

11-1-2005

Applications of Some Improved Estimators in Linear Regression

B. M. Golam Kibria

Florida International University, kibriag@fiu.edu

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Kibria, B. M. Golam (2005) "Applications of Some Improved Estimators in Linear Regression," *Journal of Modern Applied Statistical Methods*: Vol. 5 : Iss. 2 , Article 11.

DOI: 10.22237/jmasm/1162354200

Applications of Some Improved Estimators in Linear Regression

B. M. Golam Kibria
Department of Statistics
Florida International University

The problem of estimation of the regression coefficients under multicollinearity situation for the restricted linear model is discussed. Some improved estimators are considered, including the unrestricted ridge regression estimator (URRE), restricted ridge regression estimator (RRRE), shrinkage restricted ridge regression estimator (SRRRE), preliminary test ridge regression estimator (PTRRE), and restricted Liu estimator (RLIUE). They were compared based on the sampling variance-covariance criterion. The RRRE dominates other ridge estimators when the restriction does or does not hold. A numerical example was provided. The RRRE performed equivalently or better than the RLIUE in the sense of having smaller sampling variance.

Key words: Bias; non-Central F, preliminary test, restricted Liu estimator, ridge regression, shrinkage estimation, variance matrix.

Introduction

Multiple linear regression model plays an important role in statistical inference and is used extensively in business, environmental, industrial, and social sciences. In linear regression model, one usually assumes that the explanatory variables are independent. However, in practice, there may be strong or near to strong linear relationships among the explanatory variables. In that case, the independence assumptions are no longer valid, which causes the problem of multicollinearity. In the presence of multicollinearity, it is difficult to estimate the unique effects of individual variables in the

regression equation. Moreover, the estimated regression coefficient will have large sampling variance which affects both inference and prediction. In the literature, there are various methods existing to solve this problem. Among them, ridge regression is the most popular one which has much usefulness in real life. To describe the ridge regression, the following multiple linear regression model was considered:

$$y = X\beta + e, \quad (1)$$

where y is a random vector of length n , β is a $p \times 1$ vector of fixed but unknown regression coefficients, X is a $n \times p$ known design matrix of rank p and e is a $n \times 1$ vector random variable, which is distributed as multivariate normal with mean vector 0 and covariance matrix $\sigma^2 \times I_n$, I_n is an identity matrix of order n , and $\sigma^2 > 0$ is unknown error variance. Of primary interest is an estimation of the regression coefficient β , when it is *a priori* suspected that β may be restricted to the subspace

$$H\beta = h, \quad (2)$$

B. M. Golam Kibria is Assistant Professor. He is a fellow of the Royal Statistical Society and the recipient of the 2005 Chapter Service Recognition Award. He is the Associate Editor and member of the editorial board of three international statistical journals. He has authored/co-authored about 40 articles in statistical journals.

where H is a $q \times p$ known matrix of rank $q (< p)$ and h is an $q \times 1$ vector of known constants. The usual least squares estimate (ULSE) or the maximum likelihood estimate (MLE) of β is given by $\tilde{\beta} = C^{-1}X'y$, which will heavily depend on the characteristics of the matrix $C = X'X$. If the C matrix is ill conditioned ($\det(X'X \approx 0)$), the LSE are sensitive to a number of errors, for example, some of the regression coefficients may be statistically insignificant with wrong sign and meaningful statistical inference become difficult for the practitioners. To solve this problem, Hoerl and Kennard (1970) suggested use of $C(k) = C + kI_p$, ($k \geq 0$), instead of C for estimating β . The resulting estimators are given as:

$$\tilde{\beta}(k) = (C + kI_p)^{-1} X'y, \quad (3)$$

which are known as unrestricted ridge regression estimators (URRE). The constant, $k > 0$ is known as shrinkage or biasing or ridge parameter. Though these estimators result in biasing for $k \geq 0$, they yield minimum mean square error (MSE) compared to the unrestricted least squares estimator (URLSE) (for examples, see Hoerl & Kennard, 1970, Saleh & Kibria, 1993). The objective of this article is to compare the performance of some ridge regression estimators under the sampling variance-covariance criterion. The findings will be illustrated with a numerical example.

Proposed Estimators, Biases and Variances

In this section, some ridge type estimators are considered for β , when $H\beta = h$ is *a priori* suspected in the case of multicollinearity. Biases and variance-covariance matrices of the estimators were also provided.

Unrestricted Ridge Regression Estimator (URRE):

The URRE in (3), can be re-written as follows:

$$\tilde{\beta}(k) = W\tilde{\beta}, \quad W = [I_p + kC^{-1}]^{-1}; \quad k \geq 0, \quad (4)$$

where $\tilde{\beta}$ is the unrestricted least square estimator (URLSE) of β . The bias and the variance matrix of URRE are as follows:

$$\begin{aligned} \text{Bias} = B_1(k) &= E(\tilde{\beta}(k)) - \beta = -kC^{-1}(k)\beta; \\ C(k) &= [C + kI_p]; \\ V(\tilde{\beta}(k)) = V_1(k) &= \sigma^2(WC^{-1}W'). \end{aligned} \quad (5)$$

Restricted Ridge Regression Estimator (RRRE):
Sarkar (1992) proposed the following RRRE:

$$\hat{\beta}(k) = W\hat{\beta}, \quad (6)$$

where $\hat{\beta} = \tilde{\beta} - C^{-1}H'(HC^{-1}H')^{-1}(H\tilde{\beta} - h)$ is the restricted least squares estimator (RLSE) of β . The bias and the variance matrix of RRRE are as follows:

$$\begin{aligned} B_2(k) &= -W\eta - kC^{-1}(k)\beta; \\ V_2(k) &= \sigma^2(WC^{-1}W') - \sigma^2(WAW'), \end{aligned} \quad (7)$$

where

$$A = C^{-1}H'(HC^{-1}H')^{-1}HC^{-1},$$

$$\eta = C^{-1}H'(HC^{-1}H')^{-1}\delta$$

and

$$\delta = H\beta - h.$$

Shrinkage Restricted Ridge Regression Estimator (SRRRE):

Haq and Kibria (1996) proposed the following SRRRE,

$$\hat{\beta}^{SR}(k) = W\hat{\beta}^{SR} \tag{8}$$

where

$$\hat{\beta}^{SR} = d\tilde{\beta} + (1-d)\hat{\beta}, \quad 0 \leq d \leq 1,$$

is the shrinkage restricted least squares estimator (SRLSE) of β and the coefficient d is known as the coefficient of distrust and may completely be determined by the user, depending on the degree of belief about the null hypothesis. For $d = 1$, $\tilde{\beta}$ is obtained and for $d = 0$, $\hat{\beta}$ is obtained. Therefore, $\hat{\beta}^{SR}(k)$ is a convex combination of $\tilde{\beta}$ and $\hat{\beta}$. The bias and variance matrix of SRRRE are given as follows:

$$\begin{aligned} B_3(k, d) &= -(1-d)W\eta - kC^{-1}(k)\beta; \\ V_3(k, d) &= \sigma^2(WC^{-1}W) - (1-d^2)\sigma^2(WAW). \end{aligned} \tag{9}$$

Preliminary Test Ridge Regression Estimator (PTRRE):

Saleh and Kibria (1993), first introduced the preliminary test ridge regression estimator,

$$\hat{\beta}^{PT}(k) = W\hat{\beta}^{PT} = \begin{cases} W\tilde{\beta} & \text{if } H_0 \text{ is false} \\ W\hat{\beta} & \text{if } H_0 \text{ is true,} \end{cases} \tag{10}$$

where $\hat{\beta}^{PT} = \tilde{\beta}I(Ln \geq F_\alpha) + \hat{\beta}I(Ln < F_\alpha)$ is the preliminary test least squares estimator (PTLSE) of β , which is introduced by Bancraft (1944) and $I(X)$ is the indicator function of X . Here Ln is the test-statistic for testing the null-hypothesis $H_0 : H\beta = h$ against $H_1 : H\beta \neq h$ and is given by

$$Ln = \frac{(H\tilde{\beta}_n - h)'(HC^{-1}H')^{-1}(H\tilde{\beta}_n - h)}{q\tilde{\sigma}_e^2}, \tag{11}$$

where

$$\tilde{\sigma}_e^2 = (n - p)^{-1}(Y - X\tilde{\beta}_n)'(Y - X\tilde{\beta}_n)$$

is the unrestricted least squares estimator of σ^2 . The corresponding restricted least squares estimator of σ^2 is given by

$$\hat{\sigma}_e^2 = (n - p + q)^{-1}(Y - X\hat{\beta}_n)'(Y - X\hat{\beta}_n).$$

Under H_0 , the test-statistic Ln follows a central F -distribution with $(q, n - p)$ degrees of freedom and F_α is the upper α -level critical value from this distribution. However, when H_0 does not hold, Giles (1991) had shown that the probability density function of Ln is non-central F -distribution with non-centrality parameter $\frac{\Delta}{2}$, where

$$\Delta = \frac{(H\beta - h)'(HC^{-1}H')^{-1}(H\beta - h)}{\sigma^2} = \frac{\eta' C \eta}{\sigma^2}. \tag{12}$$

is called the departure parameter. For the power of the F -test, readers are referred to Sutradhar (1988) and Giles (1991) among others. Preliminary test and ridge regression approaches have been studied by Bancroft (1944, 1964), Bock and Yancey (1973), Giles (1991), Han and Bancroft (1968), Judge and Bock (1978), Kibria and Saleh (1993), Kibria (2003), Kibria (2004a,b), Kibria and Saleh (2004, 2005a, 2005b, 2006), Obenchain et al. (1975) and Saleh and Han (1990) to mention a few. The bias and variance matrix of the PTRRE are as follows:

$$\begin{aligned} B_4(k, \Delta, \alpha) &= -W\eta G_{q+2, n-p}(I_\alpha^*; \Delta) \\ &\quad - kC^{-1}(k)\beta; \\ V_4(k, \Delta, \alpha) &= \sigma^2(WC^{-1}W) \\ &\quad - \sigma^2(WAW)G_{q+2, n-p}(I_\alpha^*; \Delta) \\ &\quad + (W\eta\eta'W)G(\Delta, \alpha), \end{aligned} \tag{13}$$

where

$$G(\Delta, \alpha) = \begin{bmatrix} 2G_{q+2, n-p}(I^*_\alpha; \Delta) \\ -G_{q+4, n-p}(I^{**}_\alpha; \Delta) \\ -G^2_{q+2, n-p}(I^*_\alpha; \Delta) \end{bmatrix} \quad (14)$$

and $G_{m,n}(\cdot; \Delta)$ is the cumulative density function (*cdf*) of a non-central F-distribution with (m, n) degrees of freedom and the non-centrality parameter Δ . Also $l^*_\alpha = \frac{q}{q+2}F_\alpha$ and $l^{**}_\alpha = \frac{q}{q+4}F_\alpha$. When $\alpha = 1$, then $F_\alpha = 0$ and $G_{m,n}(\cdot; \Delta) = 0$, consequently, the H_0 will be rejected and thus PTRRE becomes URRE. On the other hand, when $\alpha = 0$, then $F_\alpha = \infty$ and $G_{m,n}(\cdot; \Delta) = 1$, consequently, the H_0 will not be rejected and thus PTRRE becomes RRRE.

Restricted Liu Estimator (RLIUE):

Kaciranlar et al. (1999) proposed the following restricted Liu estimator (RLIUE):

$$\hat{\beta}^{RL}(b) = F_b \hat{\beta}, \quad (15)$$

based on the corresponding unrestricted Liu estimator (ULIUE) (Liu 1993), $\hat{\beta}^{UL}(b) = F_b \tilde{\beta}$, where $F_b = (C + I)^{-1}(C + bI)$ and b is the constant to be determined. Note that for $b = 1$, the ULIUE and RLIUE become corresponding URLSE and RLSE respectively. It also noted that for $k = 1$ and $b = 0$ both RRRE and RLIUE become the same. The bias and the variance matrix of the RLIUE are as follows:

$$\begin{aligned} B_5(b, \Delta) &= (b-1)(C + I)^{-1} \beta + F_b \eta; \\ V_5(b) &= \sigma^2 \{F_b (C^{-1} - A) F'_b\}. \end{aligned} \quad (16)$$

In the following section, the performance of the proposed ridge type

estimators will be proposed under the restriction (2) and using the sampling variance criterion. This is because the value of the test statistics depends on the sampling variance than the mean squares error (MSE). Because the detailed analysis about ULIUE and RLIUE are available in Kaciranlar et al. (1999), for convenience, the comparison of ridge regression estimators with that of RLIUE would be provided only by an example.

Performance of the estimators under H_0

Comparison between $\tilde{\beta}$ and $\tilde{\beta}(k)$

The difference between the sampling variance of URLSE and URRE is

$$\begin{aligned} Var(\tilde{\beta}) - Var(\tilde{\beta}(k)) &= \sigma^2 [C^{-1} - C^{-1}(k)W'] \\ &= \sigma^2 W [W^{-1}C^{-1}W'^{-1} - C^{-1}] W' \\ &= \sigma^2 W [(I_p + kC^{-1})C^{-1}(I_p + kC^{-1})' - C^{-1}] W' \\ &= \sigma^2 W [2kC^{-2} + k^2C^{-3}] W' \\ &= k\sigma^2 C^{-1}(k)EC^{-1}(k)', \end{aligned} \quad (17)$$

where $C(k) = [C + kI]$ and $E = (2I + kC^{-1})$ are the positive definite (p.d.) matrices. Thus, the difference in (17) is non-negative definite (n.n.d) for $k > 0$ and therefore, $\tilde{\beta}(k)$ is superior to $\tilde{\beta}$ with respect to sampling variance. It further notes that the variance of $\tilde{\beta}$ will be equal to that of $\tilde{\beta}(k)$ iff $k = 0$.

Comparison between $\hat{\beta}$ and $\hat{\beta}(k)$

The difference between the sampling variance of RLSE and RRRE is

$$\begin{aligned} V_2(0) - V_2(k) &= \sigma^2 [(C^{-1} - A) - W(C^{-1} - A)W'] \\ &= k\sigma^2 C^{-1}(k)[E - F_0]C^{-1}(k)', \end{aligned} \quad (18)$$

where $F_0 = (CA + AC + kA)$ is a symmetric matrix. Then, following Anderson (1984)

(Theorem A.2.2, p.589), there existed a nonsingular matrix P such that $P'EP = I$ and $P'F_0P = \Lambda$, where Λ is a diagonal matrix and its elements are roots of the polynomial equation $|E - \lambda F_0| = 0$. Because

$(P'EP - P'F_0P) = P'(E - F_0)P = I - \Lambda$ is a non-negative definite matrix, then $(1 - \lambda_i) \geq 0$. That means, $(P'EP - P'F_0P)$ to be the n.n.d. matrix, it is necessary that $\lambda_{max}(E^{-1}F_0) \leq 1$.

Let $\lambda_{max}(E^{-1}F_0) \leq 1$. Because E is positive definite matrix and F_0 is symmetric matrix, then $\lambda_p \leq \frac{x'F_0x}{x'Ex} \leq \lambda_1$, where

$\lambda_1(E^{-1}F_0) \geq \lambda_2(E^{-1}F_0) \leq \dots \geq \lambda_p(E^{-1}F_0)$ are the roots of $|E - \lambda F_0| = 0$ (see Anderson, 1984, Theorem A.2.4, p. 590). Then, it is obvious that $V_2(0) - V_2(k)$ is non-negative definite for $k \geq 0$ if and only if $\lambda_{max}(E^{-1}F_0) \leq 1$. Therefore, sampling variance of RRRE is always less than or equal to that of RLSE if and only if $\lambda_{max}(E^{-1}F_0) \leq 1$.

If the sampling variance of URRE is compared with that of RRRE, then

$$V_1(k) - V_2(k) = \sigma^2(WAW'); k \geq 0,$$

which is a positive semidefinite matrix. Therefore, RRRE is always superior to that of URRE, when the sampling variance criterion is considered.

Comparison between $\hat{\beta}^{SR}$ and $\hat{\beta}^{SR}(k)$

The difference between sampling variance of SRLSE and SRRRE is

$$\begin{aligned} &V_3(0,d) - V_3(k,d) \\ &= \sigma^2[(C^{-1} - (1-d^2)A) - W(C^{-1} - (1-d^2)A)W'] \\ &= k\sigma^2 C^{-1}(k)[E - F_1]C^{-1}(k)', \end{aligned} \tag{19}$$

where $E = 2I + kC^{-1}$ and $F_1 = (1-d^2)[CA + AC + kA]$. Then, it is

obvious that $V_3(0,d) - V_3(k,d)$ is non-negative definite for $k \geq 0$ and $0 \leq d \leq 1$ if and only if $\lambda_{max}(E^{-1}F_1) \leq 1$.

If the sampling variance of SRRRE is compared with that of URRE, then

$$V_1(k) - V_3(k,d) = \sigma^2(1-d^2)WC^{-1}W',$$

which is always positive semi definite matrix for all $k \geq 0$ and $0 \leq d \leq 1$. Thus, SRRRE is superior to URRE. If the sampling variance of SRRRE is compared with that of RRRE, then

$$V_2(k) - V_3(k,d) = -\sigma^2 d^2 WC^{-1}W',$$

which is negative semidefinite. Therefore, RRRE is superior to SRRRE for all d and k when the restriction holds.

Comparison between $\hat{\beta}^{PT}$ and $\hat{\beta}^{PT}(k)$

The difference between the sampling variance of PTLSE and PTRRE is

$$\begin{aligned} &V_4(0,0,\alpha) - V_4(k,0,\alpha) \\ &= k\sigma^2 C^{-1}(k)[E - F_2]C^{-1}(k)', \end{aligned} \tag{20}$$

where $F_2 = [CA + AC + kA]G_{q+2,n-p}(I_\alpha^*; 0)$.

Then, the sampling variance of $\hat{\beta}^{PT}(k)$ is always less than or equal to that of $\hat{\beta}^{PT}$ if and only if $\lambda_{max}(E^{-1}F_2) \leq 1$ for all α . If the sampling variance of PTRRE is compared with that of URRE, then

$$\begin{aligned} &V_1(k) - V_4(k,0,\alpha) \\ &= \sigma^2(WAW')G_{q+2,n-p}(I_\alpha^*; 0), \end{aligned}$$

which is always positive semi definite matrix for all $k \geq 0$. Thus, PTRRE is superior to URRE. If PTRRE is compared with that of RRRE, then

$$\begin{aligned} &V_2(k) - V_4(k,0,\alpha) \\ &= \sigma^2(WAW')(G_{q+2,n-p}(I_\alpha^*; 0) - 1), \end{aligned}$$

which is negative semidefinite for all $\alpha \in (0,1)$. Therefore, RRRE is superior to PTRRE for all α and k when the restriction hold. If the variance of PTRRE is compared with that of SRRRE, then

$$V_3(k, d) - V_4(k, 0, \alpha) = \sigma^2(WAW') \begin{Bmatrix} G_{q+2, n-p}(I_\alpha^*; 0) \\ -(1-d^2) \end{Bmatrix}$$

Thus, the PTRRE will dominates SRRRE if

$$F_\alpha > \left(\frac{q+2}{q} \right) G_{q+2, n-p}^{-1}(1-d^2), \quad (21)$$

otherwise, SRRRE will dominate PTRRE.

It is observed from equations (5), (7), (9), (13) and (16), that under the restrictions, all the proposed ridge estimators are biased and the amount of biases are same for the ridge regression estimators. For $k=0$, all proposed ridge estimators become corresponding usual least squares estimator and under restrictions they are unbiased. It is also observed that the RRRE dominates all other estimators under the variance criterion and when the restriction holds. In the following section, the performance of the estimators will be compared when the restriction does not hold.

Performance of the estimators under H_a

Comparison between $\hat{\beta}^{PT}$ and $\hat{\beta}^{PT}(k)$

The difference between the sampling variance of PTLSE and PTRRE is

$$V_4(0, \Delta, \alpha) - V_4(k, \Delta, \alpha) = k\sigma^2 C^{-1}(k)[E - F_3]C^{-1}(k)', \quad (22)$$

where $E = (2I + kC^{-1})$ is positive definite matrix and

$$F_3 = (CA + AC + kA)G_{q+2, n-p}(I_\alpha^*; \Delta) - \frac{1}{\sigma^2}(\eta\eta'C + C\eta\eta' + k\eta\eta')G(\Delta, \alpha)$$

is a symmetric matrix and $G(\Delta, \alpha)$ is greater than 0 for all α and Δ . Then, following Anderson (1974, Theorems A.2.2 and A.2.4), the difference in (22) will be non-negative definite if and only if $\lambda_{max}(E^{-1}F_3) \leq 1$. Therefore, under the alternative hypothesis, sampling variance of $\hat{\beta}^{PT}(k)$ ($k \geq 0$) is always less than or equal to that of $\hat{\beta}^{PT}$ if and only if $\lambda_{max}(E^{-1}F_3) \leq 1$.

Comparison between PTRRE and URRE

$$V_1(k) - V_4(k, \Delta, \alpha) = \sigma^2(WAW')G_{q+2, n-p}(I_\alpha^*; \Delta) - W\eta\eta'W'G(\alpha, \Delta), \quad (23)$$

where $G(\alpha, \Delta)$ is defined in (14).

Let P be the orthogonal matrix with eigen vectors of C so that

$$P'CP = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p).$$

Because (WAW') and $(W\eta\eta'W')$, are the $p \times p$ symmetric matrices, the following may be written

$$WAW' = P[\Lambda + kI]^{-1}\Lambda^*\Lambda[\Lambda + kI]^{-1}P'$$

where $P'AP = A^*$. Then, the i^{th} diagonal element of the matrix WAW' is $\frac{a_i^* \lambda_i^2}{(\lambda_i + k)^2}$.

Similarly, the i^{th} diagonal element of the matrix $W\eta\eta'W'$ is $\frac{\eta_i^{*2} \lambda_i^2}{(\lambda_i + k)^2}$, where η_i^* is the i^{th} element if the vector $\eta^* = \eta'P$. Then, the difference in (23) is positive semidefinite if,

$$a_i^* \geq \left[\frac{2G_{q+2, n-p}(I_\alpha^*; \Delta) - G_{q+4, n-p}(I_\alpha^{**}; \Delta) - G_{q+2, n-p}^2(I_\alpha^*; \Delta)}{G_{q+2, n-p}(\alpha, \Delta)} \right] \eta_i^{*2}. \quad (24)$$

Thus, PTRRE will dominate URRE when (24) holds.

Special Cases on Δ

Consider $k = 0$ in (23), and then following Judge and Bock (1978, p.115-117), the sampling variance of PTLSE is obtained and is smaller than that of the URLSE when,

$$\Delta \leq \frac{G_{q+2,n-p}(I_{\alpha}^*; \Delta)}{2[2G_{q+2,n-p}(I_{\alpha}^*; \Delta) - G_{q+4,n-p}(I_{\alpha}^{**}; \Delta) - G_{q+2,n-p}^2(I_{\alpha}^*; \Delta)]}$$

This result coincides with that of Judge and Bock (1978, p.116).

Comparison between PTRRE and RRRE

$$V_2(k) - V_5(k, d, \Delta, \alpha) = - \left\{ \frac{\sigma^2 (WAW') [1 - G_{q+2,n-p}(I_{\alpha}^*; \Delta)]}{+ [(W\eta\eta'W)G(\alpha, \Delta)]} \right\}$$

It is observed that the above difference is always negative semi-definite for all $\alpha \in (0, 1)$, $\Delta \in (0, \infty)$ and $d \in (0, 1)$. Therefore, the RRRE dominates PTRRE, SRRRE and URRE whether the restriction does or does not holds.

Comparison between PTRRE and SRRRE

$$V_3(k, d) - V_4(k, \Delta, \alpha) = - \left\{ \frac{\sigma^2 (WAW') \{1 - d^2 - G_{q+2,n-p}(I_{\alpha}^*; \Delta)\}}{+ W\eta\eta'W'G(\alpha, \Delta)} \right\} \tag{25}$$

The difference in (25) will be negative semidefinite and therefore, the SRRRE will dominates PTRRE if

$$d \leq \sqrt{1 - G_{q+2,n-p}(I_{\alpha}^*; \Delta)},$$

otherwise, PTRRE will dominates SRRRE.

Numerical Example

An example is considered that demonstrates the performance of the proposed estimators. Accordingly, the dataset on Portland cement, originally due to Woods et al. (1932), was analyzed. This dataset has widely been analyzed by Hald (1952, p.635-652), Gorman and Torman (1966, p.735), Nomura (1988, p.735), Piepel and Redgate (1998) and recently Kaciranlar et al. (1999) among many researchers. This data came from an experimental investigation of the heat evolved during the setting and hardening of the Portland cements of varied composition as a function of the percentage of four compounds in the clinkers from which the cement was produced. Woods et al. (1932) considered the following four compounds: Tricalcium aluminate ($3CaO.Al_2O_3$), Tricalcium silicate ($3CaO.SiO_2$), Tetracalcium aluminoferrite ($4CaO.Al_2O_3.Fe_2O_3$), and β -dicalcium silicate ($2CaO.SiO_2$). The heat evolved after 180 days of curing is measured in calories per gram of cement. For details about the dataset, please see Woods et al. (1932). Consider the following linear model to analyze the cement dataset.

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + e_i; \quad i = 1, 2, \dots, 13, \tag{26}$$

where y_i denotes the i^{th} measurement for the heat evolved in calories per gram of cement. x_1 =Tricalcium aluminate ($3CaO.Al_2O_3$), x_2 =Tricalcium silicate ($3CaO.SiO_2$), x_3 =Tetracalcium aluminoferrite ($4CaO.Al_2O_3.Fe_2O_3$), x_4 = β -dicalcium silicate ($2CaO.SiO_2$) and e is a random error which is distributed as $N(0, \sigma^2)$. The model (26) is known as *inhomogeneous* linear regression model and will be used to analyze the cement dataset to compare the performance of the proposed estimators and compared them with that of Kaciranlar et al. (1999). The least square estimates and their corresponding

estimated standard error (SE)s, together with associated t_8 -statistic and two-sided P-values are given in Table 1.

Table 1: The Regression Analysis

Coefs	SE	t -value	P-value
62.4054	70.0710	0.8906	0.3991
1.5511	0.7450	2.0821	0.0709
0.5102	0.7239	0.7048	0.5010
0.1019	0.7550	0.1350	0.8960
-0.1441	0.7092	-0.2031	0.8441

Note that S-plus software was used to analyze the dataset. It was found that $R^2 = 0.9824$, which implies that 98.24% of

Moreover, the singular values for the design matrix are: 211.36746660, 77.23614495, 28.45965700, 10.26736008, and 0.03490017 and the condition number of X is $211.36746660/0.03490017 = 6056.3744$. So, the X matrix is severely ill-conditioned (Kaciranlar, 1999), which probably make the regression coefficients insignificant and therefore, it is hard to make a valid inference or prediction using usual least squares method. To make a valid inference and to compare the performance of the proposed estimators, the cement dataset would be the most appropriate one to analyze.

Table 2: The correlation matrix

	x_1	x_2	x_3	x_4	y
x_1	1.0000	0.2286	-0.8241	-0.2454	0.7307
x_2	0.2286	1.0000	-0.1392	-0.9730	0.8163
x_3	-0.8241	-0.1392	1.0000	0.0295	-0.5347
x_4	-0.2454	-0.9730	0.0295	1.0000	-0.8213
y	0.7307	0.8163	-0.5347	-0.8213	1.0000

the total variation has been explained by the regressors, $F_{0.05,4,8} = 111.48$, which is highly significant. However, none of the regression coefficients is statistically significant at 5% level. This could possibly be due to the high correlation among the explanatory variables, which can be observed from the correlation matrix provided in Table 2.

To compare the performance of the proposed ridge estimators with that of ULIUE and RLIUE, consider the following parametric restriction: $\beta_1 - \beta_2 + \beta_3 = 0$ as that of Kaciranlar (1999), which can be expressed as

$$\begin{aligned} H_0 & : H\beta = 0 \\ H_a & : H\beta \neq 0, \end{aligned} \tag{27}$$

where $H = (0,1,-1,1,0)$ is an 1×5 known matrix. In the notation, $q = 1$ and $p = 5$. To test the hypotheses in (27), use (11) and observe that $L_n = 1.92$. Because $F_{0.05,1,8} = 5.32$, one would not reject H_0 at 5% significance level. It is noted that the proposed ridge regression estimators depend on the value of k . Many different techniques for estimating k have been proposed or suggested by different researchers. Among them, Hoerl et al (1975), Lawless and Wang (1976), are notable.

For detailed about the estimation technique of k , see Kibria (2003) and references therein. The estimated values of k by using, Hoerl et al. (1975, p.107), and Lawless and Wang (1976, p.311) are 0.0077 and 0.00025 respectively. The estimated σ^2 by using URLSE and RLSE are 5.982955, 6.595198 respectively. The sampling variance of the proposed estimators depends on the value of σ^2 .

Therefore, 5.982955 was used to estimate the sampling variance of URLSE, URRE and URLIUE and 6.595198 for ridge regression estimators and RLIUE. The estimated standard error (SE) has been provided for all proposed estimators along with unrestricted Liu Estimator (ULIUE) and restricted Liu estimator (RLIUE) in Table 3. It may be observed from Table 3

that all the proposed ridge regression estimators dominate corresponding usual least squares estimators in the sense of having smaller sampling variances. Assume $d = 0.5$; the natural choice for the analysis.

It was also found from Table 3 that for any value of $k = 0, 0.00025$ and 0.0077 ,

the RRRE produces the smallest sampling variances compared to other ridge estimators. Specifically, for $k = 0.0077$, all ridge estimators dominate the restricted Liu estimator. Thus, using ridge regression technique, one would be able to improve the estimators. From Table 3, it was also observed that the sampling variance of RLIUE is smaller than that of ULIUE, which is the consequences of Theorem 3.1.1. (Kaciranlar 1999, p.448). Finally, using $k = 0.0077$, the data was analyzed and URRRE along with corresponding least squares estimators (in the parentheses) are provided in Table 4. Table 5 provides the regression analysis for RRRE and the restricted Liu estimator (in the parentheses).

It may be seen from Table 4, that under the OLS method, the regression coefficients are insignificant, yet under the ridge regression approach, all the regression coefficients are highly significant. It was also observed that the sign of the estimated regression coefficients has been changed, which is expected for highly correlated data (see Hoerl & Kennard, 1970). When the RLIUE was compared with that of RRRE, it was seen from Table 5 that all the regression coefficients, with the exception of β_2 under RLIUE, is not significant.

Next, the following parametric restrictions were considered: $\beta_1 - \beta_2 + \beta_3 = 0$, $\beta_2 - \beta_3 + \beta_4 = 0$, $\beta_1 - \beta_2 - \beta_4 = 0$, which yield the following hypotheses,

$$\begin{aligned} H_0 & : H\beta = 0 \\ H_a & : H\beta \neq 0, \end{aligned} \tag{28}$$

Table 3: Estimated standard error of the proposed estimators

URLSE	RLSE	SRLSE	PTLSE	ULIUE
$\alpha = 0.05 \quad d = 0.5 \quad k=0$				
70.0710	31.6114	44.4629	43.5948	55.7471
0.7450	0.1789	0.4025	0.3899	0.6032
0.7239	0.3674	0.4817	0.4733	0.5766
0.7550	0.2074	0.4183	0.4062	0.6083
0.7092	0.3464	0.4648	0.4561	0.5647
URRE	RRRE	SRRRE	PTRRE	RLIUE
$\alpha = 0.05 \quad d=0.5 \quad k=0.00025$				
58.1381	26.2285	36.8912	36.1710	31.6215
0.6269	0.1265	0.3317	0.3209	0.1721
0.6008	0.3114	0.4037	0.3975	0.3690
0.6325	0.1549	0.3435	0.3332	0.2031
0.5891	0.2933	0.3886	0.3821	0.3481
URRE	RRRE	SRRRE	PTRRE	RLIUE
$\alpha = 0.05 \quad d=0.5 \quad k=0.0077$				
9.5703	4.3195	6.0738	5.9553	31.6215
0.2121	0.1304	0.1549	0.1517	0.1721
0.1095	0.0894	0.0949	0.0949	0.3690
0.1897	0.1049	0.1342	0.1304	0.2031
0.1049	0.0775	0.0837	0.0837	0.3481

Table 4: Ridge regression (Usual Least Squares) analysis without restriction

Coefficients	SE	t-value	P-value
8.5642 (62.4054)	9.5703 (70.0710)	0.8949 (0.8906)	0.3970 (0.3991)
2.1048 (1.5511)	0.2121 (0.7450)	9.9238 (2.0820)	0.0000 (0.0708)
1.0651 (0.5102)	0.1095 (0.7239)	9.7268 (0.7047)	0.0000 (0.5009)
0.6683 (0.1019)	0.1897 (0.7550)	3.5231 (0.1350)	0.0078 (0.8959)
0.3998 (-0.1441)	0.1049 (0.7092)	3.8115 (-0.2031)	0.0052 (0.8441)

Table 5: Ridge regression (Restricted Liu) analysis under the restriction

Coefficients	SE	t-value	P-value
20.4006 (142.0379)	4.3195 (31.6215)	3.3528 (4.4918)	0.0100 (0.0020)
1.8724 (0.6207)	0.1304 (0.1721)	15.2503 (3.6065)	0.0000 (0.0069)
0.9716 (-0.2819)	0.0894 (0.3690)	11.3907 (-0.7640)	0.0000 (0.4668)
0.4504 (-0.8298)	0.1049 (0.2031)	5.3322 (-4.0857)	0.0007 (0.0035)
0.2985 (-0.9302)	0.0775 (0.3481)	4.5051 (-2.6721)	0.0020 (0.0283)

where H matrix is defined as

$$\begin{vmatrix} 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 1 & -1 & 0 & -1 \end{vmatrix}$$

To test the null hypothesis in (28), use (11) and observe that $L_n = 133.522$. The upper percentage point of the non-central F distribution with 3 and 8 degrees of freedom and non-centrality parameter 10.78 is 16, so, the null hypothesis is rejected at the 5% significance level. Now, the analysis will be carried out under the alternative hypothesis. The unrestricted and restricted estimated values of σ^2 are found to be 5.982955 and 222.221 respectively. It is noticed that the estimation of σ^2 in a restriction model very much depends of the restriction on the parameters. Using (12),

the non-centrality parameter Δ is estimated by replacing β by $\tilde{\beta}$ and σ^2 by $\tilde{\sigma}^2$ as 10.78.

The estimated SE for the proposed estimators along with ULIUE and RLIUE are presented in Tables 6 and 7, where σ^2 is estimated by $\tilde{\sigma}^2$ and $\hat{\sigma}^2$ respectively. It was found that the estimated SE in Table 7 is 6 times bigger than the SE in Table 6, as $\hat{\sigma}^2$ is bigger than the $\tilde{\sigma}^2$. From both Tables 6 and 7, it was observed that the proposed ridge regression estimators have smaller sampling variances compared to usual least squares estimator for any values of k . It was also noticed that the RRRE performed better than the URRRE, SRRRE and PTRRE in the sense of having smaller as well as stable sampling variance. For the restriction (27), the PTRRE performs better than URRRE and SRRRE.

Table 6: Estimated SE of the proposed estimators using σ^2 estimated by $\tilde{\sigma}^2$

URLSE	RLSE	SRLSE	PTLSE	URLIE
$\alpha = 0.05 \quad d = 0.5 \quad k=0$				
70.0710	1.1688	35.0501	32.9698	70.0464
0.7450	0.0316	0.3742	1.0232	0.7445
0.7239	0.0000	0.3619	0.3795	0.7235
0.7550	0.0316	0.3782	0.2811	0.7545
0.7092	0.0316	0.3564	0.2608	0.7088
URRE	RRRE	SRRRE	PTRRE	RLIUE
$\alpha = 0.05 \quad d=0.5 \quad k=0.00025$				
58.1381	0.9701	29.0812	27.3548	1.1632
0.6269	0.0316	0.3146	0.9762	0.0370
0.6008	0.0000	0.3000	0.3225	0.0001
0.6325	0.0447	0.3178	0.2569	0.0371
0.5891	0.0316	0.2966	0.2121	0.0370
URRE	RRRE	SRRRE	PTRRE	RLIUE
$\alpha = 0.05 \quad d=0.5 \quad k=0.0077$				
9.5703	0.1612	4.7872	4.5010	1.1632
0.2121	0.0316	0.1095	0.8000	0.0370
0.1095	0.0000	0.0548	0.0949	0.0001
0.1897	0.0447	0.1049	0.2881	0.0371
0.1049	0.0316	0.0548	0.1378	0.0370

Table 7: Estimated SE of the proposed estimators when σ^2 is estimated $\hat{\sigma}^2$

URRE	RRRE	SRRRE	PTRRE	URLIE
$\alpha = 0.05 \quad d = 0.5$				
427.0441	7.1237	213.6112	152.6839	426.8942
4.5389	0.2258	2.2779	1.8218	4.5375
4.4111	0.0000	2.2054	1.5818	4.4095
4.5996	0.2258	2.3080	1.6462	4.5980
4.3212	0.2258	2.1696	1.5479	4.3198
URRE	RRRE	SRRRE	PTRRE	RLIUE
$\alpha = 0.05 \quad d=0.5 \quad k=0.00025$				
354.3197	5.9108	177.2338	126.6822	7.0890
3.8196	0.2168	1.9189	1.5837	0.2255
3.6636	0.0000	1.8319	1.3153	0.0004
3.8555	0.2366	1.9386	1.3889	0.2261
3.5882	0.2168	1.8039	1.2888	0.2255
URRE	RRRE	SRRRE	PTRRE	RLIUE
$\alpha = 0.05 \quad d=0.5 \quad k=0.0077$				
58.3256	0.9742	29.1750	20.8533	7.0890
1.2900	0.1789	0.6633	0.8246	0.2255
0.6701	0.0632	0.3391	0.2550	0.0004
1.1589	0.2811	0.6285	0.5404	0.2261
0.6419	0.1789	0.3564	0.3033	0.2255

The estimated optimum value of b using equations (3.10) and (3.11) of Kaciranlar (1999, p. 450) are found to be $b_{OLS} = 0.9996484$ and $b_{RLE} = 0.995122$ respectively. It was noticed that both of the estimated b are approximately equal to 1. One would therefore expect that ULIUE will give similar result like that of URLSE and RLIUE will produce similar result like that of RLSE. These results can be evident from Tables 6 and 7. It is noted from numerical analysis that the RRRE performed better than URLSE, URRE, RLSE, SRLSE, SRRRE, PTLSE, and PTRRE in the sense of having smaller sampling variance. From the analysis, it is also noted that the RRRE

performed equivalently or better than RLIUE in the sense of having smaller sampling variance.

Using $k = 0.0077$, the regression analysis for the RRRE and the RLIUE (in the parenthesis) have been presented in Tables 8 and 9 where σ^2 is estimated by $\tilde{\sigma}^2$ and $\hat{\sigma}^2$ respectively. It may be found from Table 8 that all the regression coefficients are statistically significant under both RRRE and RLIUE. However, from Table 9, it was found that all the regression coefficients are statistically significant under the restricted ridge regression model and only β_2 is significant under the RLIUE model.

Table 8: Ridge regression (Restricted Liu) analysis, when σ^2 is estimated by $\tilde{\sigma}^2$

Coefficients	SE	t-value	P-value
13.9436 (101.3373)	0.1612 (1.1632)	69.8131 (87.1194)	0.0000 (0.0000)
0.6545 (-0.2444)	0.0316 (0.0370)	43.6612 (-6.6063)	0.0000 (0.0002)
0.9058 (0.0051)	0.0000 (0.0001)	Inf (51.2085)	0.0000 (0.0000)
1.1742 (0.2547)	0.0447 (0.0371)	20.6098 (6.8648)	0.0000 (0.0001)
0.6383 (-0.2445)	0.0316 (0.0370)	16.4263 (-6.6076)	0.0000 (0.0002)

Table 9: Ridge regression (Restricted Liu) analysis, when σ^2 is estimated by $\hat{\sigma}^2$

Coefficients	SE	t-value	P-value
13.9436 (101.3373)	0.9742 (7.0890)	11.5519 (14.2950)	0.0000 (0.0000)
0.6545 (-0.2444)	0.1789 (0.2255)	7.7121 (-1.0840)	0.0001 (0.3100)
0.9058 (0.0051)	0.0632 (0.0004)	15.5926 (12.8021)	0.0000 (0.0000)
1.1742 (0.2547)	0.2811 (0.2261)	3.2773 (1.1264)	0.0112 (0.2926)
0.6383 (-0.2445)	0.1789 (0.2255)	2.9015 (-1.0842)	0.0198 (0.3099)

Conclusion

In this article, some ridge type estimators were considered for estimating the regression parameters under a parametric restriction. The proposed estimators have been studied by using the variance criterion. It was found that a sufficient and necessary condition for all the ridge estimators to have smaller sampling variance than their corresponding least squares estimators. It was found that, under both H_0 and H_a , the RRRE dominates all other ridge estimators and corresponding least squares estimators. As Δ moves away from 0, the conditions of superiority of the PTRRE and SRRRE over the URRE and URLSE are determined. The findings were illustrated by analyzing the Portland Cement dataset. From the analysis, it is evident that RRRE is superior to other proposed ridge regression estimators and as good as restricted Liu estimator.

References

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis*. (2nd Ed.). New York, N.Y.: John Wiley.

Bancroft, T. A. (1944). On biases in estimation due to use of preliminary tests of significance. *Annals of Mathematics and Statistics*, 15, 190-204.

Bancroft, T. A. (1964). Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance. *Biometrics*, 20, 427-442.

Bock, M. E., Yancey, T. A. & Judge, G. G. (1973). The statistical consequences of preliminary test estimators in regression. *Journal of the American Statistical Association*, 68, 109-116.

Giles, A. J. (1991). Pre-testing for linear restrictions in a regression model with spherically symmetric disturbances. *Journal of Econometrics*, 50, 377-398.

Gorman, J. W., & Toman, R. J. (1966). Selection of variables for fitting equations to data. *Technometrics*, 8, 27-51.

Hald, A. (1952). *Statistical theory with Engineering Applications*. New York, N.Y.: Wiley.

Han, C-P, & Bancroft, T. A. (1968). On pooling mean when variance is unknown. *Journal of the American Statistical Association*, 63, 1333-1342.

- Haq, M. S. & Kibria, B. M. G. (1996). A shrinkage estimator for the restricted linear regression model: Ridge regression approach. *Journal of Applied Statistical Science*, 3(4), 301-316.
- Hoerl, A. E. & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12, 55-67.
- Hoerl, A. E., Kennard, R. W. & Baldwin, K. F. (1975). Ridge regression: Some simulation. *Communications in Statistics*, A 4, 105-123.
- Judge, G. G. & Bock, M E. (1978). *The statisticaal implications of pre-test and Stein-rule estimators in econometrics*. Amsterdam: North-Holland Publishing Company.
- Kaciranlar, S., Akdeniz, S. S. F., Styan, G. P. & Werner, H. J. (1999). A new biased estimators in linear regression and detailed analysis of the widely-analysed dataset on portland cement. *Sankhya*, Series B, 443-459.
- Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics-Simulation and Computation*, 32, 419-435.
- Kibria, B. M. G. (2004a). Performance of the shrinkage preliminary test ridge regression estimators based on the conflicting of W, LR and LM tests. *Journal of Statistical Computation and Simulation*, 74(11), 793 – 810.
- Kibria, B. M. G. (2004b). On some ridge regression estimators under possible stochastic constraint. *Pakistan Journal of Statistics* -Special issue in the honor of Professor Dr. Mir Masoom Ali, 20(1), 1-23.
- Kibria, B. M. G. and Saleh, A. K. Md. E. (1993). Performance of shrinkage preliminary test estimator in regression analysis. *Jahangirnagar Review*, A 17, 133-148.
- Kibria, B. M. G. & Saleh, A. K. Md. E. (2004). Performance of positive rule estimator in the ill-conditioned Gaussian regression model. *Calcutta Statistical Association Bulletin*, 55 (219-220), 211-241.
- Kibria, B. M. G. & Saleh, A. K. Md. E. (2005a). Pooling multivariate data under W, LR, and LM tests. *Statistical Papers*, 47, 49-68.
- Kibria, B. M. G. & Saleh, A. K. Md. E. (2005b). Comparison between Han-Bancroft and Brook methods to determine the optimal significance level for pre-test Estimator. *Journal of Probability and Statistical Science*, 3, 293-303.
- Kibria, B. M. G. & Saleh, A. K. Md. E. (2006). Optimum critical value for pre-test estimators. *Communications in Statistics-Simulation and Computation*. 35 (2), 309-319.
- Lawless, J. F. & Wang, P. (1976). A simulation study of ridge and other regression estimators. *Communications in Statistics*, A 5, 307-323.
- Liu, K. (1993). A new class of biased estimator in linear regression. *Communications in Statistics*, A, 22, 393-402.
- Nomura, M. (1988). On the almost unbiased ridge regression estimator. *Communications in Statistics*, B, 17, 729-743.
- Obenchain, R. L. (1975). Ridge analysis following a preliminary test of the shrunken hypothesis. *Technometrics*, 17, 431-441.
- Piepel, G. & Redgate, T. (1998). A mixture experiemnt analysis of the Hald cement data. *The American Statistician*, 52, 23-30.
- Saleh, A. K. Md. E. & Han, C-P. (1990). Shrinkage estimation in regression analysis. *Estadistica*, 42-40.63.
- Saleh, A. K. Md. & Kibria, B. M. G. (1993). Performances of some new preliminary test ridge regression estimators & their properties. *Communications in Statistics-Theory and Methods*, 22, 2747-2764.
- Sarkar, N. (1992). A new estimator combining the ridge regression and the restricted least squares method of estimation. *Communications in Statistics*, A 21, 1987-2000.
- Sutradhar, B. C. (1990). A note on the power of the F-test for testing linear hypothesis with elliptic t errors. . *Communications in Statistics*, A 19, 863--867.
- Woods, H., Steinour, H. H., & Starke, H. R. (1932). Effect of composition of Portland cement on heat evolved during hardening. *Industrial and Engineering Chemistry*, 24, 1207-1214.