STATISTICS IN TRANSITION new series, Winter 2014
Vol. 15, No. 1, pp. 23–36

A MODIFIED TWO-PARAMETER ESTIMATOR IN LINEAR REGRESSION

Ashok V. Dorugade[1]

ABSTRACT

In this article, a modified two-parameter estimator is introduced for the vector of parameters in the linear regression model when data exists with multicollinearity. The properties of the proposed estimator are discussed and the performance in terms of the matrix mean square error criterion over the ordinary least squares (OLS) estimator, a new two-parameter estimator (NTP), an almost unbiased two-parameter estimator (AUTP) and other well known estimators reviewed in this article is investigated. A numerical example and simulation study are finally conducted to illustrate the superiority of the proposed estimator.

Key words: liu estimator, multicollinearity, two-parameter estimator, mean squared error matrix.

1. Introduction

In practice, there can be strong or near to strong linear relationships among the explanatory variables. In that case the independent assumptions are no longer valid, which causes the problem of multicollinearity. In the presence of multicollinearity, it is impossible to estimate the unique effects of individual variables in the regression equation. Also, the OLS estimator yields regression coefficients whose absolute values are too large and whose signs can actually reverse with negligible changes in the data (see Buonaccorsi, 1996). Therefore, multicollinearity becomes one of the serious problems in the linear regression analysis. The method of ridge regression, proposed by Hoerl and Kennard (1970a) is a popular technique for estimating the regression parameter for the ill-conditioned multiple linear regression models.

Much of the discussion on ridge regression concerns the problem of finding better alternative to the OLS estimator. Some popular numerical techniques to deal with multicollinearity are the ridge regression due to Stein estimator (Stein, 1956), contraction estimator (Mayer and Willke, 1973), modified ridge regression

[1] Y C Mahavidyalaya Halkarni, Tal-Chandgad, Kolhapur, Maharashtra, India - 416552.
E-mail: adorugade@rediffmail.com.

(MRR) estimator (Swindel, 1976), Kadiyala (1984), Ohtani (1986), Singh and Chaubey (1987), Nomura (1988), and Gruber (1998) Sing et al. (1988), Liu (1993), Akdeniz and Kaciranlar (1995), Crouse et al. (1995), Ozkale and Kaciranlar (2007), Batah et al. (2008), Sakallioglu and Kaciranlar (2008), Yang and Chang (2010), Wu and Yang (2011), Dorugade and Kashid (2011) and others.

In this paper we introduce a modified two-parameter estimator for the vector of parameters in the linear regression model when data exists with multicollinearity. The rest of the paper is organized as follows. The model and some well known estimators are reviewed in section 2. The modified two-parameter estimator is introduced in section 3. Performances of the proposed estimator with respect to the scalar MSE criterion are discussed in section 4. In section 5, we give methods to choose the biasing parameters. A simulation study to justify the superiority of the suggested estimator is given in section 6. Some concluding remarks are given in section 7.

2. Model and estimators

Consider a widely used linear regression model

$$Y = X\beta + \varepsilon\ , \tag{1}$$

where Y is an n×1 vector of observations on a response variable. β is a p×1 vector of unknown regression coefficients, X is a matrix of order (n × p) of observations on 'p' predictor (or regressor) variables and ε is an n × 1 vector of errors with E(ε) = 0 and V(ε) = $\sigma^2 I_n$. For the sake of convenience, we assume that the matrix X and the response variable Y are standardized in such a way that $X'X$ is a non-singular correlation matrix and $X'Y$ is the correlation between X and Y. The paper is concerned with data exhibited with multicollinearity leading to a high MSE for β meaning that $\hat{\beta}$ is an unreliable estimator of β.

Let Λ and T be the matrices of eigenvalues and eigenvectors of $X'X$, respectively, satisfying $T'X'XT$ = Λ = diagonal ($\lambda_1, \lambda_2, ..., \lambda_p$), where λ_i being the ith eigenvalue of $X'X$ and $T'T = TT' = I_p$. We obtain the equivalent model

$$Y = Z\alpha + \varepsilon\ , \tag{2}$$

where Z = XT. It implies that $Z'Z = \Lambda$, and α = $T'\beta$ (see Montgomery et al., 2006).

Then, the OLS estimator of α is given by

$$\hat{\alpha} = (Z'Z)^{-1} Z'Y = \Lambda^{-1}\ Z'Y \tag{3}$$

Therefore, the OLS estimator of β is given by

$$\hat{\beta} = T\hat{\alpha}$$

2.1. Ordinary ridge estimator (ORR)

A popular estimator for combating multicollinearity is the ridge estimator, originally introduced by Hoerl and Kennard (1970a) as

$$\hat{\beta}_R = \mathrm{T}\,\hat{\alpha}_R = \mathrm{T}\left[I - k(\Lambda + kI)^{-1}\right]\hat{\alpha} \tag{4}$$

where k is the ridge parameter (or biasing constant), and it normally lies between 0 and 1. $\hat{\alpha}_i$ is the i[th] element of $\hat{\alpha}$, $i = 1,2,\ldots,p$ and $\hat{\sigma}^2$ is the OLS estimator of σ^2 i.e. $\hat{\sigma}^2 = (Y'Y - \hat{\alpha}'Z'Y)/(n-p-1)$.

The ridge regression method has been considered by various researchers. The drawback of the ridge regression method is that it is a complicated function of *k*. To overcome this problem Liu (1993) proposed an estimator which combines the benefit of both the estimators given by Hoerl and Kennard (1970a) and Stein (1956), respectively.
It is given as

$$\hat{\alpha}_{Liu} = (\Lambda + I)^{-1}(\Lambda + dI)\hat{\alpha} \qquad 0 < d < 1 \tag{5}$$

Liu estimator has been considered by several researchers several times for different perspectives. Following Liu many researchers propose two-parameter ridge estimators. Ozkale and Kaciranlar (2007) obtained the two-parameter (TP) estimator given as

$$\hat{\alpha}_{TP} = (\Lambda + kI)^{-1}(\Lambda + kdI)\hat{\alpha} \tag{6}$$

MSE of $\hat{\alpha}_{TP}$ is given as

$$MSE(\hat{\alpha}_{TP}) = \sigma^2 \sum_{i=1}^{p}\left[\frac{(\lambda_i + kd)^2}{\lambda_i(\lambda_i + k)^2}\right] + \sum_{i=1}^{p}\left[\frac{k^2(1-d)^2}{(\lambda_i + k)^2}\right]\alpha_i^2 \tag{7}$$

Sakallioglu and Kaciranlar (2008) suggested the following two-parameter estimator:

$$\hat{\alpha}_{LTE(3)} = (\Lambda + I)^{-1}(\Lambda + (d+k)I)^{-1}(\Lambda + kI)^{-1}Z'Y\,. \tag{8}$$

MSE of $\hat{\alpha}_{LTE(3)}$ is given as

$$MSE(\hat{\alpha}_{LTE(3)}) = \sigma^2 \sum_{i=1}^{p}\left[\frac{\lambda_i(\lambda_i + (k+d))^2}{(\lambda_i + 1)^2(\lambda_i + k)^2}\right] + \sum_{i=1}^{p}\left[\frac{(\lambda_i(1-d) + k)^2}{(\lambda_i + 1)^2(\lambda_i + k)^2}\right]\alpha_i^2 \tag{9}$$

since the ridge parameter $\hat{k} = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p}\hat{\alpha}_i^2$ given by Hoerl et al. (1975) performs fairly well and the well-known estimate of ‘ d ’proposed by Liu (1993) is given as

$$\hat{d} = \sum_{i=1}^{p}(\hat{\alpha}_i^2 - \hat{\sigma}^2)/(\lambda_i + 1)^2 \Big/ \sum_{i=1}^{p}(\hat{\sigma}^2 + \lambda_i\hat{\alpha}_i^2)/(\lambda_i + 1)^2\lambda_i$$

The above calculated values of $\hat{k}$ and $\hat{d}$ are used in determination of estimators given in equations (6) and (8).

On the other hand, Yang and Chang (2010) introduce a new two-parameter (NTP) estimator given as

$$\hat{\alpha}_{NTP} = \left(\Lambda + I\right)^{-1}\left(\Lambda + dI\right)\left(\Lambda + kI\right)^{-1} Z^{'}Y \,, \tag{10}$$

where $\hat{k} = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p} \hat{\alpha}_i^2$

and $$\hat{d} = \frac{\sum_{i=1}^{p} \left\{\left[(k+1)\lambda_i + k\right]\lambda_i \hat{\alpha}_i^2 - \lambda_i^{\,2}\hat{\sigma}^2\right\} \Big/ \left[(\lambda_i+1)^2(\lambda_i+k)^2\right]}{\sum_{i=1}^{p} (\hat{\sigma}^2 + \lambda_i\hat{\alpha}_i^2) \Big/ \left[(\lambda_i+1)^2(\lambda_i+k)^2\right]} .$$

It includes the OLS, RR, and Liu estimators as special cases and provides an alternative method to overcome multicollinearity in linear regression.
Also, MSE of $\hat{\alpha}_{NTP}$ is given as

$$MSE\left(\hat{\alpha}_{NTP}\right) = \sigma^2 \sum_{i=1}^{p} \left[\frac{\lambda_i(\lambda_i+d)^2}{(\lambda_i+1)^2(\lambda_i+k)^2}\right] + \sum_{i=1}^{p} \left\{\frac{\left[(k+1-d)\lambda_i + k\right]^2}{(\lambda_i+1)^2(\lambda_i+k)^2}\right\}\alpha_i^2 \tag{11}$$

Recently Wu and Yang (2011) introduced an almost unbiased two-parameter (AUTP) estimator alternative to the OLS estimator in the presence of multicollinearity. These estimators are given as

$$\hat{\alpha}_{AUTP} = \hat{\alpha}_{TP} + k(1-d)\left(\Lambda + kI\right)^{-1}\hat{\alpha}_{TP} \,, \tag{12}$$

where $\hat{d} < 1 - \min\left(\hat{\sigma} \Big/ \sqrt{\lambda_i\hat{\alpha}_i^2 + \hat{\sigma}^2}\right)$ and $\hat{k} = \lambda_i\hat{\sigma} \Big/ \left[(1-d)\sqrt{\lambda_i\hat{\alpha}_i^2 + \hat{\sigma}^2} - \hat{\sigma}\right]$

Also, MSE of $\hat{\alpha}_{AUTP}$ is given as

$$MSE\left(\hat{\alpha}_{AUTP}\right) = \sigma^2 \sum_{i=1}^{p} \left\{\frac{\left[\lambda_i(\lambda_i+2k) + d\,k^2(2-d)\right]^2}{\lambda_i\left(\lambda_i+k\right)^4}\right\} + \sum_{i=1}^{p} \left[\frac{k^4(1-d)^4}{\left(\lambda_i+k\right)^4}\right]\alpha_i^2 \tag{13}$$

Estimators given in equations (6), (8), (10) and (12) used for estimating α are used in section 6.

3. Proposed ridge estimator

In this article we introduce a modified two-parameter estimator and it can be computed in two steps. Initially, following a similar method proposed by Liu (1993), Kaciranlar et al. (1999) and Yang and Chang (2010) we introduce two-parameter estimator as

$$\hat{\alpha}^{*} = \left(\Lambda + k\,d\,I\right)^{-1} Z^{'}Y \tag{14}$$

Then, following Kadiyala (1984), Ohtani (1986) and Wu and Yang (2011) the estimator defined in equation (14) can be rewritten as

$$\hat{\alpha}_{MTP} = \hat{\alpha}^* + k(1-d)\left(\Lambda + kdI\right)^{-1} \hat{\alpha}^* \tag{15}$$

or

$$\hat{\alpha}_{MTP} = \left[I + k(1-d)(\Lambda + kdI)^{-1}\right]\left[I - kd\,(\Lambda + kdI)^{-1}\right]\hat{\alpha}\,.$$

It is termed as a modified two-parameter (MTP) estimator of α .

Thus, the coordinate wise estimators can be written as

$$\hat{\alpha}_{iMTP} = \left[\frac{\lambda_i(\lambda_i + k)}{(\lambda_i + kd)^2}\right]\hat{\alpha}_i \qquad i = 1,2,\dots,p \tag{16}$$

where $\hat{\alpha}_i$ are the individual components of $\hat{\alpha}$.

We can see that it is a general estimator which includes the OLS and RR estimators as special cases:

at $\left(k = 0 \;\; or \;\; d = 0\right)$ $\quad \hat{\alpha}_{MTP} = \Lambda^{-1} Z^{'}Y$, the OLS estimator

at $d = 1$ $\quad \hat{\alpha}_{MTP} = \left(\Lambda + kI\right)^{-1} Z^{'}Y$, the RR estimator

Obviously,

$\hat{\alpha}_{iMTP} = (\hat{\alpha}_i)_{OLS}$ at $\left(k = 0 \;\; or \;\; d = 0\right)$

and $\hat{\alpha}_{iMTP} = (\hat{\alpha}_i)_R$ at $d = 1$

3.1. Bias, variance and MSE of MTP estimator

It is clear that $\hat{\alpha}_{MTP}$ is a biased estimator, with the bias of the MTP estimator is given by:

$$Bias\left(\hat{\alpha}_{MTP}\right) = E[\hat{\alpha}_{MTP}] - \alpha \;\; = [k(1-2d)(\Lambda + kdI)^{-1} - k^2 d\,(1-d)(\Lambda + kdI)^{-2}]\alpha$$

$$= \sum_{i=1}^{p}\left\{\frac{k\left[(1-2d)\lambda_i - kd^2\right]}{\left(\lambda_i + kd\right)^2}\right\}\alpha_i \tag{17}$$

$$V\left(\hat{\alpha}_{MTP}\right) = \sigma^2 V \Lambda^{-1} V^{'}$$

where $V = \left[I + k(1-d)(\Lambda + kdI)^{-1}\right]\left[I - kd\,(\Lambda + kdI)^{-1}\right]$

$$= \sigma^2 \sum_{i=1}^{p}\left[\frac{\lambda_i\left(\lambda_i + k\right)^2}{\left(\lambda_i + kd\right)^4}\right] \tag{18}$$

The MSE of MTP estimator is

$$MSE\left(\hat{\alpha}_{MTP}\right) = V\left(\hat{\alpha}_{MTP}\right) + \left[Bias\left(\hat{\alpha}_{MTP}\right)\right]\left[Bias\left(\hat{\alpha}_{MTP}\right)\right]^{'},$$

$$MSE\left(\hat{\alpha}_{MTP}\right) = \sigma^2 \sum_{i=1}^{p}\left[\frac{\lambda_i\left(\lambda_i + k\right)^2}{\left(\lambda_i + kd\right)^4}\right] + \sum_{i=1}^{p}\left\{\frac{k^2\left[(1-2d)\lambda_i - kd^2\right]^2}{\left(\lambda_i + kd\right)^4}\right\}\alpha_i^2 \tag{19}$$

Setting $k=0$ or $d=0$ in equation (19), we obtain

$$MSE(\hat{\alpha}) = \sigma^2 \sum_{i=1}^{p} \frac{1}{\lambda_i} \tag{20}$$

Also, setting $d=1$ in equation (19), we obtain

$$MSE(\hat{\alpha}_R) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+k)^2} + k^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i+k)^2} \tag{21}$$

4. Performance of proposed estimator

This section compares the performance of the $\hat{\alpha}_{MTP}$ with the $\hat{\alpha}$, $\hat{\alpha}_{AUTP}$ and $\hat{\alpha}_{NTP}$ using smaller MSE criteria.

4.1. Comparison between $\hat{\alpha}_{MTP}$ and $\hat{\alpha}$

In order to compare $\hat{\alpha}_{MTP}$ with $\hat{\alpha}$ in the MSE sense, using equations (19) and (20) we investigate the following difference:

$$MSE(\hat{\alpha}) - MSE(\hat{\alpha}_{MTP}) = \sigma^2 \sum_{i=1}^{p} \left[\frac{1}{\lambda_i} - \frac{\lambda_i(\lambda_i+k)^2}{(\lambda_i+kd)^4} \right] - k^2 \sum_{i=1}^{p} \left\{ \frac{\left[(1-2d)\lambda_i - k^2 d^2\right]^2}{(\lambda_i+kd)^4} \right\} \alpha_i^2$$

$$= \sum_{i=1}^{p} \left\{ \frac{\sigma^2\left[(\lambda_i+kd)^4 - \lambda_i^2(\lambda_i+k)^2\right] - \lambda_i \alpha_i^2 k^2 \left[(1-2d)\lambda_i - k^2 d^2\right]^2}{\lambda_i(\lambda_i+kd)^4} \right\}$$

From above equation it can be seen that $MSE(\hat{\alpha}_{OLS}) \geq MSE(\hat{\alpha}_{MTP})$ if and only if

$$\sigma^2\left[(\lambda_i+kd)^4 - \lambda_i^2(\lambda_i+k)^2\right] \geq \lambda_i \alpha_i^2 k^2 \left[(1-2d)\lambda_i - k^2 d^2\right]^2$$

4.2. Comparison between $\hat{\alpha}_{MTP}$ and $\hat{\alpha}_{AUTP}$

Wu and Yang (2011) proposes the almost unbiased two-parameter estimator ($\hat{\alpha}_{AUTP}$) given in equation (12). Also, they compare performance of their estimator with the OLS estimator and the two-parameter estimator given in equation (6). To compare $\hat{\alpha}_{MTP}$ with $\hat{\alpha}_{AUTP}$ in the MSE sense, using equations (19) and (13) we investigate the following difference:

$$MSE(\hat{\alpha}_{AUTP}) - MSE(\hat{\alpha}_{MTP})$$

$$= \sigma^2 \sum_{i=1}^{p} \left\{ \frac{\left[\lambda_i(\lambda_i+2k) + d\,k^2(2-d)\right]^2}{\lambda_i(\lambda_i+k)^4} - \frac{\lambda_i(\lambda_i+k)^2}{(\lambda_i+kd)^4} \right\}$$

$$+ k^2 \sum_{i=1}^{p} \left\{ \frac{k^2(1-d)^4}{(\lambda_i+k)^4} - \frac{\left[(1-2d)\lambda_i - kd^2\right]^2}{(\lambda_i+kd)^4} \right\} \alpha_i^2$$

$$= \sigma^2 \sum_{i=1}^{p} \left\{ \frac{\left[\lambda_i(\lambda_i + 2k) + d\,k^2(2-d)\right]^2 (\lambda_i + kd)^4 - \lambda_i^{\,2}(\lambda_i + k)^6}{(\lambda_i + kd)^4 \lambda_i (\lambda_i + k)^4} \right\}$$

$$+ k^2 \sum_{i=1}^{p} \left\{ \frac{k^2(1-d)^4(\lambda_i + kd)^4 - \left[(1-2d)\lambda_i - kd^2\right]^2 (\lambda_i + k)^4}{(\lambda_i + kd)^4 (\lambda_i + k)^4} \right\} \alpha_i^2$$

$$= \sum_{i=1}^{p} \left\{ \frac{\left\{\sigma^2\left[\lambda_i(\lambda_i + 2k) + d\,k^2(2-d)\right]^2 + \lambda_i k^4 (1-d)^4 \alpha_i^2\right\}(\lambda_i + kd)^4 - \left\{\sigma^2 \lambda_i^{\,2}(\lambda_i + k)^2 + \lambda_i \alpha_i^2\left[(1-2d)\lambda_i - kd^2\right]^2\right\}(\lambda_i + k)^4}{(\lambda_i + kd)^4 \lambda_i (\lambda_i + k)^4} \right\}$$

From above equation it can be seen that $MSE(\hat{\alpha}_{AUTP}) \geq MSE(\hat{\alpha}_{MTP})$ if and only if

$$\left\{\sigma^2\left[\lambda_i(\lambda_i + 2k) + d\,k^2(2-d)\right]^2 + \lambda_i k^4 (1-d)^4 \alpha_i^2\right\}(\lambda_i + kd)^4$$
$$\geq \left\{\sigma^2 \lambda_i^{\,2}(\lambda_i + k)^2 + \lambda_i \alpha_i^2\left[(1-2d)\lambda_i - kd^2\right]^2\right\}(\lambda_i + k)$$

4.3. Comparison between $\hat{\alpha}_{MTP}$ and $\hat{\alpha}_{NTP}$

Yang and Chang (2010) introduced a new two-parameter (NTP) estimator and studied superiority of their estimator over the OLS estimator, Liu estimator and the two-parameter estimator. In order to compare $\hat{\alpha}_{MTP}$ with $\hat{\alpha}_{NTP}$ in the MSE sense, using equations (19) and (11) we investigate the following difference:

$$MSE(\hat{\alpha}_{NTP}) - MSE(\hat{\alpha}_{MTP})$$

$$= \sigma^2 \sum_{i=1}^{p} \left\{ \frac{\lambda_i(\lambda_i + d)^2}{(\lambda_i + 1)^2(\lambda_i + k)^2} - \frac{\lambda_i(\lambda_i + k)^2}{(\lambda_i + kd)^4} \right\}$$

$$+ \sum_{i=1}^{p} \left\{ \frac{\left[(k+1-d)\lambda_i + k\right]^2}{(\lambda_i + 1)^2(\lambda_i + k)^2} - \frac{\left[(1-2d)\lambda_i - kd^2\right]^2}{(\lambda_i + kd)^4} \right\} \alpha_i^2$$

$$= \sigma^2 \sum_{i=1}^{p} \left\{ \frac{\left[\lambda_i(\lambda_i + d)^2(\lambda_i + kd)^4\right] - \lambda_i(\lambda_i + 1)^2(\lambda_i + k)^4}{(\lambda_i + kd)^4(\lambda_i + 1)^2(\lambda_i + k)^2} \right\}$$

$$+ \sum_{i=1}^{p} \left\{ \frac{((k+1-d)\lambda_i + k)^2(\lambda_i + kd)^4 - k^2\left[(1-2d)\lambda_i - kd^2\right]^2(\lambda_i + k)^2(\lambda_i + 1)^2}{(\lambda_i + kd)^4(\lambda_i + k)^2(\lambda_i + 1)^2} \right\} \alpha_i^2$$

$$= \sum_{i=1}^{p} \left\{ \frac{\left\{\sigma^2\left[\lambda_i(\lambda_i + d)^2\right] + ((k+1-d)\lambda_i + k)^2 \alpha_i^2\right\}(\lambda_i + kd)^4 - \left\{\sigma^2 \lambda_i(\lambda_i + k)^2 + \alpha_i^2 k^2\left[(1-2d)\lambda_i - kd^2\right]^2\right\}(\lambda_i + k)^2(\lambda_i + 1)^2}{(\lambda_i + kd)^4(\lambda_i + k)^2(\lambda_i + 1)^2} \right\}$$

From above equation it can be seen that $MSE(\hat{\alpha}_{NTP}) \geq MSE(\hat{\alpha}_{MTP})$ if and only if

$$\left\{\sigma^2\left[\lambda_i(\lambda_i+d)^2\right]+((k+1-d)\lambda_i+k)^2\alpha_i^2\right\}(\lambda_i+kd)^4$$
$$\geq\left\{\sigma^2\lambda_i(\lambda_i+k)^2+\alpha_i^2k^2\left[(1-2d)\lambda_i-kd^2\right]^2\right\}(\lambda_i+k)^2(\lambda_i+1)^2$$

5. Determination of ridge parameters k and d

In ridge regression the additional parameter, the ridge parameter k, plays a vital role to control the bias of the regression towards the mean of the response variable. Although these estimators result in a bias for certain value of k they yield minimum mean squared error (MSE) compared to the OLS estimator (see Hoerl and Kennard, 1970a). Similarly, d is another ridge parameter which serves the same role as k used in determination of two-parameter estimators (see Liu, 1993).

In order to determine and evaluate the performance of our proposed estimator $\hat{\alpha}_{MTP}$ as compared to the OLS estimator and others, we will find the optimal values of k and d. Let k be fixed and determined using one of the available methods for choosing the ridge parameter value. Some of the well known methods are listed below.

$$k_1 = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p}\hat{\alpha}_i^2 \qquad \text{(Hoerl et al., (1975))} \qquad (22)$$

$$k_2 = Median\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^{2'}}\right) \qquad i=1,2,...,p \qquad \text{(Kibria, (2003))} \qquad (23)$$

$$k_3 = p\hat{\sigma}^2 \Big/ \sum_{i=1}^{p}\left\{\hat{\alpha}_i^{\,2}\Big/\left[\left[\left(\hat{\alpha}_i^{\,4}\lambda_i^2/4\hat{\sigma}^2\right)+\left(6\hat{\alpha}_i^{\,4}\lambda_i/\hat{\sigma}^2\right)\right]^{1/2}-\left(\hat{\alpha}_i^{\,2}\lambda_i/2\hat{\sigma}^2\right)\right]\right\}$$

(Batah et al., (2008)) (24)

Then, the optimal value of d can be considered to be the d that minimize $MSE(\hat{\alpha}_{MTP})$.

Let $g(k,d) = MSE(\hat{\alpha}_{MTP}) = \sigma^2\sum_{i=1}^{p}\left[\frac{\lambda_i(\lambda_i+k)^2}{(\lambda_i+kd)^4}\right]+\sum_{i=1}^{p}\left\{\frac{k^2\left[(1-2d)\lambda_i-kd^2\right]^2}{(\lambda_i+kd)^4}\right\}\alpha_i^2$

Then, by differentiating $g(k,d)$ w.r.t. d and equating to 0, we have

$$d=\sum_{i=1}^{p}\left[\frac{(\lambda_i+k)(\sigma^2+\lambda_i\alpha_i^2)-\lambda_i}{k\alpha_i^2}\right] \qquad (25)$$

Unfortunately, d depends on the unknown σ^2 and α_i. For practical purposes we replace them with their unbiased estimator $\hat{\sigma}^2$ and $\hat{\alpha}_i$, and obtain

$$\hat{d} = \sum_{i=1}^{p}\left[\frac{(\lambda_i + k)(\hat{\sigma}^2 + \lambda_i\hat{\alpha}_i^2) - \lambda_i}{k\hat{\alpha}_i^2}\right] \tag{26}$$

6. Comparative study

6.1. Numerical illustration

In this section we demonstrate the performance of the proposed estimator by considering a numerical example; we use Hald cement data (see Montgomery et al., 2006). We use the ridge parameters given in equations (22) to (24) and d given in equation (26) to compute our modified two-parameter (MTP) estimator. Also, $\hat{\alpha}_{OLS}$, $\hat{\alpha}_{TP}$, $\hat{\alpha}_{LTE(3)}$, $\hat{\alpha}_{NTP}$, $\hat{\alpha}_{AUTP}$ estimators are computed and their estimated MSE values are obtained by replacing all unknown model parameters respectively with their OLS estimators in the corresponding expressions, and the values are reported in Table 1.

Table 1. Values of estimates and MSE

Estimator	$\hat{\alpha}_{OLS}$	$\hat{\alpha}_{TP}$	$\hat{\alpha}_{LTE(3)}$	$\hat{\alpha}_{NTP}$	$\hat{\alpha}_{AUTP}$	$\hat{\alpha}_{MTP}$ at k_1	$\hat{\alpha}_{MTP}$ at k_2	$\hat{\alpha}_{MTP}$ at k_3
MSE	1.3709	0.1485	1.2605	0.1484	1.3662	0.1492	0.1487	0.1490

From Table 1 we can see that the estimated MSE value of the modified two-parameter estimator is always smaller than the one of the OLS, AUTP and LTE(3) estimators. However, we also find that the estimated MSE value of the modified two-parameter estimator for each choice of the ridge parameter is approximately equal to those of the TP and NTP estimators. The results agree with our theoretical findings in section 4.

6.2. Simulation study

Here, we examine the performance of the modified two-parameter estimator $(\hat{\alpha}_{MTP})$ over different estimators $\hat{\alpha}_{OLS}$, $\hat{\alpha}_{TP}$, $\hat{\alpha}_{LTE(3)}$, $\hat{\alpha}_{NTP}$, $\hat{\alpha}_{AUTP}$. We examine the average MSE (AMSE) ratio of the $\hat{\alpha}_{MTP}$ and other estimators over the OLS estimator. We will discuss the simulation study that compares the performance of different estimators under several degrees of multicollinearity. We consider the true model as $Y = X\beta + \varepsilon$. Here, ε follows a normal distribution $N(0, \sigma^2 I_n)$ and the explanatory variables are generated (see Batah et al., 2008) from

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

where u_{ij} are independent standard normal random numbers and ρ^2 is the correlation between x_{ij} and $x_{ij'}$ for j, $j' <$ p and $j \neq j'$. j, j' = 1, 2,…, p. When j or j' = p, the correlation is be ‘ ρ ’. Here, we consider predictor variables p = 4 and ρ = 0.9, 0.95. These variables are standardized such that $X'X$ is in the correlation matrix form and it is used for the generation of Y with $\beta = (2,1,4,5)'$. We have simulated the data with sample sizes n = 20, 50 and 100. The variance of the error terms is taken as σ^2 = 1, 5, 10 and 25. Estimators $\hat{\alpha}_{OLS}$, $\hat{\alpha}_{TP}$, $\hat{\alpha}_{LTE(3)}$, $\hat{\alpha}_{NTP}$, $\hat{\alpha}_{AUTP}$ are computed. The modified two-parameter estimator ($\hat{\alpha}_{MTP}$) is computed for different choices of ridge parameters given in equations (22) to (24) and d given in equation (26). The experiment is repeated 1500 times and obtains the average MSE (AMSE) of estimators using the following expression:

$$AMSE(\hat{\alpha}) = \frac{1}{1500}\sum_{i=1}^{4}\sum_{j=1}^{1500}(\hat{\alpha}_{ij} - \alpha_i)^2$$

where $\hat{\alpha}_{ij}$ denotes the estimator of the i[th] parameter in the j[th] replication and α_i, i=1,2,3, 4 are the true parameter values.

Firstly, we have computed the AMSE ratios ($AMSE(\hat{\alpha}_{OLS})/AMSE(\hat{\alpha})$) of the OLS estimator over different estimators for various values of triplet (ρ, n, σ^2) and reported them in Table 2. We consider the method that leads to the maximum AMSE ratio to be the best from the MSE point of view.

The same procedure for another choice of p = 3 and $\beta = (1, 2, 5)'$ is performed and AMSE ratios are computed and reported in Table 3.

Table 2. Ratio of AMSE of OLS over various two-parameter estimators (p = 4 and $\beta = (2, 1, 4, 5)'$)

ρ = 0.90													
n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
$\hat{\alpha}$	$\hat{\alpha}_{TP}$	1.185	1.2908	1.3352	1.472	1.228	1.2809	1.3539	1.478	1.235	1.2563	1.3283	1.484
	$\hat{\alpha}_{LTE(3)}$	1	1.0009	1.0023	1.004	1	1.0003	1.0007	1.001	1	1.0002	1.0008	1.001
	$\hat{\alpha}_{NTP}$	1.359	1.5054	1.5857	1.978	1.473	1.5372	1.6313	2.067	1.496	1.4705	1.576	2.089
	$\hat{\alpha}_{AUTP}$	0.997	0.9903	0.9758	1.002	0.999	0.9946	1.032	0.989	1.003	0.9959	0.9883	0.983
	$\hat{\alpha}_{MTP}$ at k_1	1.369	1.5169	1.6342	2.094	1.477	1.5402	1.6379	2.195	1.497	1.4848	1.607	2.164
	$\hat{\alpha}_{MTP}$ at k_2	1.368	1.5069	1.6059	2.044	1.471	1.5367	1.6325	2.110	1.498	1.4835	1.605	2.172
	$\hat{\alpha}_{MTP}$ at k_3	1.369	1.5167	1.6344	2.097	1.470	1.5401	1.638	2.141	1.497	1.4848	1.6071	2.165

Table 2. Ratio of AMSE of OLS over various two-parameter estimators (p = 4 and $\beta = (2, 1, 4, 5)'$) (cont.)

ρ = 0.95													
n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
$\hat{\alpha}$	$\hat{\alpha}_{TP}$	1.214	1.2792	1.4212	1.554	1.211	1.256	1.3406	1.495	1.222	1.2798	1.3456	1.502
	$\hat{\alpha}_{LTE(3)}$	1	1.0002	1.0006	1.001	1	1.0003	1.0006	1.001	1	1.0003	1.001	1.002
	$\hat{\alpha}_{NTP}$	1.448	1.5177	1.8593	2.434	1.438	1.4786	1.6337	2.128	1.46	1.5168	1.6237	2.12
	$\hat{\alpha}_{AUTP}$	1	0.9974	0.9929	1.005	0.998	1.001	1.002	0.987	0.999	0.9947	0.9954	0.983
	$\hat{\alpha}_{MTP}$ at k_1	1.447	1.5282	1.8693	2.485	1.442	1.5075	1.6492	2.172	1.468	1.5207	1.6325	2.182
	$\hat{\alpha}_{MTP}$ at k_2	1.447	1.5271	1.872	2.499	1.443	1.5053	1.6367	2.176	1.467	1.5186	1.6458	2.175
	$\hat{\alpha}_{MTP}$ at k_3	1.447	1.5281	1.8697	2.486	1.442	1.5074	1.6493	2.174	1.468	1.5206	1.6227	2.183
ρ = 0.99													
n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
$\hat{\alpha}$	$\hat{\alpha}_{TP}$	1.2819	1.3621	1.397	1.552	1.2745	1.3373	1.466	1.563	1.2615	1.37	1.457	1.565
	$\hat{\alpha}_{LTE(3)}$	1.0011	1.0021	1.003	1.004	1.0007	1.0023	1.004	1.004	1.0009	1.0026	1.004	1.005
	$\hat{\alpha}_{NTP}$	1.4962	1.6641	1.798	2.355	1.4737	1.5729	1.939	2.37	1.4394	1.6433	1.9	2.382
	$\hat{\alpha}_{AUTP}$	0.9847	1.0003	1.003	0.963	0.9923	0.9761	0.963	1.003	0.9929	0.9782	1.002	0.958
	$\hat{\alpha}_{MTP}$ at k_1	1.5321	1.7232	1.909	2.55	1.4708	1.6249	2.071	2.617	1.4374	1.6832	2.004	2.621
	$\hat{\alpha}_{MTP}$ at k_2	1.5293	1.7102	1.918	2.592	1.479	1.6116	2.081	2.544	1.4319	1.6666	2.030	2.671
	$\hat{\alpha}_{MTP}$ at k_3	1.532	1.7237	1.911	2.555	1.47	1.6253	2.073	2.622	1.4376	1.6835	2.006	2.627

Table 3. Ratio of AMSE of OLS over various two-parameter estimators (p = 3 and $\beta = (1, 2, 5)'$)

ρ = 0.90													
n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
$\hat{\alpha}$	$\hat{\alpha}_{TP}$	1.1772	1.3769	1.554	1.637	1.1193	1.3248	1.511	1.58	1.113	1.3425	1.539	1.671
	$\hat{\alpha}_{LTE(3)}$	1.0047	1.0063	1.01	1.01	1.0007	1.0017	1.002	1.002	1.0012	1.0016	1.003	1.002
	$\hat{\alpha}_{NTP}$	1.1245	1.5709	2.098	2.572	1.0138	1.4679	2.218	2.588	1.0059	1.493	2.247	2.905
	$\hat{\alpha}_{AUTP}$	1.003	0.9369	0.92	0.93	0.9882	1.0002	0.975	0.978	1.002	0.9788	0.975	1.003

Table 3. Ratio of AMSE of OLS over various two-parameter estimators (p = 3 and $\beta = (1,2,5)'$) (cont.)

ρ = 0.90													
n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
$\hat{\alpha}$	$\hat{\alpha}_{MTP}$ at k_1	1.1551	1.6875	2.429	3.004	1.0116	1.4822	2.347	2.712	1.0033	1.5099	2.343	3.081
	$\hat{\alpha}_{MTP}$ at k_2	1.1336	1.5861	2.52	3.12	1.0017	1.4733	2.326	2.754	1.003	1.5067	2.268	2.972
	$\hat{\alpha}_{MTP}$ at k_3	1.1545	1.6884	2.437	3.016	1.0112	1.4823	2.349	2.715	1.0029	1.51	2.346	3.085
ρ = 0.95													
n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
$\hat{\alpha}$	$\hat{\alpha}_{TP}$	1.217	1.2942	1.4217	1.543	1.147	1.214	1.2792	1.4212	1.209	1.299	1.4227	1.559
	$\hat{\alpha}_{LTE(3)}$	1	1.0001	1.0006	1.001	1	1	1.0002	1.0006	1	1.0003	1.0007	1.001
	$\hat{\alpha}_{NTP}$	1.457	1.5275	1.8817	2.38	1.314	1.448	1.5177	1.8593	1.434	1.5311	1.834	2.445
	$\hat{\alpha}_{AUTP}$	1	0.9971	0.9921	1.002	1	1	0.9974	0.9929	1.020	0.9968	1.003	0.99
	$\hat{\alpha}_{MTP}$ at k_1	1.454	1.5328	1.9085	2.391	1.31	1.447	1.5282	1.8693	1.435	1.5322	1.845	2.508
	$\hat{\alpha}_{MTP}$ at k_2	1.454	1.5322	1.9105	2.398	1.32	1.457	1.5271	1.872	1.439	1.5394	1.8461	2.45
	$\hat{\alpha}_{MTP}$ at k_3	1.454	1.5327	1.9089	2.392	1.31	1.457	1.5281	1.8697	1.441	1.5322	1.8454	2.509
ρ = 0.99													
n		20				50				100			
$\hat{\sigma}^2$		1	5	10	25	1	5	10	25	1	5	10	25
$\hat{\alpha}$	$\hat{\alpha}_{TP}$	1.3596	1.534	1.624	1.65	1.2018	1.3461	1.561	1.654	1.1787	1.3767	1.577	1.575
	$\hat{\alpha}_{LTE(3)}$	1.0053	1.006	1.008	1.01	1.0033	1.005	1.008	1.008	1.004	1.0065	1.008	1.009
	$\hat{\alpha}_{NTP}$	1.5565	2.09	2.489	2.66	1.1904	1.5099	2.133	2.555	1.1288	1.5669	2.173	2.31
	$\hat{\alpha}_{AUTP}$	1.001	1.002	0.933	0.93	0.9607	0.9514	1.002	0.933	0.9541	1.003	0.927	0.927
	$\hat{\alpha}_{MTP}$ at k_1	1.6313	2.376	2.972	3.24	1.2261	1.592	2.394	2.949	1.1592	1.638	2.499	2.843
	$\hat{\alpha}_{MTP}$ at k_2	1.6133	2.379	2.989	3.42	1.2051	1.5716	2.41	3.062	1.1312	1.6141	2.564	2.658
	$\hat{\alpha}_{MTP}$ at k_3	1.6317	2.381	2.984	3.26	1.2255	1.5926	2.399	2.957	1.1585	1.6385	2.506	2.854

From Tables 2 and 3 we observe that the performance of our proposed modified two-parameter estimator $\hat{\alpha}_{MTP}$ is better than $\hat{\alpha}_{OLS}$, $\hat{\alpha}_{TP}$, $\hat{\alpha}_{LTE(3)}$, and $\hat{\alpha}_{AUTP}$. At the same time $\hat{\alpha}_{MTP}$ perform equivalently and is slightly better than $\hat{\alpha}_{NTP}$ for all combinations of correlation between predictors (ρ), the numbers of explanatory variables (p), the sample size (n), the choice of the ridge parameter (k) and the variance of the error term (σ^2) used in this simulation study.

7. Conclusion

In this article a modified two-parameter estimator alternative to the OLS estimator is proposed for estimating the regression parameter in the presence of multicollinearity. The performance of the proposed estimator is evaluated in terms of scalar mean-squared error criterion. Through the simulation study the performance of the proposed estimator is evaluated, for different combinations of ρ, p, n, k and σ^2 over the OLS and other two-parameter estimators reviewed in this article. Finally, it is found that the performance of the proposed estimator is satisfactory over the other estimators in the presence of multicollinearity.

Acknowledgements

The author is very grateful to the reviewers and the editor for their valuable comments and constructive suggestions which certainly improved the quality of the paper in the present version.

REFERENCES

AKDENIZ, F., EROL, H., (2003). Mean squared error matrix comparisons of some biased estimators in linear regression. Commun. Statist. Theor. Meth. 32(23), 89–2413.

AKDENIZ, F., KACIRANLAR, S., (1995). On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. Commun. Statist. Theor. Meth. 24, 1789–1797.

BATAH, F. S., RAMNATHAN, T., GORE, S. D., (2008). The efficiency of modified jackknife and ridge type regression estimators: a comparison. Surveys in Mathematics and its Applications 24(2), 157–174.

BUONACCORSI, J. P., (1996). A Modified estimating equation approach to correcting for measurement error in regression. Biometrika 83, 433–440.

CROUSE, R. H., JIN, C., HANUMARA, R. C., (1995). Unbiased ridge estimation with prior information and ridge trace. Commun. Statist. Theor. Meth. 24, 2341–2354.

DORUGADE, A. V., KASHID, D. N., (2011). Parameter estimation method in Ridge Regression. Interstat May 2011.

HOERL, A. E., KENNARD, R. W., (1970a). Ridge regression: biased estimation for nonorthogonal problems. Tech. 12, 55–67.

HOERL, A. E., KENNARD, R. W., BALDWIN, K. F., (1975). Ridge regression: Some Simulations. Commun. Statist. 4, 105–123.

KACIRANLAR, S., SAKALLIOGLU, S., AKDENIZ, F., STYAN, G. P. H., WERNER, H. J., (1999). A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland Cement. Sankhya Ind. J. Statist. 61, 443–459.

KADIYALA, K., (1984). A class almost unbiased and efficient estimators of regression coefficients. Econom. Lett. 16, 293–296.

KIBRIA, B. M., (2003). Performance of some new ridge regression estimators. Commun. Statist. –Simulation 32 (2), 419–435.

LIU, K., (1993). A new class of biased estimate in linear regression. Commun. Statist. Theor. Meth. 22, 393–402.

MASSY, W. F., (1965). Principal components regression in exploratory statistical research. JASA 60, 234–266.

NOMURA, M., (1988). On the almost unbiased ridge regression estimation. Commun. Statist. –Simulation 17(3), 729–743.

MAYER, L. S., WILLKE, T. A., (1973). On biased estimation in linear models. Technometrics 15, 497–508.

MONTGOMERY, D. C., PECK, E. A., VINING, G. G., (2006). Introduction to linear regression analysis. John Wiley and Sons, New York.

OZKALE, M. R., KACIRANLAR, S., (2007). The restricted and unrestricted two-parameter estimators. Commun. Statist. Theor. Meth. 36, 2707–2725.

OHTANI, K., (1986). On small sample properties of the almost unbiased generalized ridge estimator. Commun. Statist. Theor. Meth. 15, 1571–1578.

SAKALLIOGLU, S., KACIRANLAR, S., (2008). A new biased estimator based on ridge estimation. Stat Papers 49, 669–689.

STEIN, C., (1956). Inadmissibility of the usual estimator for mean of multivariate normal distribution. Proc. Third Berkeley Symp. Math. Statist. Probab. 1, 197–206.

SWINDEL, B. F., (1976). Good ridge estimators based on prior information. Commun. Statist. Theor. Meth. A5, 1065–1075.

SINGH, B., CHAUBEY, Y. P., (1987). On some improved ridge estimators. Stat Papers 28, 53–67.

YANG, H., CHANG, X., (2010). A New Two-Parameter Estimator in Linear Regression. Commun. Statist. Theor. Meth. 39, 923–934.

WU, J., YANG, H., (2011). Efficiency of an almost unbiased two-parameter estimator in linear regression model. Statistics 47(3), 535–545.