



OPEN

## On the mixed Kibria–Lukman estimator for the linear regression model

Hongmei Chen<sup>1</sup> & Jibo Wu<sup>2</sup>✉

This paper considers a linear regression model with stochastic restrictions, we propose a new mixed Kibria–Lukman estimator by combining the mixed estimator and the Kibria–Lukman estimator. This new estimator is a general estimation, including OLS estimator, mixed estimator and Kibria–Lukman estimator as special cases. In addition, we discuss the advantages of the new estimator based on MSEM criterion, and illustrate the theoretical results through examples and simulation analysis.

Consider the following linear regression model:

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

where  $y$  is the response variable vector of  $n \times 1$ ,  $X$  is the column full rank independent variables matrix of  $n \times (p + 1)$ ,  $\beta$  is the unknown coefficient vector of  $p \times 1$ ,  $\varepsilon$  is the random error vector of  $n$  dimension such that  $E(\varepsilon) = 0$  and  $\text{Cov}(\varepsilon) = \sigma^2 I$ , where  $\sigma^2 > 0$  is mean squared error.

In the estimation of unknown coefficient vector  $\beta$ , the OLS estimator is the most commonly used:

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'y \quad (2)$$

It is easy to know from formula (2),  $E\hat{\beta} = \beta$ , and the OLS estimator has been widely used because of its unbiased nature and concise form. However, the ill condition of the design matrix  $X$  caused by the increasing number of dependent predictors often makes the OLS estimates unstable.

Massy<sup>1</sup> proposed principal component estimator. Hoerl and Kennard<sup>2</sup> obtained the ridge estimation by introducing a ridge parameter  $k$  into the design  $X'X$  matrix calculation. Swindel<sup>3</sup> proposed a modified ridge estimator with prior information while Lukman et al.<sup>4</sup> proposed the two-parameter form of the ridge estimator called the modified ridge estimator (MRT). Liu<sup>5</sup> obtained a linearized form of the ridge estimator called the Liu estimator. Akdeniz and Kaciranlara<sup>6</sup> proposed the generalized Liu estimator. Liu<sup>7</sup> obtained a two-parameter form of the Liu estimator.

Many scholars have found that a new estimator can be obtained by combining the two estimators, which generally have good statistical properties. Baye and Parker<sup>8</sup> proposed  $r$ - $k$  estimator by combining ridge estimator and principal component estimator. Kaciranlar and Sakalloglu<sup>9</sup> proposed  $r$ - $d$  estimator by combining Liu estimator and principal component estimator. Ozkale and Kaciranlar<sup>10</sup> proposed two parameter estimator by combining the James–Stein Shrinkage estimator and the modified ridge estimator proposed by Swindel. Batah et al.<sup>11</sup> proposed a modified  $r$ - $k$  estimator combining unbiased ridge estimator and principal component estimator. Yang and Chang<sup>12</sup> proposed another two parameter estimator based on ridge estimator and Liu estimator. Lukman et al.<sup>13</sup> proposed a new estimator by combining modified ridge estimator (MRT) and principal component estimator. Kibria and Lukman<sup>14</sup> proposed Kibria–Lukman estimator by combining ridge estimator and Liu estimator.

In practice, in addition to the sample information given by model (1), additional information about parameters in the sample information, such as certain deterministic or stochastic restrictions on unknown parameters, can also be considered. This method can also overcome the complex collinearity problem. Theil and Goldberger<sup>15</sup> and Theil<sup>16</sup> proposed mixed estimator by comprehensively considering sample information and constraints. Schiffrin and Toutenburg<sup>17</sup> proposed weighted mixed estimator for the different importance of sample information and prior information.

In recent years, biased estimation and estimation methods with prior information are often combined to form a broader biased estimation. Hubert and Wijekoon<sup>18</sup> proposed a stochastic restricted Liu estimator by combining Liu estimator and mixed estimator. Yang and Xu<sup>19</sup> obtained another stochastic mixed Liu estimator. In the same year, Yang and Chang further studied the stochastic mixed Liu estimator and obtained the weighted mixed Liu

<sup>1</sup>College of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing, China. <sup>2</sup>School of Mathematics and Big Data, Chongqing University of Arts and Sciences, Chongqing, China. ✉email: linfen52@126.com

estimator. Yang and Li<sup>12</sup> proposed another stochastic mixed ridge estimator. Ozbay and Kaciranlar<sup>20</sup> integrated two parameter estimator and mixed estimator and proposed a two parameter mixed estimator.

In this paper, a new mixed KL estimator under stochastic restrictions is proposed, and its excellent properties under certain conditions are proved theoretically. The above theoretical results are verified and analyzed by examples and data simulation.

### The proposed estimator

Hoerl and Kennard<sup>2</sup> proposed the ridge estimator (RE):

$$\hat{\beta}_{RE} = (X'X + kI)^{-1}X'y \quad (3)$$

where  $k > 0$  is the parameter. In fact, ridge estimator is obtained by solving the following extreme value problem:

$$(y - X\beta)'(y - X\beta) + k(\beta'\beta - c)$$

where  $c$  is constant,  $k$  is the Lagrange constant.

Kibria and Lukman<sup>14</sup> proposed the Kibria Lukman (KL) estimator:

$$\hat{\beta}_{KL} = (X'X + kI)^{-1}(X'y - k\hat{\beta}) \quad (4)$$

where  $k > 0$  is the parameter. KL estimator is obtained by solving the following extreme value problem:

$$(y - X\beta)'(y - X\beta) + k[(\beta + \hat{\beta})'(\beta + \hat{\beta}) - c] \quad (5)$$

where  $c$  is constant,  $k$  is the Lagrange constant.

Consider the following stochastic restrictions:

$$r = R\beta + e, e \sim (0, \sigma^2\psi), \quad (6)$$

where  $r$  is the known random vector of  $j \times 1$ ,  $R$  is the row full rank sample data matrix of  $j \times p$ , let  $e$  be the  $j \times 1$  random error vector and independent of each other, and  $\psi$  be the known positive definite matrix.

Theil and Goldberger<sup>15</sup> and Theil<sup>16</sup> proposed the mixed estimator by integrating sample information and constraints. The derivation idea is to rewrite models (1) and (6) into a new linear model:

$$\begin{pmatrix} y \\ r \end{pmatrix} = \begin{pmatrix} X \\ R \end{pmatrix} \beta + \begin{pmatrix} \varepsilon \\ e \end{pmatrix}$$

If  $\tilde{y} = \begin{pmatrix} y \\ r \end{pmatrix}$ ,  $\tilde{X} = \begin{pmatrix} X \\ R \end{pmatrix}$ ,  $\tilde{\varepsilon} = \begin{pmatrix} \varepsilon \\ e \end{pmatrix}$ , above model is transformed into

$$\tilde{y} = \tilde{X}\beta + \tilde{\varepsilon} \quad (7)$$

By applying the least square estimator to the new linear model (7), the mixed estimator (ME) of parameter  $\beta$  is obtained:

$$\hat{\beta}_{ME} = (X'X + R'\psi^{-1}R)^{-1}(X'y + R'\psi^{-1}r) \quad (8)$$

Combined mixed estimator and ridge estimator and proposed stochastic mixed ridge estimation (RME):

$$\hat{\beta}_{MRE} = (X'X + kI + R'\psi^{-1}R)^{-1}(X'y + R'\psi^{-1}r) \quad (9)$$

The estimator proposed in this paper is obtained by solving the following extreme value problem:

$$\Phi^* = (y - X\beta)'(y - X\beta) + k[(\beta - d\hat{\beta})'(\beta - d\hat{\beta}) - c] + (r - R\beta)'\psi^{-1}(r - R\beta) \quad (10)$$

where  $c$  is constant,  $k$  is Lagrange constant.

Regular equations can be obtained:

$$X'X\beta - X'y + k(\beta - d\hat{\beta}) + R'\psi^{-1}R - R'\psi^{-1}r = 0 \quad (11)$$

$$(\beta - d\hat{\beta})'(\beta - d\hat{\beta}) = c \quad (12)$$

from Eqs. (11) and (12), we can get the mixed KL estimator:

$$\hat{\beta}_{MKL} = (X'X + kI + R'\psi^{-1}R)^{-1}(X'y - k\hat{\beta} + R'\psi^{-1}r), k > 0 \quad (13)$$

It can be seen from Eq. (13) that mixed estimator, KL estimator and OLS estimator can be regarded as special cases of mixed KL estimator. Namely

When  $k = 0$ ,  $\hat{\beta}_{ME} = \hat{\beta}_{MKL} = (X'X + R'\psi^{-1}R)^{-1}(X'y + R'\psi^{-1}r)$  is mixed estimator;

When  $R = 0$ ,  $\hat{\beta}_{KL} = \hat{\beta}_{MKL} = (X'X + kI)^{-1}(X'y - k\hat{\beta})$  is KL estimator;

When  $k = 0, R = 0, \hat{\beta}_{OLS} = \hat{\beta}_{MKL} = (X'X)^{-1}X'y$  is OLS estimator.

### The performance of the new estimator

If  $\hat{\beta}$  is the estimation of  $\beta$ , then the mean square error matrix (MSEM) of  $\hat{\beta}$  is given as:

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

where  $Cov(\hat{\beta})$  is the covariance matrix of  $\hat{\beta}$ , and  $Bias(\hat{\beta}) = E(\hat{\beta}) - \beta$  is the deviation vector. Two estimates  $\hat{\beta}_1$  and  $\hat{\beta}_2, \hat{\beta}_2$  are better than  $\hat{\beta}_1$  under MSEM criterion if and only if:

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

**Lemma 3.1** Suppose two  $n \times n$  matrix  $M > 0, N \geq 0$ , then  $M > N \Leftrightarrow \lambda_1(NM^{-1}) < 1$ , where  $\lambda_1(NM^{-1})$  is the maximum eigenvalue of matrix  $NM^{-1}$ .

The mean square error matrix of mixed KL estimator  $\hat{\beta}_{MKL}$  is calculated as follows:

$$\begin{aligned} E(\hat{\beta}_{MKL}) &= E[(X'X + kI + R'\psi^{-1}R)^{-1}(X'y - k\hat{\beta} + R'\psi^{-1}r)] \\ &= A_k E(X'y - k\hat{\beta} + R'\psi^{-1}r) \\ &= A_k E(X'y + k\hat{\beta} - 2k\hat{\beta} + R'\psi^{-1}r) \\ &= A_k(A_k^{-1} - 2k)\beta \\ &= \beta - 2kA_k\beta \end{aligned} \tag{14}$$

where  $A_k = (X'X + kI + R'\psi^{-1}R)^{-1}$ .

Deviation vector:  $Bias(\hat{\beta}_{MKL}) = E(\hat{\beta}_{MKL}) - \beta = -2kA_k\beta$ .

$$\begin{aligned} Cov(\hat{\beta}_{MKL}) &= Cov[(X'X + kI + R'\psi^{-1}R)^{-1}(X'y - k\hat{\beta} + R'\psi^{-1}r)] \\ &= Cov[A_k(X'y - k\hat{\beta} + R'\psi^{-1}r)] \\ &= A_k Cov(X'y - k\hat{\beta} + R'\psi^{-1}r)A_k \\ &= A_k(\sigma^2 X'X - k\sigma^2 S^{-1} + \sigma^2 R'\psi^{-1}R)A_k \\ &= \sigma^2 A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k \end{aligned} \tag{15}$$

Therefore,

$$\begin{aligned} MSEM(\hat{\beta}_{MKL}) &= Cov(\hat{\beta}_{MKL}) + Bias(\hat{\beta}_{MKL})Bias(\hat{\beta}_{MKL})' \\ &= \sigma^2 A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k + 4k^2 A_k \beta \beta' A_k \\ &= \sigma^2 A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k + b_1 b_1' \end{aligned} \tag{16}$$

where  $b_1 = -2kA_k\beta$ .

By substituting  $k = 0$  into Eq. (16), the mean square error matrix of the mixed estimator can be obtained:

$$\begin{aligned} MSEM(\hat{\beta}_{ME}) &= \sigma^2 (X'X + R'\psi^{-1}R)^{-1} (X'X + R'\psi^{-1}R) (X'X + R'\psi^{-1}R)^{-1} \\ &= \sigma^2 (X'X + R'\psi^{-1}R)^{-1} \\ &= \sigma^2 M^{-1} \end{aligned} \tag{17}$$

where  $M = X'X + R'\psi^{-1}R$ .

By substituting  $R = 0$  into Eq. (16), the mean square error matrix of the KL estimator can be obtained:

$$\begin{aligned}
 \text{MSEM}(\hat{\beta}_{KL}) &= \sigma^2(X'X + kI)^{-1}(X'X - kS^{-1})(X'X + kI)^{-1} \\
 &\quad + 4k^2(X'X + kI)^{-1}\beta\beta'(X'X + kI)^{-1} \\
 &= \sigma^2S_k^{-1}(X'X - kS^{-1})S_k^{-1} + 4k^2S_k^{-1}\beta\beta'S_k^{-1} \\
 &= \sigma^2S_k^{-1}(X'X - kS^{-1})S_k^{-1} + b_2b_2'
 \end{aligned}
 \tag{18}$$

where  $S_k = X'X + kI, b_2 = -2kS_k^{-1}\beta$ .

By substituting  $k = 0, R = 0$  into Eq. (16), the mean square error matrix of the OLS estimator can be obtained:

$$\text{MSEM}(\hat{\beta}_{OLS}) = \sigma^2S^{-1}
 \tag{19}$$

Mean square error matrix of mixed ridge estimator:

$$\begin{aligned}
 E(\hat{\beta}_{MRE}) &= E[(X'X + kI + R'\psi^{-1}R)^{-1}(X'y + R'\psi^{-1}r)] \\
 &= A_kE(X'y + R'\psi^{-1}r) \\
 &= A_kE(X'y + k\hat{\beta} - k\hat{\beta} + R'\psi^{-1}r) \\
 &= A_k(A_k^{-1} - kI)\beta \\
 &= \beta - kA_k\beta
 \end{aligned}
 \tag{20}$$

Deviation vector: Bias( $\hat{\beta}_{MRE}$ ) =  $E(\hat{\beta}_{MRE}) - \beta = -kA_k\beta$ .

$$\begin{aligned}
 \text{Cov}(\hat{\beta}_{MRE}) &= \text{Cov}[(X'X + kI + R'\psi^{-1}R)^{-1}(X'y + R'\psi^{-1}r)] \\
 &= \text{Cov}[A_k(X'y + R'\psi^{-1}r)] \\
 &= A_k\text{Cov}(X'y + R'\psi^{-1}r)A_k \\
 &= A_k(\sigma^2X'X + \sigma^2R'\psi^{-1}R)A_k \\
 &= \sigma^2A_k(X'X + R'\psi^{-1}R)A_k
 \end{aligned}$$

Therefore,

$$\text{MSEM}(\hat{\beta}_{MRE}) = \sigma^2A_k(X'X + R'\psi^{-1}R)A_k + k^2A_k\beta\beta'A_k.
 \tag{21}$$

**Comparison between mixed KL estimator and mixed estimator.** From Eqs. (16) and (17), we make

$$\begin{aligned}
 \Delta_1 &= \text{MSEM}(\hat{\beta}_{ME}) - \text{MSEM}(\hat{\beta}_{MKL}) \\
 &= \sigma^2M^{-1} - \sigma^2A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k - b_1b_1' \\
 &= \sigma^2M^{-1} - \sigma^2A_k(M - kS^{-1})A_k - b_1b_1' \\
 &= \sigma^2[M^{-1} - A_k(M - kS^{-1})A_k] - b_1b_1'
 \end{aligned}
 \tag{22}$$

Because

$$\begin{aligned}
 M^{-1} - A_k(M - kS^{-1})A_k &= A_kA_k^{-1}M^{-1}A_k^{-1}A_k - A_k(M - kS^{-1})A_k \\
 &= A_k[A_k^{-1}M^{-1}A_k^{-1} - (M - kS^{-1})]A_k \\
 &= A_k[(M + kI)M^{-1}(M + kI) - (M - kS^{-1})]A_k \\
 &= A_k(M + 2kI + k^2M^{-1} - M + kS^{-1})A_k \\
 &= A_k(2kI + k^2M^{-1} + kS^{-1})A_k,
 \end{aligned}$$

from  $k > 0$ , so  $M^{-1} - A_k(M - kS^{-1})A_k > 0$ , Theorem 3.2 is obtained.

**Theorem 3.2** The necessary and sufficient conditions for mixed KL estimator  $\hat{\beta}_{MKL}$  to be superior to mixed estimator  $\hat{\beta}_{ME}$  under MSEM criterion are as follows:

$$\sigma^{-2}b_1' [M^{-1} - A_k(M - kS^{-1})A_k]^{-1}b_1 \leq 1 \tag{23}$$

**Comparison between mixed KL estimator and KL estimator.** From Eqs. (16) and (18), we make

$$\begin{aligned} \Delta_2 &= \text{MSEM}(\hat{\beta}_{KL}) - \text{MSEM}(\hat{\beta}_{MKL}) \\ &= \sigma^2 S_k^{-1}(S - kS^{-1})S_k^{-1} + b_2b_2' - \sigma^2 A_k(S - kS^{-1} + R'\psi^{-1}R)A_k - b_1b_1' \\ &= \sigma^2 [S_k^{-1}(S - kS^{-1})S_k^{-1} - A_k(S - kS^{-1} + R'\psi^{-1}R)A_k] + b_2b_2' - b_1b_1' \end{aligned} \tag{24}$$

Because

$$\begin{aligned} &S_k^{-1}(S - kS^{-1})S_k^{-1} - A_k(S - kS^{-1} + R'\psi^{-1}R)A_k \\ &= A_k [A_k^{-1}S_k^{-1}(S - kS^{-1})S_k^{-1}A_k^{-1} - (S - kS^{-1} + R'\psi^{-1}R)]A_k \\ &= A_k [(S_k + R'\psi^{-1}R)S_k^{-1}NS_k^{-1}(S_k + R'\psi^{-1}R) - (N + R'\psi^{-1}R)]A_k \\ &= A_k [(S_k + Q)S_k^{-1}NS_k^{-1}(S_k + Q) - (N + Q)]A_k \\ &= A_k [(I + QS_k^{-1})N(I + S_k^{-1}Q) - (N + Q)]A_k \\ &= A_k [N + NS_k^{-1}Q + QS_k^{-1}N + QS_k^{-1}NS_k^{-1}Q - (N + Q)]A_k \\ &= A_k (NS_k^{-1}Q + QS_k^{-1}N + QS_k^{-1}NS_k^{-1}Q - Q)A_k \\ &= A_k BA_k, \end{aligned}$$

where  $N = S - kS^{-1}$ ,  $Q = R'\psi^{-1}R$ ,  $B = NS_k^{-1}Q + QS_k^{-1}N + QS_k^{-1}NS_k^{-1}Q - Q$

According to the Lemma 3.1, it can be obtained that if  $k < \min_{i=1}^p \lambda_i^2$ , then  $N > 0$ . So  $B > 0$  if and only if  $k < \min_{i=1}^p \lambda_i^2$ ,  $\lambda_1 Q (NS_k^{-1}Q + QS_k^{-1}N + QS_k^{-1}NS_k^{-1}Q)^{-1} < 1$ .

As long as  $k < \min_{i=1}^p \lambda_i^2$ ,  $\lambda_1 Q (NS_k^{-1}Q + QS_k^{-1}N + QS_k^{-1}NS_k^{-1}Q)^{-1} < 1$ , following conclusions can be obtained:

$$\Delta_2 \geq 0 \text{ if and only if } b_1' (\sigma^2 A_k BA_k + b_2 b_2')^{-1} b_1 \leq 1. \text{ Therefore, there is Theorem 3.2.}$$

**Theorem 3.3** When  $k < \min_{i=1}^p \lambda_i^2$ ,  $\lambda_1 Q (NS_k^{-1}Q + QS_k^{-1}N + QS_k^{-1}NS_k^{-1}Q)^{-1} < 1$ , the necessary and sufficient conditions for mixed KL estimator  $\hat{\beta}_{MKL}$  to be superior to KL estimator  $\hat{\beta}_{KL}$  under MSEM criterion are as follows:

$$b_1' (\sigma^2 A_k BA_k + b_2 b_2')^{-1} b_1 \leq 1 \tag{25}$$

**Comparison between mixed KL estimator and OLS estimator.** From Eqs. (16) and (19), we make

$$\begin{aligned} \Delta_3 &= \text{MSEM}(\hat{\beta}_{OLS}) - \text{MSEM}(\hat{\beta}_{MKL}) \\ &= \sigma^2 S^{-1} - \sigma^2 A_k (X'X - kS^{-1} + R'\psi^{-1}R)A_k - b_1b_1' \\ &= \sigma^2 [S^{-1} - A_k (X'X - kS^{-1} + R'\psi^{-1}R)A_k] - b_1b_1' \end{aligned} \tag{26}$$

Because

$$\begin{aligned}
 & S^{-1} - A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k \\
 &= A_k A_k^{-1} S^{-1} A_k^{-1} A_k - A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k \\
 &= A_k [A_k^{-1} S^{-1} A_k^{-1} - (X'X - kS^{-1} + R'\psi^{-1}R)] A_k \\
 &= A_k [(S + kI + Q)S^{-1}(S + kI + Q) - (S - kS^{-1} + Q)] A_k \\
 &= A_k [(I + kS^{-1} + QS^{-1})(S + kI + Q) - (S - kS^{-1} + Q)] A_k \\
 &= A_k [S + kI + Q + (I + kS^{-1} + QS^{-1})(kI + Q) - (S - kS^{-1} + Q)] A_k \\
 &= A_k [kI + kS^{-1} + (I + kS^{-1} + QS^{-1})(kI + Q)] A_k \\
 &= A_k (2kI + kS^{-1} + k^2S^{-1} + Q + kS^{-1}Q + kQS^{-1} + QS^{-1}Q) A_k \\
 &= A_k [2kI + kS^{-1} + k^2S^{-1} + Q + k(S^{-1}Q + QS^{-1}) + QS^{-1}Q] A_k \\
 &= A_k [2kI + kS^{-1} + k^2S^{-1} + Q + kC + QS^{-1}Q] A_k
 \end{aligned}$$

where  $C = S^{-1}Q + QS^{-1}$ .

Because  $C = C'$ , and  $\lambda_i(S^{-1}Q) = \lambda_i(S^{-\frac{1}{2}}QS^{-\frac{1}{2}}) > 0$ , we can get  $C > 0$ , so  $2kI + kS^{-1} + k^2S^{-1} + Q + kC + QS^{-1}Q > 0$ , that is  $S^{-1} - A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k > 0$ , Theorem 3.4 is obtained.

**Theorem 3.4** The necessary and sufficient conditions for mixed KL estimator  $\hat{\beta}_{MKL}$  to be superior to  $\hat{\beta}_{OLS}$  under MSEM criterion are as follows:

$$\sigma^{-2}b_1' [S^{-1} - A_k(X'X - kS^{-1} + R'\psi^{-1}R)A_k] b_1 \leq 1 \tag{27}$$

**Comparison between mixed KL estimator and mixed ridge estimator.** From Eqs. (16) and (22), we make

$$\begin{aligned}
 \Delta_4 &= \text{MSEM}(\beta_{MRE}) - \text{MSEM}(\beta_{MKL}) \\
 &= \sigma^2 A_k(S + Q)A_k + k^2 A_k \beta \beta' A_k - \sigma^2 A_k(S - kS^{-1} + Q)A_k - 4k^2 A_k \beta \beta' A_k \\
 &= \sigma^2 A_k M A_k - \sigma^2 A_k(M - kS^{-1})A_k - 3k^2 A_k \beta \beta' A_k \\
 &= \sigma^2 A_k [M - (M - kS^{-1})] A_k - 3k^2 A_k \beta \beta' A_k \\
 &= k\sigma^2 A_k S^{-1} A_k - 3k^2 A_k \beta \beta' A_k \\
 &= kA_k (\sigma^2 S^{-1} - 3k\beta \beta') A_k
 \end{aligned} \tag{28}$$

**Theorem 3.5** The necessary and sufficient conditions for mixed KL estimator  $\hat{\beta}_{MKL}$  to be superior to the mixed ridge estimator  $\hat{\beta}_{MRE}$  under MSEM criterion are as follows:

$$3k\sigma^{-2}\beta'S\beta \leq 1 \tag{29}$$

### Numerical example and simulation study

In order to further explain the theoretical results, this section will verify and analyze the above theoretical results through examples.

The example analysis data adopts the percentage data of research and development expenses in GNP of several countries from 1972 to 1986 used by Gruber<sup>21</sup>, Akdeniz and Erol<sup>22</sup>, in which  $x_1$  represents France,  $x_2$  represents West Germany,  $x_3$  represents Japan,  $x_4$  represents the former Soviet Union and  $y$  represents the United States. See Table 1 for specific data.

The data in Table 1 are processed as follows

Year	$x_1$	$x_2$	$x_3$	$x_4$	$y$
1972	1.9	2.2	1.9	3.7	2.3
1975	1.8	2.2	2	3.8	2.2
1979	1.8	2.4	2.1	3.6	2.2
1980	1.8	2.4	2.2	3.8	2.3
1981	2	2.5	2.3	3.8	2.4
1982	2.1	2.6	2.4	3.7	2.5
1983	2.1	2.6	2.6	3.8	2.6
1984	2.2	2.6	2.6	4	2.6
1985	2.3	2.8	2.8	3.7	2.7
1986	2.3	2.7	2.8	3.8	2.7

**Table 1.** 1972–1986 research and development expenditure as a percentage of GNP.

	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
$\alpha_1$	0.6455	0.6102	0.6455	0.6452	0.6449
$\alpha_2$	0.0896	0.0988	0.0896	0.0894	0.0892
$\alpha_3$	0.1436	0.1577	0.1436	0.1434	0.1433
$\alpha_4$	0.1526	0.1566	0.1526	0.1530	0.1534
MSE	0.0431	0.0235	0.0561	0.0390	0.0180

**Table 2.** Estimated MSE.

$$X = \begin{pmatrix} 7 & 26 & 6 & 60 \\ 1 & 29 & 15 & 52 \\ 11 & 56 & 8 & 20 \\ 11 & 31 & 8 & 47 \\ 7 & 52 & 6 & 33 \\ 11 & 55 & 9 & 22 \\ 3 & 71 & 17 & 6 \\ 1 & 31 & 22 & 44 \\ 2 & 54 & 18 & 22 \\ 21 & 47 & 4 & 26 \\ 1 & 40 & 23 & 34 \\ 11 & 66 & 9 & 12 \\ 10 & 68 & 12 & 12 \end{pmatrix}, y = \begin{pmatrix} 78.5 \\ 74.3 \\ 104.3 \\ 87.6 \\ 95.9 \\ 109.2 \\ 102.7 \\ 72.5 \\ 93.1 \\ 115.9 \\ 83.8 \\ 113.3 \\ 109.4 \end{pmatrix}$$

Firstly, it is easy to calculate that the eigenvalue of  $X'X$  is  $\lambda_1 = 302.9626, \lambda_2 = 0.7283, \lambda_3 = 0.0446, \lambda_4 = 0.0345$ , the OLS estimator of  $\sigma^2$  is  $\hat{\sigma}^2 = 0.0015$ , and OLS estimator of  $\beta$  is  $\hat{\beta}_{OLS} = (0.6455, 0.0896, 0.1436, 0.1526)'$ .

We can use the method proposed by Kibria and Lukman<sup>14</sup> to choose the biasing parameter  $k$ , and we can also use the generalized cross validation (GCV) criterion and the cross validation (CV) to choose the biasing parameter, the reference can refer to Arashi et al.<sup>23</sup>, Roozbeh<sup>24</sup>, and Roozbeh et al.<sup>25</sup>. In this paper we use the method proposed by Kibria and Lukman<sup>14</sup> to choose the biasing parameter  $k$ , which is given as follows:

$$\hat{k}_i = \frac{\hat{\sigma}^2}{2\hat{\alpha}_i^2 + (\hat{\sigma}^2/\lambda_i)} \quad (30)$$

we take  $k = \hat{k}_{\min}$ .

Consider the following stochastic restrictions, this can refer to Roozbeh et al.<sup>26</sup> and Roozbeh and Hamzah<sup>27</sup>:

$$r = R\beta + e, R = (1 \ -2 \ -2 \ -2), r = 1, e \sim (0, \hat{\sigma}^2)$$

For the mixed estimator, KL estimator, OLS estimator, mixed ridge estimator and mixed KL estimator proposed in this paper. The MSE is presented in Table 2.

As can be seen from Table 2:

When  $k$  takes  $\hat{k}_{\min} = 0.018$ , the MSE value of mixed KL estimator  $\hat{\beta}_{MKL}$  is better than that of mixed estimator, KL estimator, OLS estimator and mixed ridge estimator. Consistent with the theoretical results of this paper, it can be concluded that adding stochastic restrictions may have better estimation effect under certain conditions. So in practice we can use the stochastic restrictions to address the multicollinearity.

Next, we consider Monte Carlo simulation analysis.

Firstly, the generation of relevant parameters and data in the process of simulation analysis is briefly described.

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.0024	0.0028	0.0028	0.0023	0.0023
0.9	0.0032	0.0042	0.0043	0.0032	0.0032
0.99	0.0197	0.0214	0.0282	0.0185	0.0160

**Table 3.** Estimated MSE when  $\sigma^2 = 0.1, n = 30$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.0012	0.0013	0.0013	0.0012	0.0012
0.9	0.0027	0.0031	0.0031	0.0027	0.0027
0.99	0.0146	0.0221	0.0245	0.0158	0.0152

**Table 4.** Estimated MSE when  $\sigma^2 = 0.1, n = 50$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.0011	0.0011	0.0011	0.0011	0.0011
0.9	0.0016	0.0019	0.0020	0.0016	0.0016
0.99	0.0126	0.0179	0.0205	0.0119	0.0015

**Table 5.** Estimated MSE when  $\sigma^2 = 0.1, n = 70$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.0009	0.0009	0.0009	0.0009	0.0009
0.9	0.0015	0.0018	0.0018	0.0015	0.0015
0.99	0.0104	0.0137	0.0142	0.0102	0.0101

**Table 6.** Estimated MSE when  $\sigma^2 = 0.1, n = 100$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.2375	0.1963	0.2710	0.1752	0.1697
0.9	0.3463	0.2594	0.4260	0.2354	0.2231
0.99	3.5782	2.7915	4.2485	1.8076	1.3056

**Table 7.** Estimated MSE when  $\sigma^2 = 1, n = 30$ .

The data generation of explanatory variables adopts the same method as McDonald and Galarneau<sup>28</sup>, Gibbons<sup>29</sup>), that is, it is generated by the following equation:

$$x_{ij} = (1 - \rho^2)^{1/2} z_{ij} + \rho z_{ip}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, p$$

where  $z_{ij}$  is the random number generated by the standard normal random variable,  $\rho$  is the given constant, and  $\rho^2$  theoretically represents the correlation between two different variables, so  $\rho^2$  reflects the degree of complex collinearity of the model to some extent. In this simulation analysis, we consider three cases  $\rho = 0.85, 0.9, 0.99$ , set  $p = 3, r = 1, R = \begin{pmatrix} 1 & -2 & -2 \end{pmatrix}, e \sim (0, \sigma^2), n = 30, 50, 70, 100$ .

For a given design matrix  $X$ , we take the orthogonalized eigenvector corresponding to the maximum eigenvalue of  $X'X$  as the real value of parameter vector  $\beta$ .

The data corresponding to the response variable is generated by the following equation:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

where  $\varepsilon_i$  is the mean of zero, and random vector with variance of  $\sigma^2 = 0.1, 1, 5, 10$ .

See Tables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 and 18 for simulation analysis and calculation results. Based on Tables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17 and 18, the following conclusions are drawn:

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.1908	0.1775	0.2012	0.1673	0.1675
0.9	0.1932	0.1968	0.2340	0.1662	0.1660
0.99	1.9234	1.3309	2.9377	1.3809	0.7301

**Table 8.** Estimated MSE when  $\sigma^2 = 1, n = 50$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.0984	0.0868	0.1033	0.0835	0.0828
0.9	0.1524	0.1533	0.1906	0.1277	0.1270
0.99	1.7543	1.1738	2.0825	1.0379	0.9163

**Table 9.** Estimated MSE when  $\sigma^2 = 1, n = 70$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.0751	0.0737	0.0823	0.0680	0.0678
0.9	0.1424	0.1314	0.1548	0.1218	0.1215
0.99	1.0463	0.4883	1.3303	0.5769	0.3634

**Table 10.** Estimated MSE when  $\sigma^2 = 1, n = 100$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	8.7289	3.3534	13.8404	3.7486	3.3308
0.9	10.7412	4.5196	13.4709	4.4088	3.7785
0.99	84.6872	115.8703	143.0686	45.0361	31.8312

**Table 11.** Estimated MSE when  $\sigma^2 = 5, n = 30$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	6.3491	2.8390	7.8722	2.7472	2.5652
0.9	4.3795	2.1116	4.8651	1.8364	1.7838
0.99	51.4274	44.6585	75.3850	30.2169	24.9511

**Table 12.** Estimated MSE when  $\sigma^2 = 5, n = 50$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	3.2597	1.7710	3.4194	1.3309	1.5452
0.9	3.9034	1.7501	4.3951	1.6719	1.5983
0.99	27.7983	32.8762	38.5172	21.0908	19.3660

**Table 13.** Estimated MSE when  $\sigma^2 = 5, n = 70$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	0.8875	1.7599	0.9286	1.9500	0.8946
0.9	1.4590	2.9916	1.5563	3.5008	1.4880
0.99	10.8178	22.8642	19.2461	31.9632	11.7280

**Table 14.** Estimated MSE when  $\sigma^2 = 5, n = 100$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	23.4609	8.7136	27.9864	8.4935	8.0303
0.9	28.7442	13.2491	31.8223	12.4392	11.8325
0.99	343.6973	450.0098	539.6959	250.2456	106.987

**Table 15.** Estimated MSE when  $\sigma^2 = 10, n = 30$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	19.1189	7.0762	23.1833	6.7891	6.8042
0.9	30.4727	10.6969	32.8809	10.1258	9.0393
0.99	226.4676	335.4425	390.911	130.0564	97.6481

**Table 16.** Estimated MSE when  $\sigma^2 = 10, n = 50$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	11.7984	3.8394	13.2484	3.9206	3.7376
0.9	18.0514	5.6690	19.9073	5.7731	5.4011
0.99	197.0861	114.1176	237.743	91.9548	54.3546

**Table 17.** Estimated MSE when  $\sigma^2 = 10, n = 70$ .

$\rho$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MRE}$	$\hat{\beta}_{MKL}$
0.85	8.6620	3.0434	8.9405	3.1080	2.9602
0.9	16.3215	5.6973	17.6756	5.6112	5.4424
0.99	120.1565	67.5786	168.6498	62.5914	47.1027

**Table 18.** Estimated MSE when  $\sigma^2 = 10, n = 100$ .

- (1) The mean square error of all estimates increases with the increase of  $\rho$  and decreases with the increase of  $n$ ;
- (2) The new estimator mixed KL estimator always has the minimum MSE when all given  $n$  and  $\sigma^2$ , and  $k$  takes  $k_{\min}$ . Consistent with the results of Theorems 3.2–3.5 in this paper, under certain conditions, mixed KL estimator  $\hat{\beta}_{MKL}$  is better than mixed estimator  $\hat{\beta}_{ME}$ , KL estimator  $\hat{\beta}_{KL}$ , least square estimator  $\hat{\beta}_{OLS}$  and mixed ridge estimator  $\hat{\beta}_{MRE}$  under MSE criterion;
- (3) Under the same conditions, mixed estimator  $\hat{\beta}_{ME}$ , mixed ridge estimator  $\hat{\beta}_{MRE}$  and mixed KL estimator  $\hat{\beta}_{MKL}$  are better than unconstrained least squares estimator  $\hat{\beta}_{OLS}$  under MSE criterion, mixed KL estimator  $\hat{\beta}_{MKL}$  is better than unconstrained KL estimator  $\hat{\beta}_{KL}$  under MSE criterion.

## Conclusions

In this paper, a new mixed KL estimator considering the prior information about parameters in sample information in linear model is proposed, and the properties of the new estimator are discussed. The necessary and sufficient conditions for KL estimator to be better than mixed estimator, KL estimator, OLS estimator and mixed ridge estimator under the criterion of mean square error matrix are given, and the proofs are given respectively. Then the theoretical results are verified by examples and simulation analysis.

Received: 3 April 2022; Accepted: 13 July 2022

Published online: 20 July 2022

## References

1. Massy, W. F. Principal components regression in exploratory statistical research. *J. Am. Stat. Assoc.* **60**, 234–256 (1965).
2. Hoerl, A. E. & Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **12**, 55–67 (1970).
3. Swindel, B. F. Good estimators based on prior information. *Commun. Stat. Theory Methods* **5**, 1065–1075 (1976).
4. Lukman, A. F., Ayinde, K., Binuomote, S. & Onate, A. C. Modified ridge-type estimator to combat multicollinearity. *J. Chemom.* **e3125**, 1–12 (2019).
5. Liu, K. J. A new class of biased estimate in linear regression. *Commun. Stat. Theory Methods* **22**, 393–402 (1993).
6. Akdeniz, F. & Kaciranlar, S. On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. *Commun. Stat. Theory Methods* **24**, 1789–1797 (1995).

7. Liu, K. J. Using Liu-type estimator to combat collinearity. *Commun. Stat. Theory Methods* **32**, 1009–1020 (2003).
8. Baye, M. R. & Parker, D. F. Combining ridge and principal component regression: A money demand illustration. *Commun. Stat. Theory Methods* **13**, 197–225 (1984).
9. Kaciranlar, S. & Sakallioglu, S. Combining the Liu estimator and the principal component regression estimator. *Commun. Stat. Theory Methods* **30**, 2699–2705 (2001).
10. Ozkale, M. R. & Kaciranlar, S. The restricted and unrestricted two-parameter estimators. *Commun. Stat. Theory Methods* **36**, 2707–2725 (2007).
11. Batah, F. M., Ozkale, M. R. & Gore, S. D. Combining unbiased ridge and principal component regressions estimators. *Commun. Stat. Theory Methods* **38**, 2201–2209 (2009).
12. Yang, H. & Chang, X. F. A new two-parameter estimator in linear regression. *Commun. Stat. Theory Methods* **39**(6), 923–934 (2010).
13. Lukman, A. F., Ayinde, K., Oludoun, O. & Onate, C. A. Combining modified ridge-type and principal component regression estimators. *Sci. Afr.* **e536**, 1–8 (2020).
14. Kibria, B. M. G. & Lukman, A. F. A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientific Reports* <https://doi.org/10.1155/2020/9758378> (2020).
15. Theil, H. & Goldberger, A. S. On pure and mixed estimation in econometrics. *Int. Econ. Rev.* **2**, 65–78 (1961).
16. Theil, H. On the use of incomplete prior information in regression analysis. *J. Am. Stat. Assoc.* **58**, 401–414 (1963).
17. Schiffrin, B. & Toutenburg, H. Weighted mixed regression. *Z. Angew. Math. Mech.* **70**, 735–738 (1990).
18. Hubert, M. H. & Wijekoon, P. Improvement of the Liu estimator in linear regression coefficient. *Stat. Pap.* **47**, 471–479 (2006).
19. Yang, H. & Xu, J. W. An alternative stochastic restricted Liu estimator in linear regression model. *Stat. Pap.* **50**, 369–647 (2009).
20. Ozbay, N. & Kaciranlar, K. S. Estimation in a linear regression model with stochastic linear restrictions: A new two-parameter-weighted mixed estimator. *J. Stat. Comput. Simul.* **88**, 1669–1683 (2018).
21. Gruber, M. H. J. *Improving Efficiency by Shrinkage: The James–Stein and Ridge Regression estimators* (Marcel Dekker Inc, 1998).
22. Akdeniz, F. & Erol, H. Mean Squared error matrix comparisons of some biased estimator in linear regression. *Commun. Stat. Theory Methods* **32**(12), 2389–2413 (2003).
23. Arashi, M. *et al.* Ridge regression and its applications in genetic studies. *PLoS One* **16**(4), e0245376 (2021).
24. Roozbeh, M. & Azen, S. P. Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion. *Comput. Stat. Data Anal.* **117**, 45–61 (2018).
25. Roozbeh, M., Arahi, M. & Hamzah, N. A. Generalized cross-validation for simultaneous optimization of tuning parameters in ridge regression. *Iran. J. Sci. Technol. Trans. A, Sci.* **44**(2), 473–485 (2020).
26. Roozbeh, M., Hesamiab, G. & Akbaric, M. G. Ridge estimation in semi-parametric regression models under the stochastic restriction and correlated elliptically contoured errors. *J. Comput. Appl. Math.* **378**, 112940 (2020).
27. Roozbeh, M. & Hamzah, N. A. Uncertain stochastic ridge estimation in partially linear regression models with elliptically distributed errors. *Statistics* **3**, 494–523 (2022).
28. McDonald, M. C. & Galarneau, D. I. A Monte Carlo evaluation of ridge-type estimators. *J. Am. Stat. Assoc.* **70**, 407–416 (1975).
29. Gibbons, D. G. A simulation study of some ridge estimators. *J. Am. Stat. Assoc.* **76**, 131–139 (1981).

## Author contributions

H.C. and J.W. wrote the main manuscript text. All authors reviewed the manuscript.

## Funding

The authors are highly obliged to the editor and the reviewers for the comments and suggestions which improved the paper in its present form. This work was sponsored by the Natural Science Foundation of Chongqing [grant number cstc2020jcyj-msxmX0028] and the Scientific Technological Research Program of Chongqing Municipal Education Commission [grant number KJQN202001321].

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to J.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022