Research Article

On the Performance of Principal Component Liu-Type Estimator under the Mean Square Error Criterion

Jibo Wu[1,2]

[1] *School of Mathematics and Finances, Chongqing University of Arts and Sciences, Chongqing 402160, China*
[2] *Department of Mathematics and KLDAIP, Chongqing University of Arts and Sciences, Chongqing 402160, China*

Correspondence should be addressed to Jibo Wu; linfen52@126.com

Wu (2013) proposed an estimator, principal component Liu-type estimator, to overcome multicollinearity. This estimator is a general estimator which includes ordinary least squares estimator, principal component regression estimator, ridge estimator, Liu estimator, Liu-type estimator, r-k class estimator, and r-d class estimator. In this paper, firstly we use a new method to propose the principal component Liu-type estimator; then we study the superior of the new estimator by using the scalar mean squares error criterion. Finally, we give a numerical example to show the theoretical results.

1. Introduction

Consider the multiple linear regression model

$$y = X\beta + \epsilon, \tag{1}$$

where y is an $n \times 1$ vector of observation, X is an $n \times p$ known matrix of rank p, β is a $p \times 1$ vector of unknown parameters, and ϵ is an $n \times 1$ vector of disturbances with expectation $E(\epsilon) = 0$ and variance-covariance matrix $\text{Cov}(\epsilon) = \sigma^2 I_n$.

According to the Gauss-Markov theorem, the classical ordinary least squares estimator (OLSE) is obtained as follows:

$$\hat{\beta} = \left(X'X\right)^{-1} X'y. \tag{2}$$

The OLSE has been regarded as the best estimator for a long time. However, when multicollinearity is present and the matrix $X'X$ is ill-conditioned, the OLSE is no longer a good estimator. To improve OLSE, many ways have been proposed. One way is to consider biased estimator, such as, principal component regression estimator [1], ridge estimator [2], Liu estimator [3], Liu-type estimator [4], two-parameter ridge estimator [5], r-k class estimator [6], r-d class estimator [7], and modified r-k class estimator [8].

An alternative method to overcome the multicollinearity is to consider the restrictions. Xu and Yang [9] introduced

a stochastic restricted Liu estimator; Li and Yang [10] introduced a stochastic restricted ridge estimator.

To overcome multicollinearity, Hoerl and Kennard [2] solve the following problem:

$$\min_{\beta} \left\{ \left(y - X\beta\right)' \left(y - X\beta\right) + k\left(\beta\beta' - c\right) \right\}, \tag{3}$$

where k is a Lagrangian multiplier and c is a constant, and obtain the ridge estimator (RE):

$$\hat{\beta}(k) = \left(X'X + kI\right)^{-1} X'y, \quad k > 0. \tag{4}$$

Liu [3] introduced the Liu estimator (LE):

$$\hat{\beta}(d) = \left(X'X + I\right)^{-1} \left(X'y + d\hat{\beta}\right), \quad 0 < d < 1, \tag{5}$$

where $\hat{\beta}$ is OLSE. This estimator can be obtained by solving the following problem:

$$\min_{\beta} \left\{ \left(y - X\beta\right)' \left(y - X\beta\right) + \left(\beta - d\hat{\beta}\right)\left(\beta - d\hat{\beta}\right)' \right\}. \tag{6}$$

This estimator can also be obtained by the following ways. Suppose that β satisfied $d\hat{\beta} = \beta + e$. Then, we use the mixed method [11]; we can obtain the Liu estimator.

Recently, Huang et al. [4] introduced a Liu-type estimator which includes the OLSE, RE, and LE, defined as follows:

$$\widehat{\beta}(k, d) = \left(X'X + kI\right)^{-1}\left(X'y + d\widehat{\beta}\right), \quad k > 0, \; 0 < d < 1, \tag{7}$$

where $\widehat{\beta}$ is OLSE. This estimator can be obtained by solving the following problem:

$$\min_{\beta}\left\{(y - X\beta)'(y - X\beta) \right.$$
$$\left. + \left(k^{1/2}\beta - \frac{d\widehat{\beta}}{k^{1/2}}\right)\left(k^{1/2}\beta - \frac{d\widehat{\beta}}{k^{1/2}}\right)'\right\}. \tag{8}$$

Let us consider the following transformation for the model (1):

$$y = XTT'\beta + \epsilon = Z\alpha + \epsilon, \tag{9}$$

where

$$X'X = \left(T_r, T_{p-r}\right)\begin{pmatrix} \Lambda_r & 0 \\ 0 & \Lambda_{p-r} \end{pmatrix}\begin{pmatrix} T'_r \\ T'_{p-r} \end{pmatrix}. \tag{10}$$

Λ_r and Λ_{p-r} are diagonal matrices such that that the main diagonal elements of the $r \times r$ matrix Λ_r are the r largest eigenvalues of $X'X$, while Λ_{p-r} are the remaining $p - r$ eigenvalues. The $p \times p$ matrix $T = (T_r, T_{p-r})$ is orthogonal with $T_r = (t_1, t_2, \ldots, t_r)$ consisting of its first r columns and $T_{p-r} = (t_{r+1}, t_{r+2}, \ldots, t_p)$ consisting of the remaining $p - r$ columns of the matrix T. The PCRE for β can be written as

$$\widehat{\beta}_r = T_r\left(T'_r X' X T_r\right)^{-1} T'_r X' y. \tag{11}$$

Baye and Parker [6] introduced the application of ridge approach to improve the PCR estimator, namely, r-k class estimator as

$$\widehat{\beta}_r(k) = T_r\left(T'_r X' X T_r + kI_r\right)^{-1} T'_r X' y, \quad k > 0. \tag{12}$$

Alternatively, Kaçıranlar and Sakallıoğlu [7] introduced the r-d class estimator which is the combination of the LE and the PCRE, which is defined as follows:

$$\widehat{\beta}_r(d) = T_r\left(T'_r X' X T_r + I_r\right)^{-1}\left(T'_r X' y + dT'_r\widehat{\beta}_r\right), \tag{13}$$
$$0 < d < 1.$$

Wu [12] proposed the principal component Liu-type estimator (PCTTE), which is defined as

$$\widehat{\beta}_r(k, d) = T_r\left(T'_r X' X T_r + kI_r\right)^{-1}\left(T'_r X' y + dT'_r\widehat{\beta}_r\right), \tag{14}$$
$$k > 0, \quad 0 < d < 1.$$

In this paper, firstly we use a new method to propose the principal component Liu-type estimator. Then, we show that, under certain conditions, the PCTTE is superior to the related estimator in the mean square error criterion. Finally, we give a numerical example to illustrate the theoretical results.

2. The Principal Component Liu-Type Estimator

Using the symbols in (9) and (10), (1) can be written as follows:

$$y = XT_r T'_r \beta + XT_{p-r} T'_{p-r}\beta + \epsilon$$
$$= Z_r \alpha_r + Z_{p-r}\alpha_{p-r} + \epsilon. \tag{15}$$

The PCRE can be obtained by omitted $Z_{p-r}\alpha_{p-r}$, and the model (15) reduced to:

$$y = Z_r \alpha_r + \epsilon. \tag{16}$$

Then, solve the following problem:

$$\left(y - Z_r\alpha_r\right)'\left(y - Z_r\alpha_r\right), \tag{17}$$

we obtain

$$\widehat{\alpha}_r = \left(Z'_r Z_r\right)^{-1} Z'_r y. \tag{18}$$

Then, transforming $\widehat{\alpha}_r$ to the original parameter space, we can get the PCRE of parameter β.

Now, we give a method to obtain the principal component Liu-type estimator. Let c be a constant and k a Lagrangian multiplier, minimizing

$$\left(y - Z_r\alpha_r\right)'\left(y - Z_r\alpha_r\right) + \frac{1}{k}\left[\left(k\alpha_r - d\widehat{\alpha}_r\right)\left(k\alpha_r - d\widehat{\alpha}_r\right)' - c\right], \tag{19}$$

where $\widehat{\alpha}_r = (Z'_r Z_r)^{-1} Z'_r y$. Then we get

$$\widehat{\alpha}_r(k, d) = \left(Z'_r Z_r + kI_r\right)^{-1}\left(Z'_r y + d\widehat{\alpha}_r\right). \tag{20}$$

After transforming back to original parameters pace, we obtain

$$\widehat{\beta}_r(k, d) = T_r\left(T'_r X' X T_r + kI_r\right)^{-1}$$
$$\times \left(T'_r X' y + dT'_r\widehat{\beta}_r\right), \quad k > 0, \; 0 < d < 1. \tag{21}$$

This estimator can also be got by minimizing the function

$$\left(k\alpha_r - d\widehat{\alpha}_r\right)\left(k\alpha_r - d\widehat{\alpha}_r\right)' + k\left[\left(y - Z_r\alpha_r\right)'\left(y - Z_r\alpha_r\right) - c\right]. \tag{22}$$

It is easy to see that the new estimator $\widehat{\beta}_r(k, d)$ has the following properties:

(1) $\widehat{\beta}_r(1, 1) = \widehat{\beta}_r = T_r(T'_r X' X T_r)^{-1} T'_r X' y$ is the PCRE;

(2) $\widehat{\beta}_p(1, 1) = \widehat{\beta}_{\text{OLS}} = (X'X)^{-1}X'y$ is the OLSE;

(3) $\widehat{\beta}_r(1, d) = \widehat{\beta}_r(d) = \widehat{\beta}_r(d) = T_r(T'_r X' X T_r + I_r)^{-1}(T'_r X' y + dT'_r\widehat{\beta}_r)$ is the r-d class estimator;

(4) $\widehat{\beta}_p(1, d) = \widehat{\beta}(d) = (X'X + I)^{-1}(X'y + d\widehat{\beta}_{\text{OLS}})$ is the LE;

(5) $\widehat{\beta}_r(k, 0) = \widehat{\beta}_r(k) = T_r(T'_r X' X T_r + kI_r)^{-1}T'_r X' y$ is the r-k class estimator;

(6) $\widehat{\beta}_p(k, 0) = \widehat{\beta}(k) = (X'X + kI)^{-1}X'y$ is the RE;

(7) $\widehat{\beta}_p(k, d) = \widehat{\beta}(k, d) = (X'X + kI)^{-1}(X'y + kd\widehat{\beta}_{\text{OLS}})$ is the LTE.

3. Superiority of the Principal Component Liu-Type Estimator over Some Estimators in the Mean Square Error Criterion

The mean square error (MSE) of an estimator $\widehat{\beta}$ is defined as

$$M\left(\widehat{\beta}\right) = E\left[\left(\widehat{\beta} - \beta\right)'\left(\widehat{\beta} - \beta\right)\right]$$
$$= \operatorname{tr} V\left(\widehat{\beta}\right) + \left[\operatorname{Bias}\left(\widehat{\beta}\right)\right]'\left[\operatorname{Bias}\left(\widehat{\beta}\right)\right],$$

(23)

where $V(\widehat{\beta})$ is the dispersion matrix and $\operatorname{Bias}(\widehat{\beta})$ is the bias vector. For two given estimators $\widehat{\beta}_1$ and $\widehat{\beta}_2$, the estimator $\widehat{\beta}_2$ is said to be superior to $\widehat{\beta}_1$ in the MSE criterion, if and only if

$$\Delta\left(\widehat{\beta}_1, \widehat{\beta}_2\right) = M\left(\widehat{\beta}_1\right) - M\left(\widehat{\beta}_2\right) \geq 0. \tag{24}$$

3.1. $\widehat{\beta}_r(k,d)$ versus $\widehat{\beta}_r$. Firstly, we compute that

$$V\left(\widehat{\beta}_r(k,d)\right) = \sigma^2 T_r S_r(k)^{-1} S_r(d) \Lambda_r^{-1} S_r(d) S_r(k)^{-1} T_r',$$
$$E\left(\widehat{\beta}_r(k,d)\right) = T_r S_r(k)^{-1} S_r(d) T_r' \beta,$$

(25)

where $S_r(k) = \Lambda_r + kI_r$. Then, the mean square error (MSE) of $\widehat{\beta}_r(k,d)$ is given as follows:

$$M\left(\widehat{\beta}_r(k,d)\right) = \operatorname{tr} V\left(\widehat{\beta}_r(k,d)\right)$$
$$+ \left[\operatorname{Bias}\left(\widehat{\beta}_r(k,d)\right)\right]'\left[\operatorname{Bias}\left(\widehat{\beta}_r(k,d)\right)\right]$$
$$= \sum_{i=1}^{r} \frac{\sigma^2(\lambda_i + d)^2 + (k-d)^2\alpha_i^2\lambda_i}{\lambda_i(\lambda_i + k)^2} + \sum_{i=r+1}^{p}\alpha_i^2.$$

(26)

Let $k = d$ in (26); we obtain the MSE of $\widehat{\beta}_r$ as follows:

$$M\left(\widehat{\beta}_r\right) = \sum_{i=1}^{r}\frac{\sigma^2}{\lambda_i} + \sum_{i=r+1}^{p}\alpha_i^2. \tag{27}$$

Now we consider the following difference:

$$M\left(\widehat{\beta}_r(k,d)\right) - M\left(\widehat{\beta}_r\right)$$
$$= \sum_{i=1}^{r}\frac{\sigma^2(\lambda_i + d)^2 + (k-d)^2\alpha_i^2\lambda_i}{\lambda_i(\lambda_i + k)^2} + \sum_{i=r+1}^{p}\alpha_i^2$$
$$- \left\{\sum_{i=1}^{r}\frac{\sigma^2}{\lambda_i} + \sum_{i=r+1}^{p}\alpha_i^2\right\}$$
$$= \sum_{i=1}^{r}(d-k)\frac{\left(\alpha_i^2\lambda_i + \sigma^2\right)d + 2\sigma^2\lambda_i + \left(\sigma^2 - \alpha_i^2\lambda_i\right)k}{\lambda_i(\lambda_i + k)^2}.$$

(28)

If $\sigma^2 - \alpha_i^2\lambda_i > 0$, then when $0 < d < k$, $M(\widehat{\beta}_r(k,d)) - M(\widehat{\beta}_r) < 0$. If $\sigma^2 - \alpha_i^2\lambda_i < 0$, then when $0 < d < k < \min\{((\alpha_i^2\lambda_i + \sigma^2)d + 2\sigma^2\lambda_i)/(\alpha_i^2\lambda_i - \sigma^2)\}$ and $\max\{((\alpha_i^2\lambda_i + \sigma^2)d + 2\sigma^2\lambda_i)/(\alpha_i^2\lambda_i - \sigma^2)\} < k < d < 1$, $M(\widehat{\beta}_r(k,d)) - M(\widehat{\beta}_r) < 0$. So we have the following theorem.

Theorem 1. *The estimator $\widehat{\beta}_r(k,d)$ is superior to the estimator for $\widehat{\beta}_r$ under the mean square error criterion for:*

(a) $0 < d < k$ *if* $\sigma^2 - \alpha_i^2\lambda_i > 0$ *for all* $i = 1, \ldots, r$,

(b) $0 < d < k < \min\{((\alpha_i^2\lambda_i + \sigma^2)d + 2\sigma^2\lambda_i)/(\alpha_i^2\lambda_i - \sigma^2)\}$ *and* $\max\{((\alpha_i^2\lambda_i + \sigma^2)d + 2\sigma^2\lambda_i)/(\alpha_i^2\lambda_i - \sigma^2)\} < k < d < 1$, $i = 1, \ldots, r$ *if* $\sigma^2 - \alpha_i^2\lambda_i < 0$ *for all* $i = 1, \ldots, r$.

3.2. $\widehat{\beta}_r(k,d)$ versus $\widehat{\beta}_r(k)$. From the definition of the $\widehat{\beta}_r(k)$, we know that let $d = 0$ in $\widehat{\beta}_r(k,d)$, and we obtain the $\widehat{\beta}_r(k)$.

Theorem 2. *Let $k\alpha_i^2 - \sigma^2 > 0$ for all $i = 1, \ldots, r$. Then, there exists a strictly positive d such that $\widehat{\beta}_r(k,d)$ is superior to $\widehat{\beta}_r(k)$ in the mean square error criterion for $0 < d < (\sum_{i=1}^{r}((k\alpha_i^2 - \sigma^2)/(\lambda_i + k)^2))/(\sum_{i=1}^{r}((\sigma^2 + \alpha_i^2\lambda_i)/(\lambda_i(\lambda_i + k)^2))) < 1$.*

Proof. We know that $M(\widehat{\beta}_r(k)) = M\widehat{\beta}_r(k,0)$, so that by continuity it is sufficient to show that $M(\widehat{\beta}_r(k,d))$ decreasing in the neighborhood of $d = 0$.

Performing the calculus for fixed k, we can see that

$$\frac{\partial M\left(\widehat{\beta}_r(k,d)\right)}{\partial d} = 2\sum_{i=1}^{r}\frac{\sigma^2(\lambda_i + d) - (k-d)\alpha_i^2\lambda_i}{\lambda_i(\lambda_i + k)^2}. \tag{29}$$

So when $0 < d < (\sum_{i=1}^{r}((k\alpha_i^2 - \sigma^2)/(\lambda_i + k)^2))/(\sum_{i=1}^{r}((\sigma^2 + \alpha_i^2\lambda_i)/(\lambda_i(\lambda_i + k)^2))) < 1$, $(\partial M(\widehat{\beta}_r(k,d)))/(\partial d) < 0$; that is to say, $M(\widehat{\beta}_r(k,d)) < M(\widehat{\beta}_r(k,0)) = M(\widehat{\beta}_r(k))$. The proof of Theorem 2 is completed. □

3.3. $\widehat{\beta}_r(k,d)$ versus $\widehat{\beta}_r(d)$. From the definition of the $\widehat{\beta}_r(d)$, we know that let $k = 1$ in $\widehat{\beta}_r(k,d)$, and we obtain that $\widehat{\beta}_r(d)$,

$$M\left(\widehat{\beta}_r(d)\right) = \sum_{i=1}^{r}\frac{\sigma^2(\lambda_i + d)^2 + (1-d)^2\alpha_i^2\lambda_i}{\lambda_i(\lambda_i + 1)^2} + \sum_{i=r+1}^{p}\alpha_i^2. \tag{30}$$

Now, we discuss the following difference:

$$\Delta = M\left(\widehat{\beta}_r(k,d)\right) - M\left(\widehat{\beta}_r(k)\right)$$
$$= \sum_{i=1}^{r}\frac{\sigma^2(\lambda_i + d)^2 + (k-d)^2\alpha_i^2\lambda_i}{\lambda_i(\lambda_i + k)^2} + \sum_{i=r+1}^{p}\alpha_i^2$$
$$- \left\{\sum_{i=1}^{r}\frac{\sigma^2(\lambda_i + d)^2 + (1-d)^2\alpha_i^2\lambda_i}{\lambda_i(\lambda_i + 1)^2} + \sum_{i=r+1}^{p}\alpha_i^2\right\}$$
$$= \sum_{i=1}^{r}\frac{\sigma^2(1-k)(\lambda_i + d)(\lambda_i + d)(2\lambda_i + k + 1)}{\lambda_i(\lambda_i + 1)^2(\lambda_i + k)^2}$$
$$- \sum_{i=1}^{r}\left(\left(\sigma^2(1-k)(\lambda_i + d)\alpha_i^2\lambda_i\right.\right.$$
$$\times\left[(k + 1 - 2d)\lambda_i + 2k - d - dk\right]\bigg)$$
$$\times\left(\lambda_i(\lambda_i + 1)^2(\lambda_i + k)^2\right)^{-1}\bigg).$$

(31)

Then by (31), if we want $\Delta < 0$, then

$$d < d_1$$

$$= \sum_{i=1}^{r} \frac{(2\lambda_i + k + 1)\sigma^2 + 2\alpha_i^2\lambda_i^2 + (1+k)\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+1)^2(\lambda_i+k)^2}$$

$$\times \left(\sum_{i=1}^{r} \frac{\lambda_i\sigma^2(2\lambda_i+k+1) - \alpha_i^2\lambda_i^2(k+1) - 2k\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+1)^2(\lambda_i+k)^2} \right)^{-1} \tag{32}$$

when $k < 1$, it is easy to know that d_1 is always less than 1, and d_1 is to be bigger than 0, if

$$\sum_{i=1}^{r} \frac{\lambda_i\sigma^2(2\lambda_i+k+1) - \alpha_i^2\lambda_i^2(k+1) - 2k\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+1)^2(\lambda_i+k)^2} > 0. \tag{33}$$

Then, we get

$$k > \frac{(2\lambda_i+1)\sigma^2 - \alpha_i^2\lambda_i}{\alpha_i^2(\lambda_i+2) - \sigma^2} \tag{34}$$

for $\alpha_i^2(\lambda_i+2) - \sigma^2 > 0$ with all $i = 1, \ldots, r$.

Thus, we obtain the following theorem.

Theorem 3. *If $\alpha_i^2(\lambda_i+2) - \sigma^2 > 0$ and $k > \{\max(((2\lambda_i+1)\sigma^2 - \alpha_i^2\lambda_i)/(\alpha_i^2(\lambda_i+2) - \sigma^2)), 0\}$ for all $i = 1, \ldots, r$, then the $\widehat{\beta}_r(k, d)$ is better than the $\widehat{\beta}_r(d)$ in the mean square error sense for $k < 1$ and*

$$d < \sum_{i=1}^{r} \frac{(2\lambda_i+k+1)\sigma^2 + 2\alpha_i^2\lambda_i^2 + (1+k)\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+1)^2(\lambda_i+k)^2}$$

$$\times \left(\sum_{i=1}^{r} \frac{\lambda_i\sigma^2(2\lambda_i+k+1) - \alpha_i^2\lambda_i^2(k+1) - 2k\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+1)^2(\lambda_i+k)^2} \right)^{-1}. \tag{35}$$

3.4. $\widehat{\beta}_r(k, d)$ versus $\widehat{\beta}$. In this subsection, we will give the comparison of the $\widehat{\beta}_r(k, d)$ and $\widehat{\beta}$ under the mean square error criterion.

Let $k = d$ and $r = p$ in (26); we obtain the MSE of $\widehat{\beta}_r$ as follows:

$$M\left(\widehat{\beta}\right) = \sum_{i=1}^{p} \frac{\sigma^2}{\lambda_i}. \tag{36}$$

In order to compare the $\widehat{\beta}_r(k, d)$ and $\widehat{\beta}$, now we consider the following difference:

$$M\left(\widehat{\beta}_r(k, d)\right) - M\left(\widehat{\beta}\right)$$

$$= \sum_{i=1}^{r} \frac{\sigma^2(\lambda_i+d)^2 + (k-d)^2\alpha_i^2\lambda_i}{\lambda_i(\lambda_i+k)^2}$$

$$+ \sum_{i=r+1}^{p} \alpha_i^2 - \sum_{i=1}^{p} \frac{\sigma^2}{\lambda_i}$$

$$= \sum_{i=1}^{r} (d-k) \tag{37}$$

$$\times \frac{\left(\alpha_i^2\lambda_i + \sigma^2\right)d + 2\sigma^2\lambda_i + \left(\sigma^2 - \alpha_i^2\lambda_i\right)k}{\lambda_i(\lambda_i+k)^2}$$

$$+ \sum_{i=r+1}^{p} \frac{\alpha_i^2\lambda_i - \sigma^2}{\lambda_i}$$

when $k > d > d_2 = (\sum_{r=1}^{r}(((\alpha_i^2\lambda_i - \sigma^2)k - 2\sigma^2\lambda_i)/(\lambda_i(\lambda_i + k)^2)))/(\sum_{r=1}^{r}((\alpha_i^2\lambda_i + \sigma^2)/(\lambda_i(\lambda_i + k)^2)))$, then

$$\sum_{i=1}^{r} (d-k) \frac{\left(\alpha_i^2\lambda_i + \sigma^2\right)d + 2\sigma^2\lambda_i + \left(\sigma^2 - \alpha_i^2\lambda_i\right)k}{\lambda_i(\lambda_i+k)^2} < 0. \tag{38}$$

Since the lower bound of d is less than 1, the lower bound of d may be less than 0. If the lower bound of of d is less than 0, then we can choose any d in $[0, 1]$. Thus, we can get the following theorem.

Theorem 4.

(1) *If $\alpha_i^2\lambda_i - \sigma^2 < 0$ for all $i = 1, 2, \ldots, p$, then $\widehat{\beta}_r(k, d)$ is superior to $\widehat{\beta}$ in the mean square error sense for any $k > 0$ and $0 < d < 1$.*

(2) *If $\alpha_i^2\lambda_i - \sigma^2 < 0$ for some $i = 1, 2, \ldots, r$, then $\widehat{\beta}_r(k, d)$ is superior to $\widehat{\beta}$ in the mean square error sense for $k > d > (\sum_{r=1}^{r}(((\alpha_i^2\lambda_i - \sigma^2)k - 2\sigma^2\lambda_i)/(\lambda_i(\lambda_i + k)^2)))/(\sum_{r=1}^{r}((\alpha_i^2\lambda_i + \sigma^2)/(\lambda_i(\lambda_i + k)^2)))$ when $\sum_{i=r+1}^{p}((\alpha_i^2\lambda_i - \sigma^2)/\lambda_i) < 0$.*

3.5. $\widehat{\beta}_r(k, d)$ versus $\widehat{\beta}(k, d)$. In this subsection, we will compare $\widehat{\beta}_r(k, d)$ with $\widehat{\beta}(k, d)$ in the mean square error sense.

Define $v_i = (\lambda_i + d)/(\lambda_i + k)$, then the mean square error of $\widehat{\beta}_r(k, d)$ can be written as

$$M\left(\widehat{\beta}_r(k, d)\right) = \sum_{1=1}^{r} \left[\frac{\sigma^2 v_i^2}{\lambda_i} + (v_i - 1)^2\alpha_i^2 \right] + \sum_{i=r+1}^{p} \alpha_i^2. \tag{39}$$

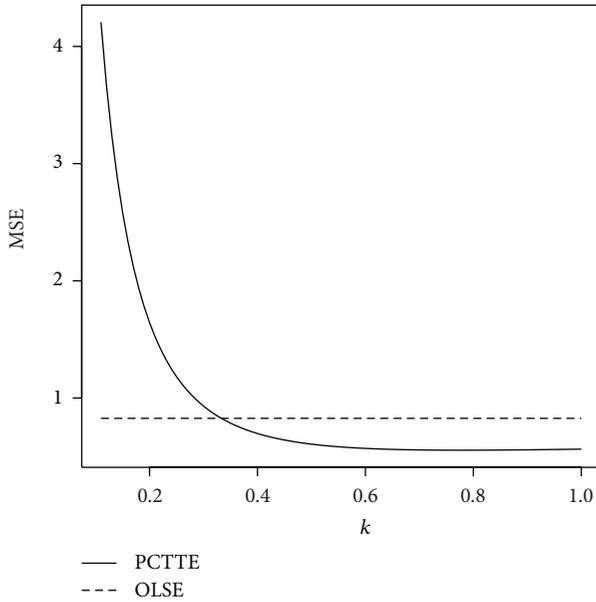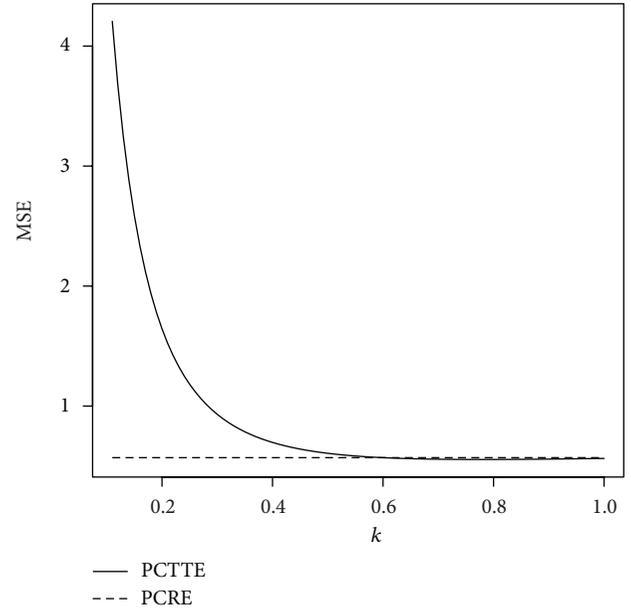FIGURE 1: The MSE values of OLSE and PCTTE when $d = 0.6$.

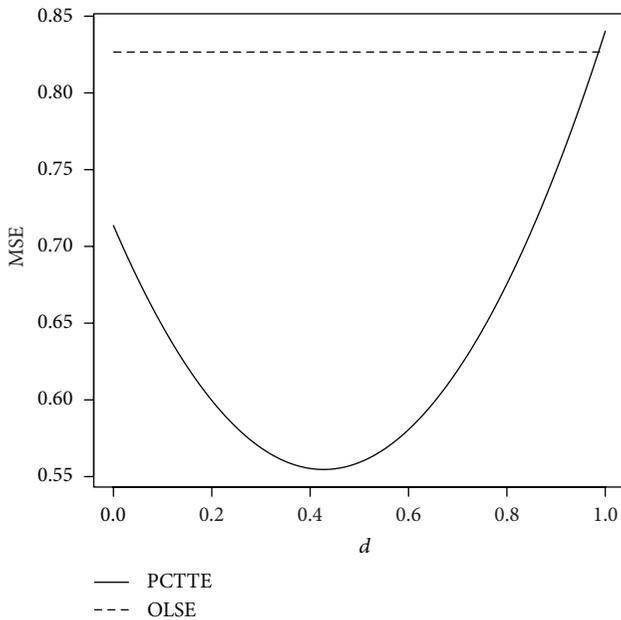FIGURE 2: The MSE values of OLSE and PCTTE when $k = 0.56$.

FIGURE 3: The MSE values of PCRE and PCTTE when $d = 0.6$.

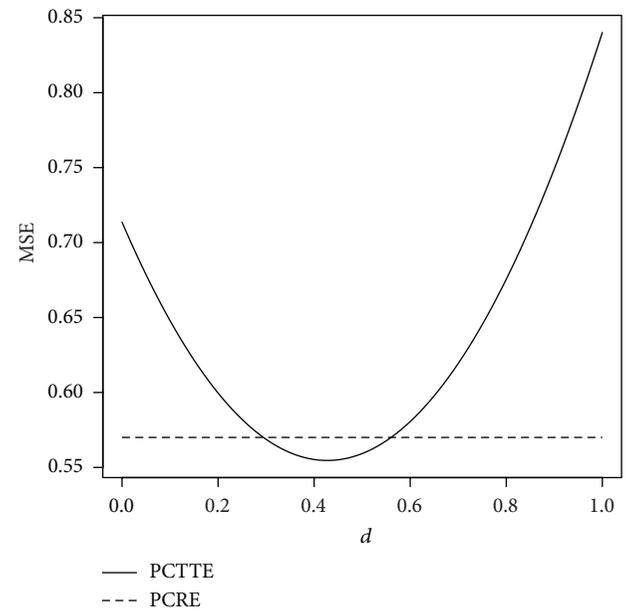FIGURE 4: The MSE values of PCRE and PCTTE when $k = 0.56$.

Firstly, we discuss the difference between $M(\widehat{\beta}_r(k, d))$ and $M(\widehat{\beta}_{r-1}(k, d))$:

$$
\begin{aligned}
& M\left(\widehat{\beta}_r(k, d)\right) - M\left(\widehat{\beta}_{r-1}(k, d)\right) \\
& = v_r\left[v_r\left(\frac{\sigma^2}{\lambda_r} + \alpha_r^2\right) - 2\alpha_r^2\right] \\
& = \frac{v_r}{\lambda_i + k}\left[d\left(\frac{\sigma^2}{\lambda_r} + \alpha_r^2\right) + \sigma^2 - \alpha_r^2\lambda_r - 2k\alpha_r^2\right].
\end{aligned}
\tag{40}
$$

So when $\sigma^2 < \alpha_r^2\lambda_r + 2k\alpha_r^2$, if $1 > d > ((\alpha_r^2\lambda_r + 2k\alpha_r^2 - \sigma^2)/((\sigma^2/\lambda_r)+\alpha_r^2)) > 0$, then $M(\widehat{\beta}_r(k, d)) - M(\widehat{\beta}_{r-1}(k, d)) > 0$.

If $0 < d < ((\alpha_r^2\lambda_r + 2k\alpha_r^2 - \sigma^2)/((\sigma^2/\lambda_r) + \alpha_r^2)) < 1$, then $M(\widehat{\beta}_r(k, d)) - M(\widehat{\beta}_{r-1}(k, d)) < 0$.

Then we have the following theorem.

Theorem 5.

(1) If $\sigma^2 < \alpha_i^2\lambda_i + 2k\alpha_i^2$ for some $i \in N_p$, then there exists a nonnegative d_3 such that $M(\widehat{\beta}(k, d)) > M(\widehat{\beta}_r(k, d))$ for $0 < d_3 < d < 1$, where $N_p = r + 1, \ldots, p$.
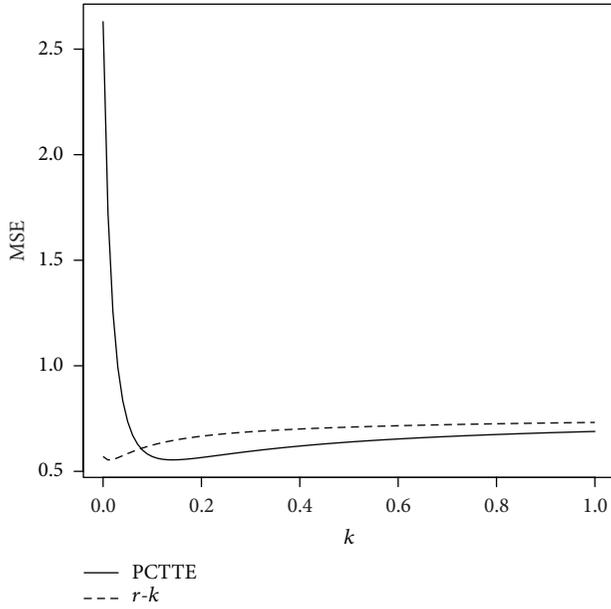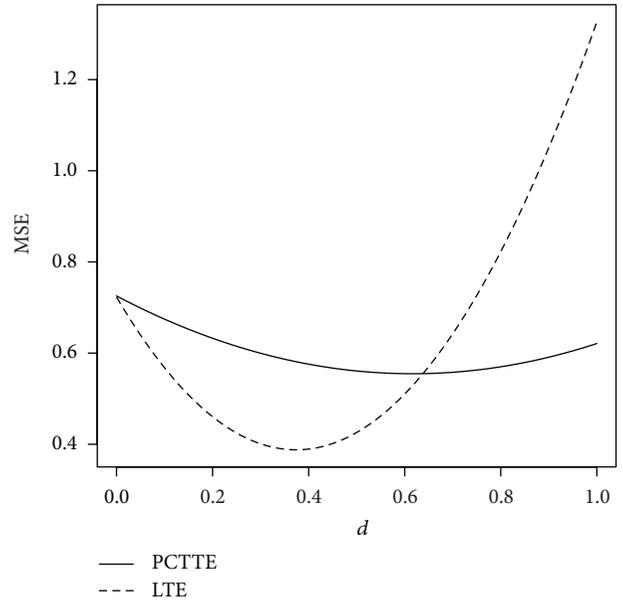
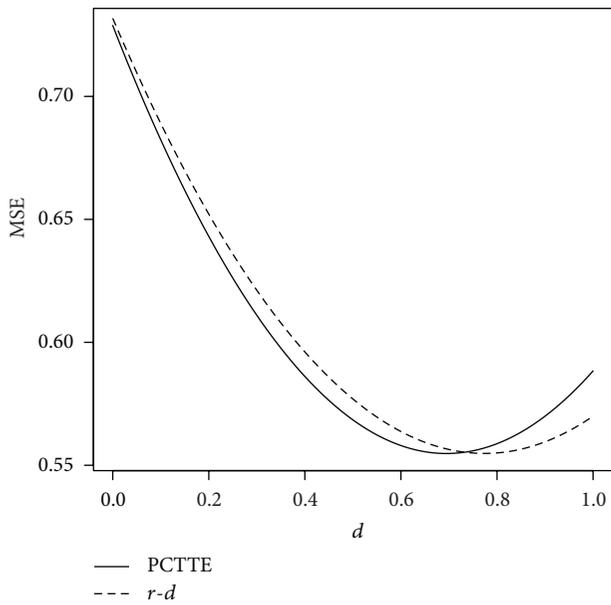FIGURE 5: The MSE values of r-k class estimator and PCTTE when $d = 0.1$.

FIGURE 6: The MSE values of r-d class estimator and PCTTE when $k = 0.9$.

(2) If $\sigma^2 < \alpha_i^2 \lambda_i + 2k\alpha_i^2$ for all $i \in N_p$, then there exists a nonnegative d_3 such that $M(\widehat{\beta}(k, d)) < M(\widehat{\beta}_r(k, d))$ for $0 < d < d_4 < 1$, where $N_p = r + 1, \ldots, p$, where

$$d_3 = \min_{i \in N_p} \left\{ 1, \max_{i \in N_p} \frac{\alpha_i^2 \lambda_i + 2k\alpha_i^2 - \sigma^2}{(\sigma^2 / \lambda_i) + \alpha_i^2} \right\},$$

$$d_4 = \min_{i \in N_p} \frac{\alpha_i^2 \lambda_i + 2k\alpha_i^2 - \sigma^2}{(\sigma^2 / \lambda_i) + \alpha_i^2}. \tag{41}$$

FIGURE 7: The MSE values of LTE and PCTTE when $k = 0.8$.

4. Numerical Example

To illustrate our theoretical results, firstly we use a numerical example to investigate the estimators studied in the dataset originally due to Gruber [13], and later considered by Akdeniz and Erol [14]. We assemble the data as follows:

$$X = \begin{pmatrix} 1.9 & 2.2 & 1.9 & 3.7 \\ 1.8 & 2.2 & 2.0 & 3.8 \\ 1.8 & 2.4 & 2.1 & 3.6 \\ 1.8 & 2.4 & 2.2 & 3.8 \\ 2.0 & 2.5 & 2.3 & 3.8 \\ 2.1 & 2.6 & 2.4 & 3.7 \\ 2.1 & 2.6 & 2.6 & 3.8 \\ 2.2 & 2.6 & 2.6 & 4.0 \\ 2.3 & 2.8 & 2.8 & 3.7 \\ 2.3 & 2.7 & 2.8 & 3.8 \end{pmatrix},$$

$$\tag{42}$$

$$y = \begin{pmatrix} 2.3 \\ 2.2 \\ 2.2 \\ 2.3 \\ 2.4 \\ 2.5 \\ 2.6 \\ 2.6 \\ 2.7 \\ 2.7 \end{pmatrix}.$$

Firstly, we can compute the OLSE of β as follows:

$$\widehat{\beta} = (0.6921, 0.6258, -0.1154, 0.2866, 0.0256)'. \tag{43}$$
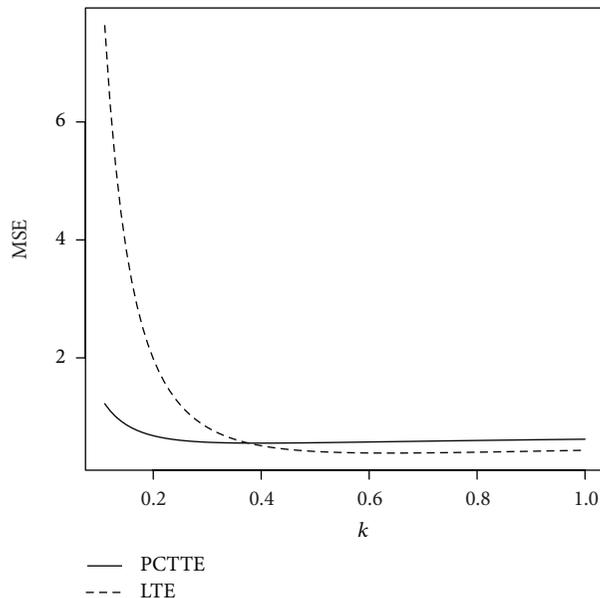
FIGURE 8: The MSE values of LTE and PCTTE when $d = 0.3$.

For the OLSE, PCRE, r-k class estimator, r-d class estimator, Liu-type estimator (LTE), and new estimator (PCTTE), their estimated mean square error (MSE) values are obtained by replacing all unknown model parameters by their, respectively, least squares estimators in corresponding expressions.

Firstly, we see the comparison of the OLSE and the PCTTE. When $d = 0.6$ is fixed, then if the values of k is big then the new estimator has smaller MSE values than OLSE; that is to say, the new estimator is better than the OLSE. So we see that the new estimator improved the OLSE.

From Figure 3, we see that when d, fixed, if $0 < d < k$, then the new estimator is better than the PCRE. In Theorem 1, we see that if $0 < d < k$, the new estimator is better. For the numerical example, when $0 < d < k$, the new estimator is more efficient than the PCRE.

By Figures 1, 2, 3, 4, 5, 6, 7, and 8, we see that when $0 < d < k$, then new estimator is better than the LTE. So in practice, we can choose bigger k and smaller d.

5. Conclusion

In this paper, we use a new method to propose the principal component Liu-type estimator. Then, we discuss the superiority of the new estimator with the OLSE, PCRE, r-k class estimator, r-d class estimator, and Liu-type estimator in the sense of mean square error. We also give a method to choose the biasing parameters. Finally, we give a numerical example to illustrate the performance the various estimators.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] W. F. Massy, "Principal components regression in exploratory statistical reseach," *Journal of the American Statistical Association*, vol. 60, pp. 234–256, 1965.

[2] A. E. Hoerl and R. W. Kennard, "Ridge regression: biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[3] K. J. Liu, "A new class of biased estimate in linear regression," *Communications in Statistics*, vol. 22, no. 2, pp. 393–402, 1993.

[4] W. H. Huang, J. J. Qi, and N. T. Huang, "Liu-type estimation for a linear regression model with linear restrictions," *Journal of Systems Science and Mathematical Sciences*, vol. 29, no. 7, pp. 937–946, 2009.

[5] S. Toker and S. Kaçıranlar, "On the performance of two parameter ridge estimator under the mean square error criterion," *Applied Mathematics and Computation*, vol. 219, no. 9, pp. 4718–4728, 2013.

[6] M. R. Baye and D. F. Parker, "Combining ridge and principal component regression: a money demand illustration," *Communications in Statistics*, vol. 13, no. 2, pp. 197–205, 1984.

[7] S. Kaçıranlar and S. Sakallıoğlu, "Combining the Liu estimator and the principal component regression estimator," *Communications in Statistics*, vol. 30, no. 12, pp. 2699–2705, 2001.

[8] M. I. Alheety and B. M. G. Kibria, "Modified Liu-type estimator based on (r-k) class estimator," *Communications in Statistics*, vol. 42, no. 2, pp. 304–319, 2013.

[9] J. Xu and H. Yang, "On the restricted almost unbiased estimators in linear regression," *Journal of Applied Statistics*, vol. 38, no. 3, pp. 605–617, 2011.

[10] Y. Li and H. Yang, "Two kinds of restricted modified estimators in linear regression model," *Journal of Applied Statistics*, vol. 38, no. 7, pp. 1447–1454, 2011.

[11] J. Durbin, "A note on regression when there is extraneous information that one of coecients," *Journal of the American Statistical Association*, vol. 48, pp. 799–808, 1953.

[12] J. B. Wu, *Research on the properties of parameter estimation in linear regression model [Ph.D. papers]*, Chongqing University, Chongqing, China, 2013.

[13] M. H. J. Gruber, *Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimators*, vol. 156, Marcel Dekker, New York, NY, USA, 1998.

[14] F. Akdeniz and H. Erol, "Mean squared error matrix comparisons of some biased estimators in linear regression," *Communications in Statistics*, vol. 32, no. 12, pp. 2389–2413, 2003.