

A New Biased Estimator in Linear Regression Model

Huiping Hao and Chunping Li

Department of Mathematics, Hubei Engineering University, Hubei Xiaogan, 432000

haohuiping@163.com, lichunping315@163.com

Keywords: Linear regression model; Stochastic restricted ridge estimator; Mixed estimator; Mean squared error matrix; Biased estimator.

Abstract. In this paper, we propose a new ridge type estimator to overcome the multicollinearity problem, and we call the new biased estimator as the stochastic restricted ridge estimator (SRRE). In the mean squared errors matrix sense SRRE will be compared with several other biased estimators. The necessary and sufficient conditions for the superiority of the new estimators SRRE over the ridge estimator (RE) and the Mixed Regression Estimator (MRE) in the mean squared error matrix criterion are derived. A numerical example with Monte Carlo simulation is given to illustrate the theoretical results.

Introduction

We consider the standard multiple linear regression model

$$\begin{cases} Y = X\beta + \varepsilon \\ \varepsilon \sim (0, \sigma^2 I_n) \end{cases} \quad (1)$$

The least squares method is applied to (1.1), we can obtain the least squares estimator of β (OLSE) as

$$\hat{\beta} = (X'X)^{-1} X'Y = S^{-1} X'Y, \quad S = X'X \quad (2)$$

In 1970, Hoerl and Kennard have introduced a biased estimator called the ridge estimator (RE) to overcome the multicollinearity problem, and defined as

$$\hat{\beta}_R(k) = (X'X + kI)^{-1} X'Y, \quad k > 0 \quad (3)$$

Theil and Goldberger (1961) have introduced an estimator for the regression coefficient vector β by adding stochastic linear restricted to model (1.1)

$$r = R\beta + \varepsilon \quad (4)$$

The Mixed Regression Estimator (MRE) has been introduced by Theil and Goldberger (1961) is based on sample information (1) with prior information (4), and is given by

$$\hat{\beta}_M = \hat{\beta} + S^{-1} R'(\Omega + RS^{-1}R')^{-1}(r - R\hat{\beta}) \quad (5)$$

The restricted Ridge estimator (RRE) has been introduced by Sarkar(1992) is based on sample information (1) with prior information (4),and is given by

$$\hat{\beta}_{RRE}(k) = W[\hat{\beta} + S^{-1} R'(RS^{-1}R')^{-1}(r - R\hat{\beta})] \quad (6)$$

where

$$W = (I + kS^{-1})^{-1}, S = X'X.$$

In this paper we introduce another alternative ridge type estimator. We call the new biased estimator as the stochastic restricted ridge estimator (SRRE). In the mean squared errors matrix sense SRRE will be compared with several other biased estimators, and the necessary and sufficient conditions for the superiority over these biased estimators will be derived.

Some Notations and Lemmas

Lemma1 Assume square matrixes A, C are not singular, and B, D are matrixes with proper orders, then

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

Proof see Rao and Toutenburg(1995).

Lemma2 Let M be a positive definite matrix, namely

$M > 0$, α be some vector, then

$$M - \alpha\alpha' \geq 0 \Leftrightarrow \alpha'M^{-1}\alpha \leq 1$$

Proof see Farebrother(1976).

Lemma3 Let $\hat{\beta}_j$, $j = 1, 2$ be two competing linear estimators of β . Suppose that

$$D = Cov(\hat{\beta}_1) - Cov(\hat{\beta}_2) > 0$$

where $Cov(\hat{\beta}_j)$ denotes the covariance matrix of $\hat{\beta}_j$. Then

$$\Delta(\hat{\beta}_1, \hat{\beta}_2) = MSEM(\hat{\beta}_1) - MSEM(\hat{\beta}_2) \geq 0 \Leftrightarrow c_2'(D + d_1d_1')c_2 \leq 1$$

where $MSEM(\hat{\beta}_j)$ denote the mean squared error matrix and bias vector of $\hat{\beta}_j$.

Proof see Trenkler and Toutenburg(1990).

Lemma4 Let $M_{n \times n} > 0$, $N_{n \times n} \geq 0$, then

$$M > N \Leftrightarrow \lambda_i(NM^{-1}) < 1$$

where $\lambda_i(M)$ denotes the i -th Eigen value of M .

Proof see Wang et al.(2006).

Model Specification and Estimation

Following Sarkar(1992), now we introduce an alternative Ridge estimator for β as

$$\hat{\beta}_{SRRE}(k) = W\hat{\beta}_M \quad (7)$$

where

$$W = (I + kS^{-1})^{-1}, S = X'X,$$

And

$$\hat{\beta}_M = \hat{\beta} + S^{-1}R'(\Omega + RS^{-1}R')^{-1}(r - R\hat{\beta})$$

is the Mixed Regression Estimator (MRE).

We call $\hat{\beta}_{SRRE}(k)$ as the stochastic restricted Ridge estimator(SRRE) of β .

The estimator $\hat{\beta}_{SRRE}(k)$ as the OLSE of β is derived by the following model

$$\begin{cases} Y = X\beta + \varepsilon \\ k^{-\frac{1}{2}}Sg = k^{\frac{1}{2}}\beta + e \end{cases} \quad (8)$$

where

$$g = S^{-1}R'(\Omega + RS^{-1}R')^{-1}(r - R\beta).$$

Then the OLSE of β from the augmented model (8) is

$$\hat{\beta}_{SRRE}(k) = (S + kI)^{-1}(X'Y + Sg) = W\hat{\beta}_M \quad (9)$$

Comparisons Among Biased Estimators

We note that for any estimator $\hat{\beta}$ of β , its mean squared errors matrix (MSEM) is defined as

$$MSEM(\hat{\beta}) = E(\hat{\beta} - \beta)(\hat{\beta} - \beta)' = Cov(\hat{\beta}) + Bias(\hat{\beta})Bias(\hat{\beta})' \quad (10)$$

where $Cov(\hat{\beta})$ is the dispersion matrix and $Bias(\hat{\beta}) = E(\hat{\beta}) - \beta$ is the bias vector.

For two given estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ of β , the estimator $\hat{\beta}_2$ is said to be superior to $\hat{\beta}_1$ in the MSEM sense, if and only if

$$\Delta(\hat{\beta}_1, \hat{\beta}_2) = MSEM(\hat{\beta}_1) - MSEM(\hat{\beta}_2) \geq 0$$

Let

$$\delta_* = R'(\Omega + RS^{-1}R')^{-1}\delta,$$

Then, the MSEM of the RE, the MRE and the SRRE are given as

$$\begin{aligned} MSEM(\hat{\beta}_M) &= \sigma^2 A + S^{-1}\delta_*\delta_*'S^{-1}, \\ MSEM(\hat{\beta}_R(k)) &= \sigma^2 WS^{-1}W + b_1b_1', \\ MSEM(\hat{\beta}_{SRRE}(k)) &= \sigma^2 WAW' + b_2b_2' \end{aligned}$$

where

$$b_1 = (W - I)\beta, \quad b_2 = (W - I)\beta + WS^{-1}\delta_*.$$

If the restriction

$$r = R\beta + e$$

is correct, then

$$\delta = 0,$$

and consequently

$$\delta^* = 0.$$

Then

$$\begin{aligned} \Delta_1 &= MSEM(\hat{\beta}_R(k)) - MSEM(\hat{\beta}_{SRRE}(k)) = \sigma^2 W(S^{-1} - A)W' \\ \Delta_2 &= MSEM(\hat{\beta}_M) - MSEM(\hat{\beta}_{SRRE}(k)) = \sigma^2 (A - WAW') - b_1b_1' \end{aligned}$$

If the restriction $r = R\beta + e$ is not correct, then $\delta \neq 0$. Then

$$\begin{aligned} \Delta_3 &= MSEM(\hat{\beta}_R(k)) - MSEM(\hat{\beta}_{SRRE}(k)) = W(D + d_1d_1' - d_2d_2')W' \\ \Delta_4 &= MSEM(\hat{\beta}_M) - MSEM(\hat{\beta}_{SRRE}(k)) = \sigma^2 (A - WAW') + S^{-1}\delta_*\delta_*'S^{-1} - b_2b_2' \end{aligned}$$

Where

$$D = \sigma^2 S^{-1}R'(\Omega + RS^{-1}R')^{-1}RS^{-1}, \quad d_1 = kS^{-1}\beta, \quad d_2 = kS^{-1}\beta - S^{-1}\delta_*$$

In the following theorem, we give the necessary and sufficient condition for the stochastic restricted Ridge estimator (SRRE) $\hat{\beta}_{SRRE}(k)$ be superior to $\hat{\beta}_R(k)$ and $\hat{\beta}_M$ in the MSEM sense.

Theorem1 In the Mean Squared Error Matrix sense,

(a) when parametric restriction $r = R\beta + e$ is true, $\hat{\beta}_{SRRE}(k)$ is always superior to $\hat{\beta}_R(k)$.

(b) when parametric restriction $r = R\beta + e$ is not true, $\hat{\beta}_{SRRE}(k)$ is always superior to $\hat{\beta}_R(k)$

if and only if $d_2'(D + d_1d_1')^{-1}d_2 \leq 1$.

Proof (a) Because

$$\begin{aligned} \Delta_1 &= \sigma^2 W(S^{-1} - A)W' \\ A^{-1} - S &= (S + R'\Omega^{-1}R) - S = R'\Omega^{-1}R \end{aligned}$$

And

$$R'\Omega^{-1}R > 0,$$

then

$$A^{-1} - S > 0,$$

so

$$S^{-1} - A > 0,$$

that is

$$\Delta_1 > 0.$$

(b) We firstly prove

$$D = \sigma^2 S^{-1}R'(\Omega + RS^{-1}R')^{-1}RS^{-1} > 0$$

According to (a) has

$$S^{-1} - A > 0$$

but

$$S^{-1} - A = S^{-1}R'(\Omega + RS^{-1}R')^{-1}RS^{-1}$$

Then

$$D > 0.$$

Because

$$d_1 d_1' = k^2 S^{-1} \beta \beta' S^{-1} \geq 0$$

by lemma3 has

$$\Delta_3 \geq 0 \Leftrightarrow d_2'(D + d_1 d_1')^{-1} d_2 \leq 1.$$

Now we make comparisons between $\hat{\beta}_{SRRE}(k)$ and the Mixed Regression Estimator (MRE) $\hat{\beta}_M$ in the following theorem.

Theorem2 In the Mean Squared Error Matrix sense, when

$$\lambda_1(WAW'A^{-1}) < 1$$

then $\hat{\beta}_{SRRE}(k)$ is superior to $\hat{\beta}_M$ if and only if (a) When parametric restriction $r = R\beta + e$ is true, then

$$b_1'(A - WAW')^{-1} b_1 \leq 1$$

(b) When parametric restriction $r = R\beta + e$ is not true, then $b_2'[\sigma^2(A - WAW') + S^{-1}\delta_*\delta_*'S^{-1}]^{-1} b_2 \leq 1$

Proof (a) For $A^{-1} - S > 0$, $A > 0$ and $WAW' \geq 0$, then when $\lambda_1(WAW'A^{-1}) < 1$, where $\lambda_1(\bullet)$ is the maximum Eigen Value of \bullet , by lemma4, we can get $\sigma^2(A - WAW') > 0$ by lemma3, then has

$$\Delta_2 \geq 0 \Leftrightarrow b_1'(A - WAW')^{-1} b_1 \leq 1.$$

(b) From (a) and $\lambda_1(WAW'A^{-1}) < 1$, We can get $\sigma^2(A - WAW') > 0$, by lemma3, then we has

$$\Delta_4 \geq 0 \Leftrightarrow b_2'[\sigma^2(A - WAW') + S^{-1}\delta_*\delta_*'S^{-1}]^{-1} b_2 \leq 1$$

Numerical Examples and Conclusions

Here we show the numerical comparison of estimators. We put

$$X' = \begin{pmatrix} 1.1 & 1.4 & 1.7 & 1.7 & 1.8 & 1.8 & 1.9 & 2.0 & 2.3 & 2.4 \\ 1.1 & 1.5 & 1.8 & 1.7 & 1.9 & 1.8 & 1.8 & 2.1 & 2.4 & 2.5 \end{pmatrix},$$

$$Y' = (16.3 \quad 16.8 \quad 19.2 \quad 18.0 \quad 19.5 \quad 20.9 \quad 21.1 \quad 20.9 \quad 20.3 \quad 22.0)$$

We firstly obtain the ordinary least square estimator of β

$$\hat{\beta} = (X'X)^{-1}X'Y = S^{-1}X'Y = (21.4169 \quad 10.598)'$$

with

$$MSE(\hat{\beta}) = 332.6172$$

and

$$\hat{\sigma}_{OLSE}^2 = 6.8495.$$

Consider the following stochastic linear restrictions $r = R\beta + \varepsilon$,

$$R = \begin{pmatrix} 1 & -1 \end{pmatrix}, \quad e \sim N(0, \hat{\sigma}_{OLSE}^2)$$

For the MRE, RE and SRRE, their estimated MSE values are obtained by replacing in the corresponding theoretical MSE expressions by their OLSE (see Table I).

Table 1

	k=0	k=0.0001	k=0.0005	k=0.0009	k=0.0010	k=0.010
MRE	22.0137	22.0137	22.0137	22.0137	22.0137	22.0317
RE	332.6172	329.4241	317.3342	306.2657	303.6478	206.0209
SRRE	22.0137	21.8146	21.2777	21.1234	21.1398	65.2605

From Table 1, we can see that the SRRE outperforms the MRE when parameter the ridge parameter k is small, while the situation is reversed when k becomes large enough. When compared to the RE, SRRE always have smaller the MSE values for different values of k . And the results agree with our theoretical findings in Theorems 1 and 2. We can see that our estimator is more meaningful in practice.

Acknowledgment

This work is supported by the Humanity and Social Science Youth Foundation of Ministry of Education of China (No. 15YJCZH055).

References

- [1] CR .Rao, H. Toutenburg , “Linear models: least squares and alternatives,” Springer.New Yor,1995.
- [2] RW. Farebrother, “Further results on the mean square error of ridge regression,”J R Stat Soc B, 38:248–250,1976.
- [3] H.Toutenburg, “Mean square error matrix comparisons between biased estimators: an overview of recent results,” Statistical Papers 31:165–179,1990.
- [4] SG .Wang et al, “Matrix inequalities,”2nd edu,Chinese Science press, Beijing, 2006.
- [5] Nityananda Sarkar. “A new estimator combining the ridge regression and the erestricted least squares methods of estimation”. Communications in statistics-Theory and Methods.21 (7), pp.1987–2000, 1992
- [6] S.Kaciranlar, S.Sakallioglu, F.Akdeniz, G. P. H.Styan, H. J Werner. “A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland Cement”. Sankhya Indian J. Stat. 61(B): 443-459, 1999.
- [7] M. H Hubert, P.Wijekoon. “Improvement of Liu estimator in linear regression model”. Stat. Pap. 47: 471-479, 2006.
- [8] H.Yang, J. W. Xu. “An alternative stochastic restricted Liu estimator in linear regression”. Stat. Pap. 50: 639-647, 2009.
- [9] Y. L. Li, H. Yang. “A new stochastic mixed ridge estimator in linear regression model”. Stat . Pap. 51: 315-323, 2010.
- [10]C. R.Rao, H.Toutenburg, H. C. Shalabh. “Linear Models and Generalizations- Least Squares and Alternatives”. Berlin: Springer, 2008.
- [11]G.Trenkler, H. Toutenburg. “Mean square error matrix comparisons between biased estimators: an overview of recent results”. Stat. Pap. 31, pp.165-179, 1990.
- [12]M. H. J. Gruber,. “Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimators”. New York: Marcell Dekker, 1998.
- [13]F. Akdeniz, H. Erol. “Mean squared error matrix comparisons of some biased estimators in linear regression”. Comm. Stat. Theor. Meth. 32: 2389-2413, 2003
- [14]H. Yang, X. Chang, D. Q. Liu. “Improvement of the Liu Estimator in Weighted Mixed Regression”. Comm. Stat. Theor. Meth. 38: 285-292, 2009.
- [15]Y. L.Li, H.Yang. “A new stochastic mixed ridge estimator in linear regression model”. Stat . Pap. 51: 315-323, 2010.