

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\gamma}_{LS}'\hat{\gamma}_{LS}}. \tag{5.2}$$

Third, CJH ridge parameter (Crouse et al. [5]) is

$$\hat{k}_{CJH} = \begin{cases} \frac{p\hat{\sigma}^2}{(\hat{\beta}_{LS}-J)'(\hat{\beta}_{LS}-J)-\hat{\sigma}^2 tr(X'X)^{-1}} & \text{if } \left(\hat{\beta}_{LS}-J\right)'\left(\hat{\beta}_{LS}-J\right) > \hat{\sigma}^2 tr\,(X'X)^{-1} \\ \frac{p\hat{\sigma}^2}{(\hat{\beta}_{LS}-J)'(\hat{\beta}_{LS}-J)} & \text{otherwise.} \end{cases}$$

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

1.OLS.
2.ORR using the HKB ridge parameter (ORR (HKB)).
3.ORR using the CJH ridge parameter (ORR (CJH)).
4.ORR using the FG ridge parameter (ORR (FG)).
5.URR using the HKB ridge parameter (URR (HKB)).
6.URR using the CJH ridge parameter (URR (CJH)).
7.URR using the FG ridge parameter (URR (FG)).
8.MUR using the HKB ridge parameter (MUR (HKB)).
9.MUR using the CJH ridge parameter (MUR (CJH)).
10.MUR using the FG ridge parameter (MUR (FG)).

6 Illustrative Example

We analyze the data generated by Hoerl and Kennard [9]. The data set is generated by taking a factor structure from a real data set, and choosing $\beta_1 = 9.0269$, $\beta_2 = 8.3384$, $\beta_3 = 3.0903$, $\beta_4 = 3.3411$, and $\beta_5 = 11.3258$ at random with constraint $\beta'\beta = 300$ and a standard normal error ϵ is added to form the observed response variable.β_1, β_2, β_3, β_4, β_5 are random with the constraint $\beta'\beta = 300$ and normal

	β_1	β_2	β_3	β_4	β_5	$SMSE$
Population	9.0269	8.3384	3.0903	3.3411	11.3258	
OLS	7.9567	16.6563	2.6446	-5.9090	12.3692	144.6341
$\hat{\beta}(\hat{k}_{HKB})$	7.4966	12.5610	1.4810	0.0517	11.8267	111.7236
$\hat{\beta}(\hat{k}_{CJH})$	6.9439	10.2121	1.4999	3.4230	11.0095	157.6149
$\hat{\beta}(\hat{k}_{FG})$	6.8922	10.0442	1.5541	3.6480	10.9105	162.3889
$\hat{\beta}(\hat{k}_{HKB}, \bar{J})$	7.5058	12.6224	1.5332	-0.0062	11.8634	88.1927
$\hat{\beta}(\hat{k}_{CJH}, \bar{J})$	6.9812	10.3265	1.6252	3.3621	11.1187	52.3663
$\hat{\beta}(\hat{k}_{FG}, \bar{J})$	6.9342	10.1645	1.6888	3.5901	11.0296	49.6882
$\hat{\beta}_{\bar{J}}(\hat{k}_{HKB})$	7.1433	10.4899	1.1554	3.1145	11.4138	94.5319
$\hat{\beta}_{\bar{J}}(\hat{k}_{CJH})$	6.9342	10.1645	1.6888	3.5901	11.0296	147.5083
$\hat{\beta}_{\bar{J}}(\hat{k}_{FG})$	6.4187	8.4265	2.2479	5.9084	9.9834	152.8586

Table 1:
Values of estimates and SMSE for $\hat{k}_{HKB} = 0.0133$, $\hat{k}_{CJH} = 0.0436$ and $\hat{k}_{FG} = 0.0481$ where $SMSE$ shows the SMSE for estimators

error e has zero mean and $\sigma^2 = 1$. The resulting model is $Y = X\beta + \epsilon$, and ϵ is normally distributed as $N\left(0, \sigma^2 I\right)$.

The data was then used by Crouse et al.[5] to compare the SMSE performance of URR, ORR and OLS. Recently, Batah et al. [2] and Batah et al. [4] 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. 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. [5] suggested $\bar{J} = [\sum_{i=1}^{5} \hat{\beta}_{iLS}/5]1_{5\times 1}$ 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 corrects 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 Noncentral-F Distribution (see website http://www.danielsoper.com/statcalc/calc06.aspx). Here, the noncentral F_{CDF} is equal to 0.031118. Then H_0 is accepted and the condition in Result (2) holds. That is, MUR has smaller MMSE than URR.

7 Conclusion

In this article we have introduced modified unbiased ridge (MUR). 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 theorical 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.

Acknowledgement. The authors would like to acknowledge the editor and the referee for their valuable comments, which improved the paper substantially.

References

[1] F. Batah and S. Gore, *Improving Precision for Jackknifed Ridge Type Estimation*, Far East Journal of Theoretical Statistics **24** (2008), 157 – 174. MR2474325. Zbl pre05495192.

[2] F. Batah, S. Gore and M. Verma, *Effect of Jackknifing on Various Ridge Type Estimators*, Model Assisted Statistics and Applications **3** (2008a), 121 – 130.

[3] F. Batah, T. Ramanathan and S. Gore, *The Efficiency of Modified Jackknife and Ridge Type Regression Estimators: A Comparison*, Surveys in Mathematics and its Applications **3** (2008b), 111 – 122. MR2438671.

[4] F. Batah, S. Gore and M. Özkale, *Combining Unbiased Ridge and Principal Component Regression Estimators*, Communication in Statistics - Theory and Methods **38** (2009), 2201 – 2209.

[5] R. Crouse, C. Jin and R. Hanumara, *Unbiased Ridge Estimation With Prior Information and Ridge Trace*, Communication in Statistics - Theory and Methods **24** (1995), 2341 – 2354. MR1350665(96i:62073). Zbl 0937.62616.

[6] R. Farebrother, *Further Results on the Mean Squared Error of the Ridge Regression*, Journal of the Royal Statistical Society B. **38** (1976), 248-250.

[7] A. Hoerl and R. Kennard, *Ridge Regression: Biased Estimation for Non orthogonal Problems*, Technometrics **12** (1970), 55 – 67. Zbl 0202.17205.

[8] A. Hoerl, R. Kennard and K. Baldwin, *Ridge Regression: Some Simulations*, Communication in Statistics - Theory and Methods **4** (1975), 105–123.

[9] A. Hoerl and R. Kennard, *Ridge Regression: 1980 Advances, Algorithms, and Applications*, Amer.J. Mathemat. Manage. Sci. **1** (1981), 5 – 83. MR0611894(82d:62117). Zbl 0536.62054.

[10] M. Özkale and S. Kaçiranlar, *Comparisons of the Unbiased Ridge Estimation to the Other Estimations*, Communication in Statistics - Theory and Methods **36** (2007), 707 – 723. MR2391894(2009a:62292). Zbl pre05150072.

[11] N. Sarkar, *Mean Squared Error Matrix Comparison of Some Estimators in Linear Regression With Multicollinearity*, Statistics and Probabilty Letters **21** (1996), 1987 – 2000.

[12] B. Swindel, *Good Ridge Estimators Based on Prior Information*, Communication in Statistics - Theory and Methods **11** (1976), 1065 – 1075. MR0440806(55:13675). Zbl 0342.62035.

Feras Shaker Mahmood Batah
Department of Statistics,
University of Pune, India.
Department of Mathematics,
University of Alanber, Iraq.
e-mail: ferashaker2001@yahoo.com

Sharad Damodar Gore
Department of Statistics,
University of Pune, India.
e-mail: sdgore@stats.unipune.ernet.in.