DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 9 | Issue 1 Article 22

5-1-2010

A New Biased Estimator Derived from Principal Component Regression Estimator

Set Foong Ng
Universiti Teknologi MARA, Malaysia, setfoong@yahoo.com

Heng Chin Low
Universiti Sains Malaysia, hclow@cs.usm.my

Soon Hoe Quah
Universiti Sains Malaysia, shquah@gmail.com

Journal of Modern Applied Statistical Methods
May 2010, Vol. 9, No. 1, 227-234

1538 – 9472/10/$95.00

A New Biased Estimator Derived from Principal Component Regression Estimator

Set Foong Ng
Universiti Teknologi MARA, Malaysia

Heng Chin Low   Soon Hoe Quah
Universiti Sains Malaysia

A new biased estimator obtained by combining the Principal Component Regression Estimator and the special case of Liu-type estimator is proposed. The properties of the new estimator are derived and comparisons between the new estimator and other estimators in terms of mean squared error are presented.

Key words: Linear regression model, Principal Component Regression Estimator, special case of Liu-type estimator, mean squared error.

Introduction

Multicollinearity is one of the problems faced in linear regression models. When multicollinearity is detected in data and the regressors that caused it are identified, one solution is to eliminate the regressors that are causing the multicollinearity. However, deleting regressors is not a safe strategy: there is no warning for extrapolation and there is no data to support a prediction in the region away from the multicollinearity.

Principal component regression is an alternative to regression deletion. Principal component regression is one type of biased regression method and its purpose is to eliminate those dimensions (which usually correspond to very small eigenvalues) causing the multicollinearity problem. Principal component regression approaches the problem of multicollinearity by dropping the dimension defined by a linear combination of the independent variables but not by a single independent variable (Rawlings, et al., 1998). Thus, the Principal Component Regression Estimator (Massy, 1965; Marquardt, 1970; Hawkins, 1973; Greenberg, 1975) is a biased alternative to the unbiased Ordinary Least Squares Estimator in the presence of multicollinearity.

Motivated by the idea of combining different estimators that might produce a better estimator, the r-k Class Estimator was proposed by Baye & Parker (1984). It has been shown that theoretical gains exist from combining the principal component regression and the ridge regression techniques. In addition, Kaciranlar and Sakallioglu (2001) proposed the r-d Class Estimator by combining the Liu Estimator and the Principal Component Regression Estimator.

A linear regression model with a dependent variable and p independent variables is given by

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{1.1}$$

where $\mathbf{Y}$ is an n×1 standardized observed random vector, $\mathbf{Z}$ is an n×p standardized known matrix with p independent variables, $\boldsymbol{\gamma}$ is an p×1 vector of parameter and $\boldsymbol{\varepsilon}$ is an n×1 vector of errors such that $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$.

If the matrix $\boldsymbol{\lambda}$ is a $p \times p$ diagonal matrix whose diagonal elements are eigenvalues of $\mathbf{Z}'\mathbf{Z}$, where the eigenvalues of $\mathbf{Z}'\mathbf{Z}$ are denoted by $\lambda_1, \lambda_2, ..., \lambda_p$, and if the matrix $\mathbf{T}$ is

Set Foong Ng is a Lecturer in the Department of Computer and Mathematical Sciences. Email him at: setfoong@yahoo.com. Heng Chin Low is a Professor in the School of Mathematical Sciences. Email him at: hclow@cs.usm.my. Soon Hoe Quah is a Professor in the School of Mathematical Sciences. Email him at: shquah@gmail.com.

a $p \times p$ orthonormal matrix consisting of the p eigenvectors of $\mathbf{Z}'\mathbf{Z}$, then the Principal Component Regression Estimator of parameter $\boldsymbol{\gamma}$ is given by

$$\hat{\boldsymbol{\gamma}}_r = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r)^{-1}\mathbf{T}_r'\mathbf{Z}'\mathbf{Y}, \qquad (1.2)$$

where r < p, $\mathbf{T}_r$ are the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after deleting p-r of the columns of $\mathbf{T}$. This satisfies $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$.

From the Liu-type estimator proposed by Liu (2003), a special case of Liu-type estimator was derived by Ng, et al. (2007)

$$\hat{\boldsymbol{\gamma}}_c = (\mathbf{Z}'\mathbf{Z} + c\mathbf{I})^{-1}(\mathbf{Z}'\mathbf{Y} + \hat{\boldsymbol{\gamma}}), \qquad (1.3)$$

where $c > 1$, $\hat{\boldsymbol{\gamma}} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}$ is the Ordinary Least Squares Estimator.

Methodology

In this article, a new estimator was derived by combining the advantage of the principal component regression, $\hat{\boldsymbol{\gamma}}_r$, and the special case of the Liu-type Estimator, $\hat{\boldsymbol{\gamma}}_c$. Here, the estimator, $\hat{\boldsymbol{\gamma}}$, in $\hat{\boldsymbol{\gamma}}_c = (\mathbf{Z}'\mathbf{Z} + c\mathbf{I})^{-1}(\mathbf{Z}'\mathbf{Y} + \hat{\boldsymbol{\gamma}})$ was substituted by the Principal Component Regression Estimator, $\hat{\boldsymbol{\gamma}}_r$. Hence, a new expression of $\hat{\boldsymbol{\gamma}}_c$ was obtained, that is

$$\hat{\boldsymbol{\gamma}}_c(new) = (\mathbf{Z}'\mathbf{Z} + c\mathbf{I})^{-1}(\mathbf{Z}'\mathbf{Y} + \hat{\boldsymbol{\gamma}}_r). \qquad (1.4)$$

Note that the eigenvalues and the eigenvectors are ordered so that $\lambda_1 > \lambda_2 > ... > \lambda_p$. The purpose of principal component regression is to eliminate those dimensions that usually correspond to eigenvalues that are very small. Thus, the concept of principal component regression eliminates p-r of the columns of $\mathbf{T}$ which correspond to the smallest p-r eigenvalues. Hence, $\mathbf{T}_r = [\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_r]$ is the matrix of the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ while $\boldsymbol{\lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_r)$ is the matrix of the remaining eigenvalues of $\mathbf{Z}'\mathbf{Z}$ after deleting $p - r$ of the columns of $\mathbf{T}$; once again, this satisfies $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_r)$.

By including the matrix $\mathbf{T}_r$ in the new expression of $\hat{\boldsymbol{\gamma}}_c$, $\hat{\boldsymbol{\gamma}}_c(new) = (\mathbf{Z}'\mathbf{Z} + c\mathbf{I})^{-1}(\mathbf{Z}'\mathbf{Y} + \hat{\boldsymbol{\gamma}}_r)$, a new biased estimator, $\hat{\boldsymbol{\gamma}}_r(c)$, was obtained which is given by:

$$\hat{\boldsymbol{\gamma}}_r(c) = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r + c\mathbf{I}_r)^{-1}(\mathbf{T}_r'\mathbf{Z}'\mathbf{Y} + \mathbf{T}_r'\hat{\boldsymbol{\gamma}}_r), \qquad (1.5)$$

where $c > 1$, $r < p$, $\hat{\boldsymbol{\gamma}}_r = \mathbf{T}_r(\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r)^{-1}\mathbf{T}_r'\mathbf{Z}'\mathbf{Y}$, $\mathbf{T}_r$ are the remaining eigenvectors of $\mathbf{Z}'\mathbf{Z}$ after deleting p-r of the columns of $\mathbf{T}$. This satisfies $\mathbf{T}_r'\mathbf{Z}'\mathbf{Z}\mathbf{T}_r = \boldsymbol{\lambda}_r = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$.

Properties of the New Estimator

The properties of the proposed new estimator are as follows:

(1) Bias of $\hat{\gamma}_r(c)$:

$$\mathrm{bias}(\hat{\boldsymbol{\gamma}}_r(c)) = -\mathbf{T}_{p-r}\mathbf{T}_{p-r}'\boldsymbol{\gamma} - \mathbf{T}_r(\boldsymbol{\lambda}_r + c\mathbf{I}_r)^{-1}(c-1)\mathbf{T}_r'\boldsymbol{\gamma} \qquad (2.1)$$

(2) Variance-covariance matrix of $\hat{\boldsymbol{\gamma}}_r(c)$:

$$\mathrm{Var}(\hat{\boldsymbol{\gamma}}_r(c)) = [\mathbf{T}_r(\boldsymbol{\lambda}_r + c\mathbf{I}_r)^{-1}(\boldsymbol{\lambda}_r + \mathbf{I}_r)\mathbf{T}_r']^2\mathbf{T}_r\boldsymbol{\lambda}_r^{-1}\mathbf{T}_r'\sigma^2 \qquad (2.2)$$

(3) Mean squared error of $\hat{\gamma}_r(c)$:

$$\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) = \sum_{j=1}^{r}\left[\frac{(\lambda_j + 1)^2\sigma^2}{\lambda_j(\lambda_j + c)^2}\right] + \sum_{j=r+1}^{p}\gamma_j^2 + \sum_{j=1}^{r}\left[\frac{(c-1)^2\gamma_j^2}{(\lambda_j + c)^2}\right] \qquad (2.3)$$

(4) When r = p, the new estimator, $\hat{\boldsymbol{\gamma}}_r(c)$, is equal to the special case of Liu-type estimator, $\hat{\boldsymbol{\gamma}}_c$, that is,

$$\hat{\boldsymbol{\gamma}}_p(c) = \hat{\boldsymbol{\gamma}}_c. \qquad (2.4)$$

(5) When r = p and c = 1, the new estimator, $\hat{\boldsymbol{\gamma}}_r(c)$, is equal to the Ordinary Least Squares Estimator, $\hat{\boldsymbol{\gamma}}$, that is,

$$\hat{\boldsymbol{\gamma}}_p(1) = \hat{\boldsymbol{\gamma}}. \qquad (2.5)$$

(6) When c = 1, the new estimator, $\hat{\boldsymbol{\gamma}}_r(c)$, is equal to the Principal Component Regression Estimator, $\hat{\boldsymbol{\gamma}}_r$, that is

$$\hat{\boldsymbol{\gamma}}_r(1) = \hat{\boldsymbol{\gamma}}_r. \qquad (2.6)$$

Results

The new estimator, $\hat{\boldsymbol{\gamma}}_r(c)$, was compared with the special case of Liu-type estimator, $\hat{\boldsymbol{\gamma}}_c$, the Ordinary Least Squares Estimator, $\hat{\boldsymbol{\gamma}}$, and the Principal Component Regression Estimator, $\hat{\boldsymbol{\gamma}}_r$, in terms of mean squared error in order to evaluate the performance of the new estimator.

The mean squared errors $\hat{\boldsymbol{\gamma}}_c$, $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\gamma}}_r$ are shown in Equations (3.1) to (3.3), respectively:

$$\operatorname{mse}(\hat{\boldsymbol{\gamma}}_c) = \sum_{j=1}^{p}\left[\frac{(\lambda_j+1)^2\sigma^2}{\lambda_j(\lambda_j+c)^2} + \frac{(c-1)^2\gamma_j^2}{(\lambda_j+c)^2}\right] \qquad (3.1)$$

$$\operatorname{mse}(\hat{\boldsymbol{\gamma}}) = \sum_{j=1}^{p}\frac{\sigma^2}{\lambda_j} \qquad (3.2)$$

$$\operatorname{mse}(\hat{\boldsymbol{\gamma}}_r) = \sum_{j=1}^{r}\frac{\sigma^2}{\lambda_j} + \sum_{j=r+1}^{p}\gamma_j^2 \qquad (3.3)$$

From the properties of the new estimator, $\hat{\boldsymbol{\gamma}}_r(c)$ is equivalent to $\hat{\boldsymbol{\gamma}}_c$ when r = p.

Theorem 3.1: Comparison Between the $\hat{\boldsymbol{\gamma}}_r(c)$ and $\hat{\boldsymbol{\gamma}}_{r+1}(c)$

If $\sigma^2 > \gamma_{r+1}^2\lambda_{r+1}$,

(a)

$$\operatorname{mse}(\hat{\boldsymbol{\gamma}}_{r+1}(c)) > \operatorname{mse}(\hat{\boldsymbol{\gamma}}_r(c)) \text{ for } 1 < c < a_1,$$

(b)

$$\operatorname{mse}(\hat{\boldsymbol{\gamma}}_{r+1}(c)) < \operatorname{mse}(\hat{\boldsymbol{\gamma}}_r(c)) \text{ for } c > a_1,$$

where

$$a_1 = \frac{(\lambda_{r+1}+1)^2\sigma^2 + \lambda_{r+1}\gamma_{r+1}^2(1-\lambda_{r+1}^2)}{2\lambda_{r+1}\gamma_{r+1}^2(1+\lambda_{r+1})}.$$

Theorem 3.1 Proof

Consider the difference between the mean squared errors of $\hat{\boldsymbol{\gamma}}_r(c)$ and $\hat{\boldsymbol{\gamma}}_{r+1}(c)$:

$$\mathrm{D}_1 = \operatorname{mse}(\hat{\boldsymbol{\gamma}}_{r+1}(c)) - \operatorname{mse}(\hat{\boldsymbol{\gamma}}_r(c))$$

$$= \sum_{j=1}^{r+1}\left[\frac{(\lambda_j+1)^2\sigma^2}{\lambda_j(\lambda_j+c)^2}\right] + \sum_{j=r+2}^{p}\gamma_j^2 + \sum_{j=1}^{r+1}\left[\frac{(c-1)^2\gamma_j^2}{(\lambda_j+c)^2}\right] - \sum_{j=1}^{r}\left[\frac{(\lambda_j+1)^2\sigma^2}{\lambda_j(\lambda_j+c)^2}\right] - \sum_{j=r+1}^{p}\gamma_j^2 - \sum_{j=1}^{r}\left[\frac{(c-1)^2\gamma_j^2}{(\lambda_j+c)^2}\right]$$

$$= \frac{(\lambda_{r+1}+1)^2\sigma^2}{\lambda_{r+1}(\lambda_{r+1}+c)^2} + \frac{(c-1)^2\gamma_{r+1}^2}{(\lambda_{r+1}+c)^2} - \gamma_{r+1}^2$$

$$= \frac{(\lambda_{r+1}+1)^2\sigma^2 + \lambda_{r+1}(c-1)^2\gamma_{r+1}^2 - \lambda_{r+1}(\lambda_{r+1}+c)^2\gamma_{r+1}^2}{\lambda_{r+1}(\lambda_{r+1}+c)^2} \qquad (3.4)$$

Thus, $\mathrm{D}_1 > 0$ if and only if

$$(\lambda_{r+1}+1)^2\sigma^2 + \lambda_{r+1}(c-1)^2\gamma_{r+1}^2 - \lambda_{r+1}(\lambda_{r+1}+c)^2\gamma_{r+1}^2 > 0. \qquad (3.5)$$

The solution for the inequality (3.5) is

$$1 < c < a_1, \qquad (3.6)$$

where

$$a_1 = \frac{(\lambda_{r+1}+1)^2\sigma^2 + \lambda_{r+1}\gamma_{r+1}^2(1-\lambda_{r+1}^2)}{2\lambda_{r+1}\gamma_{r+1}^2(1+\lambda_{r+1})}.$$

Because $c > 1$, it requires $a_1 > 1$, that is

$$a_1 > 1$$

$$\frac{(\lambda_{r+1}+1)^2\sigma^2 + \lambda_{r+1}\gamma_{r+1}^2(1-\lambda_{r+1}^2)}{2\lambda_{r+1}\gamma_{r+1}^2(1+\lambda_{r+1})} > 1$$

$$\sigma^2 > \gamma_{r+1}^2\lambda_{r+1}. \quad (3.7)$$

Thus, if $\sigma^2 > \gamma_{r+1}^2\lambda_{r+1}$, $\text{mse}(\hat{\boldsymbol\gamma}_{r+1}(c)) > \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $1 < c < a_1$. Similarly, $\text{mse}(\hat{\boldsymbol\gamma}_{r+1}(c)) < \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $c > a_1$. Hence, the proof for Theorem 3.1 is completed.

Theorem 3.2: Comparison between the New Estimator, $\hat{\boldsymbol\gamma}_r(c)$, and the Special Case of Liu-Type Estimator, $\hat{\boldsymbol\gamma}_c$

If $\sigma^2 > \gamma_j^2\lambda_j$ for $j \in \{r+1, r+2, \cdots, p\}$,

(a)

$\text{mse}(\hat{\boldsymbol\gamma}_c) > \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $1 < c < \min\{(a_2)_j\}$,

(b)

$\text{mse}(\hat{\boldsymbol\gamma}_c) < \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $c > \max\{(a_2)_j\}$,

where

$$(a_2)_j = \frac{(\lambda_j+1)^2\sigma^2 + \lambda_j\gamma_j^2(1-\lambda_j^2)}{2\lambda_j\gamma_j^2(1+\lambda_j)}$$

for

$$j \in \{r+1, r+2, \cdots, p\}.$$

Theorem 3.2 Proof

From Theorem 3.1(a), $\text{mse}(\hat{\boldsymbol\gamma}_{r+1}(c)) > \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $1 < c < a_1$, where $a_1 = \frac{(\lambda_{r+1}+1)^2\sigma^2 + \lambda_{r+1}\gamma_{r+1}^2(1-\lambda_{r+1}^2)}{2\lambda_{r+1}\gamma_{r+1}^2(1+\lambda_{r+1})}$. Thus, the $\text{mse}(\hat{\boldsymbol\gamma}_p(c)) > ... > \text{mse}(\hat{\boldsymbol\gamma}_{r+1}(c)) > \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $1 < c < \min\{(a_2)_j\}$, where $(a_2)_j = \frac{(\lambda_j+1)^2\sigma^2 + \lambda_j\gamma_j^2(1-\lambda_j^2)}{2\lambda_j\gamma_j^2(1+\lambda_j)}$ for $j \in \{r+1, r+2, \cdots, p\}$. From the properties of the new estimator, when r = p, $\hat{\boldsymbol\gamma}_p(c) = \hat{\boldsymbol\gamma}_c$. Thus, $\text{mse}(\hat{\boldsymbol\gamma}_c) > \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $1 < c < \min\{(a_2)_j\}$. Similarly, $\text{mse}(\hat{\boldsymbol\gamma}_c) < \text{mse}(\hat{\boldsymbol\gamma}_r(c))$ for $c > \max\{(a_2)_j\}$. The proof for Theorem 3.2 is completed.

Theorem 3.3: Comparison between the New (Biased) Estimator and the Unbiased Ordinary Least Squares Estimator

(a) If $\sigma^2 > \gamma_j^2\lambda_j$ for all $j \in \{1, 2, \cdots, r\}$ and $\sum_{j=r+1}^{p}\frac{\gamma_j^2\lambda_j - \sigma^2}{\lambda_j} \le 0$, $\text{mse}(\hat{\boldsymbol\gamma}_r(c)) < \text{mse}(\hat{\boldsymbol\gamma})$ for $c > 1$.

(b) If $\sigma^2 < \gamma_j^2\lambda_j$ for some $j \in \{1, 2, \cdots, r\}$ and $\sum_{j=r+1}^{p}\frac{\gamma_j^2\lambda_j - \sigma^2}{\lambda_j} \le 0$, $\text{mse}(\hat{\boldsymbol\gamma}_r(c)) < \text{mse}(\hat{\boldsymbol\gamma})$ for $1 < c < \min\{(a_3)_j\}$, $\text{mse}(\hat{\boldsymbol\gamma}_r(c)) > \text{mse}(\hat{\boldsymbol\gamma})$ for $c > \max\{(a_3)_j\}$, where $(a_3)_j = \frac{\lambda_j\gamma_j^2 + 2\lambda_j\sigma^2 + \sigma^2}{\lambda_j\gamma_j^2 - \sigma^2}$ for $j \in \{1, 2, \cdots, r\}$.

Theorem 3.3 Proof

Consider the difference between the mean squared errors of $\hat{\boldsymbol\gamma}_r(c)$ and $\hat{\boldsymbol\gamma}$:

$$\text{D}_2 = \text{mse}(\hat{\boldsymbol\gamma}_r(c)) - \text{mse}(\hat{\boldsymbol\gamma})$$

$$= \sum_{j=1}^{r}\left[\frac{(\lambda_j+1)^2\sigma^2}{\lambda_j(\lambda_j+c)^2}\right] + \sum_{j=r+1}^{p}\gamma_j^2 + \sum_{j=1}^{r}\left[\frac{(c-1)^2\gamma_j^2}{(\lambda_j+c)^2}\right] - \sum_{j=1}^{p}\frac{\sigma^2}{\lambda_j}$$

$$= \sum_{j=1}^{r}\left[\frac{(\lambda_j+1)^2\sigma^2 + \lambda_j(c-1)^2\gamma_j^2 - \sigma^2(\lambda_j+c)^2}{\lambda_j(\lambda_j+c)^2}\right] + \sum_{j=r+1}^{p}\frac{\gamma_j^2\lambda_j - \sigma^2}{\lambda_j} \quad (3.8)$$

Thus, $\mathrm{D}_2 < 0$ if and only if

$$\sum_{j=r+1}^{p} \frac{\gamma_j^2 \lambda_j - \sigma^2}{\lambda_j} \le 0 \tag{3.9}$$

and

$$(\lambda_j + 1)^2 \sigma^2 + \lambda_j (c-1)^2 \gamma_j^2 - \sigma^2 (\lambda_j + c)^2 < 0 . \tag{3.10}$$

The inequality (3.10) can also be written as

$$(\gamma_j^2 \lambda_j - \sigma^2) c^2 - 2(\gamma_j^2 \lambda_j + \lambda_j \sigma^2) c + (\gamma_j^2 \lambda_j + 2\lambda_j \sigma^2 + \sigma^2) < 0 \tag{3.11}$$

Solving the equation

$$\begin{aligned}(\gamma_j^2 \lambda_j - \sigma^2) c^2 - 2(\gamma_j^2 \lambda_j + \lambda_j \sigma^2) c \\ + (\gamma_j^2 \lambda_j + 2\lambda_j \sigma^2 + \sigma^2) = 0\end{aligned} \tag{3.12}$$

the solutions $c = \frac{\lambda_j \gamma_j^2 + 2\lambda_j \sigma^2 + \sigma^2}{\lambda_j \gamma_j^2 - \sigma^2}$ and $c = 1$ are obtained.

Let $(a_3)_j = \frac{\lambda_j \gamma_j^2 + 2\lambda_j \sigma^2 + \sigma^2}{\lambda_j \gamma_j^2 - \sigma^2}$ where the values of $(a_3)_j$ could be less than 1 or greater than 1. The condition requiring for $(a_3)_j < 1$ is given by

$$(a_3)_j < 1$$

$$\frac{\lambda_j \gamma_j^2 + 2\lambda_j \sigma^2 + \sigma^2}{\lambda_j \gamma_j^2 - \sigma^2} < 1$$

$$\frac{2\lambda_j \sigma^2 + 2\sigma^2}{\lambda_j \gamma_j^2 - \sigma^2} < 0$$

$$\lambda_j \gamma_j^2 - \sigma^2 < 0$$

$$\sigma^2 > \gamma_j^2 \lambda_j . \tag{3.13}$$

Similarly, the condition requiring for $(a_3)_j > 1$ is given by

$$\sigma^2 < \gamma_j^2 \lambda_j . \tag{3.14}$$

For the first situation where $(a_3)_j < 1$ for all $j \in \{1, 2, \cdots, r\}$, the solution for the inequality (3.11) is

$$c > 1 \text{ if } \sigma^2 > \gamma_j^2 \lambda_j \text{ for all } j \in \{1, 2, \cdots, r\} . \tag{3.15}$$

Thus, $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) < \mathrm{mse}(\hat{\boldsymbol{\gamma}})$ for $c > 1$ if $\sigma^2 > \gamma_j^2 \lambda_j$ for all $j \in \{1, 2, \cdots, r\}$ and $\sum_{j=r+1}^{p} \frac{\gamma_i^2 \lambda_i - \sigma^2}{\lambda_i} \le 0$. The proof for Theorem 3.3(a) is completed.

For the second situation where $(a_3)_j > 1$ for some $j \in \{1, 2, \cdots, r\}$, the solution for the inequality (3.11) is: $1 < c < \min\{(a_3)_j\}$ if $\sigma^2 < \gamma_j^2 \lambda_j$ for some $j \in \{1, 2, \cdots, r\}$, where $(a_3)_j = \frac{\lambda_j \gamma_j^2 + 2\lambda_j \sigma^2 + \sigma^2}{\lambda_j \gamma_j^2 - \sigma^2}$ for $j \in \{1, 2, \cdots, r\}$.

Thus, $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) < \mathrm{mse}(\hat{\boldsymbol{\gamma}})$ for $1 < c < \min\{(a_3)_j\}$ and $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) > \mathrm{mse}(\hat{\boldsymbol{\gamma}})$ for $c > \max\{(a_3)_j\}$ if $\sigma^2 < \gamma_j^2 \lambda_j$ for some $j \in \{1, 2, \cdots, r\}$ and $\sum_{j=r+1}^{p} \frac{\gamma_j^2 \lambda_j - \sigma^2}{\lambda_j} \le 0$. The proof for Theorem 3.3(b) is completed.

Theorem 3.4: Comparison between the New Estimator and the Principal Component Regression Estimator in terms of Mean Squared Error

(a) If $\sigma^2 > \gamma_j^2 \lambda_j$ for all $j \in \{1, 2, \cdots, r\}$, $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) < \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r)$ for $c > 1$.

(b) If $\sigma^2 < \gamma_j^2 \lambda_j$ for some $j \in \{1, 2, \cdots, r\}$, $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) < \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r)$ for $1 < c < \min\{(a_3)_j\}$, $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) > \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r)$ for $c > \max\{(a_3)_j\}$, where

$$(a_3)_j = \frac{\lambda_j \gamma_j^2 + 2\lambda_j \sigma^2 + \sigma^2}{\lambda_j \gamma_j^2 - \sigma^2} \quad \text{for } j \in \{1, 2, \cdots, r\}.$$

Theorem 3.4 Proof

The first derivative of $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c))$ with respect to c is given by

$$\frac{\mathrm{d}}{\mathrm{d}c}\left[\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c))\right] = \frac{\mathrm{d}}{\mathrm{d}c}\left\{ \sum_{j=1}^{r}\left[\frac{(\lambda_j+1)^2\sigma^2}{\lambda_j(\lambda_j+c)^2}\right] + \sum_{j=r+1}^{p}\gamma_j^2 + \sum_{j=1}^{r}\left[\frac{(c-1)^2\gamma_j^2}{(\lambda_j+c)^2}\right] \right\}$$

$$= 2\sum_{j=1}^{r}\left[\frac{-\sigma^2(\lambda_j+1)^2}{\lambda_j(\lambda_j+c)^3} + \frac{(c-1)(\lambda_j+1)\gamma_j^2}{(\lambda_j+c)^3}\right]. \tag{3.16}$$

When $c = 1$,

$$\frac{\mathrm{d}}{\mathrm{d}c}\left[\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c))\right] = -2\sigma^2\sum_{j=1}^{r}\left[\frac{1}{\lambda_j(\lambda_j+1)}\right] < 0. \tag{3.17}$$

It was found that the function of $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c))$ has a negative slope of approximately $c = 1$. Furthermore, the new estimator, $\hat{\boldsymbol{\gamma}}_r(c)$, is equivalent to the Principal Component Regression Estimator, $\hat{\boldsymbol{\gamma}}_r$, when $c = 1$. Thus, $c > 1$ exists such that $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) < \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r)$.

Consider the difference between the mean squared errors of $\hat{\boldsymbol{\gamma}}_r(c)$ and $\hat{\boldsymbol{\gamma}}_r$:

$$\begin{aligned} \mathrm{D}_3 &= \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) - \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r) \\ &= \sum_{j=1}^{r}\left[\frac{(\lambda_j+1)^2\sigma^2}{\lambda_j(\lambda_j+c)^2}\right] + \sum_{j=r+1}^{p}\gamma_j^2 \\ &\quad + \sum_{j=1}^{r}\left[\frac{(c-1)^2\gamma_j^2}{(\lambda_j+c)^2}\right] - \sum_{j=1}^{r}\frac{\sigma^2}{\lambda_j} - \sum_{j=r+1}^{p}\gamma_j^2 \\ &= \sum_{j=1}^{r}\left[\frac{(\lambda_j+1)^2\sigma^2 + \lambda_j(c-1)^2\gamma_j^2 - \sigma^2(\lambda_j+c)^2}{\lambda_j(\lambda_j+c)^2}\right] \end{aligned} \tag{3.18}$$

Thus, $\mathrm{D}_3 < 0$ if and only if

$$(\lambda_j+1)^2\sigma^2 + \lambda_j(c-1)^2\gamma_j^2 - \sigma^2(\lambda_j+c)^2 < 0. \tag{3.19}$$

The inequality (3.19) can also be written as

$$(\gamma_j^2\lambda_j - \sigma^2)c^2 - 2(\gamma_j^2\lambda_j + \lambda_j\sigma^2)c + (\gamma_j^2\lambda_j + 2\lambda_j\sigma^2 + \sigma^2) < 0 \tag{3.20}$$

The solution for the inequality (3.20) is the same as the solution for the inequality (3.11). Thus, $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) < \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r)$ for $c > 1$ if $\sigma^2 > \gamma_j^2\lambda_j$ for all $j \in \{1, 2, \cdots, r\}$.

By contrast, if $\sigma^2 < \gamma_j^2\lambda_j$ for some $j \in \{1, 2, \cdots, r\}$, $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) < \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r)$ for $1 < c < \min\{(a_3)_j\}$ and $\mathrm{mse}(\hat{\boldsymbol{\gamma}}_r(c)) > \mathrm{mse}(\hat{\boldsymbol{\gamma}}_r)$ for $c > \max\{(a_3)_j\}$ where

$$(a_3)_j = \frac{\lambda_j \gamma_j^2 + 2\lambda_j \sigma^2 + \sigma^2}{\lambda_j \gamma_j^2 - \sigma^2} \quad \text{for } j \in \{1, 2, \cdots, r\}.$$

The proof for Theorem 3.4 is completed.

Numerical Example

A numerical example illustrates Theorems 3.2, 3.3 and 3.4. The data set is from Ryan (1997, pp. 402-403). The data consists of one dependent variable and six independent variables. The regression model with standardized variables is

$$\mathbf{Y} = \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{4.1}$$

where $\mathbf{Y}$ is a 50×1 standardized observed random vector, $\mathbf{Z}$ is a 50×6 standardized known matrix with six independent variables, $\boldsymbol{\gamma}$ is a 6×1 vector of parameters and $\boldsymbol{\varepsilon}$ is a 50×1 vector of errors.

Multicollinearity diagnostic indicates the presence of multicollinearity in the data. The least squares estimates are given by $\hat{\gamma}_1$ = -5.218, $\hat{\gamma}_2$ = -0.376, $\hat{\gamma}_3$ = 8.869, $\hat{\gamma}_4$ = -1.755, $\hat{\gamma}_5$ = -0.320 and $\hat{\gamma}_6$ = -1.178. The estimated variance of the error term is given by $\hat{\sigma}^2 = 0.000655$ while the eigenvalues are given by λ_1 =5.80831,

$\lambda_2 = 0.11749$, $\lambda_3 = 0.04812$, $\lambda_4 = 0.02501$, $\lambda_5 = 0.00081$ and $\lambda_6 = 0.00025$.

For practical purpose, γ_i and σ^2 in Theorems 3.2, 3.3 and 3.4 are substituted by the estimated $\hat{\gamma}_i$ and $\hat{\sigma}^2$ in this numerical example. In this numerical example, the Principal Component Regression Estimator has the smallest mean squared error when $r = 4$. Thus, $p = 6$ and $r = 4$ are used throughout this example.

It was found that the condition for Theorem 3.2 is satisfied, that is, $\sigma^2 > \gamma_j^2 \lambda_j$ for $j \in \{5,6\}$, and $(a_2)_j = \{4.43639, 1.42839\}$ for $j \in \{5,6\}$ was obtained. Taking the value of $c = 1.3$ where $1 < c < \min\{(a_2)_j\}$, it was found that $\text{mse}(\hat{\boldsymbol{\gamma}}_{\mathbf{r}}(\mathbf{c})) = 5.62615$, that is, less than $\text{mse}(\hat{\boldsymbol{\gamma}}_c) = 6.21735$. This agrees with Theorem 3.2(a). On the other hand, taking $c = 4.5$ where $c > \max\{(a_2)_j\}$, it was found that $\text{mse}(\hat{\boldsymbol{\gamma}}_c) = 52.7144$ and $\text{mse}(\hat{\boldsymbol{\gamma}}_r(c)) = 53.1363$, thus, $\text{mse}(\hat{\boldsymbol{\gamma}}_c) < \text{mse}(\hat{\boldsymbol{\gamma}}_r(c))$. This is also in line with Theorem 3.2(b).

For Theorem 3.3, the value of $\sum_{j=5}^{6} \frac{\gamma_j^2 \lambda_j - \sigma^2}{\lambda_j} = -1.8931 < 0$. It was also found that the values for $\gamma_j^2 \lambda_j = \{158.13, 0.01662, 3.78485, 0.07705\}$ for $j = \{1, 2, 3, 4\}$. This shows that the condition for Theorem 3.3(a) is not satisfied since $\hat{\sigma}^2 < \hat{\gamma}_j^2 \lambda_j$ for $j = \{1, 2, 3, 4\}$. Thus, Theorem 3.3(b) will be used to illustrate Theorem 3.3.

The values of $(a_3)_j = \{1.00006, 1.09168, 1.00036, 1.01758\}$ for $j = \{1, 2, 3, 4\}$. Choosing $c = 1.00003$ where $1 < c < \min\{(a_3)_j\}$, $\text{mse}(\hat{\boldsymbol{\gamma}}_r(c)) = 1.536355$, was obtained that is, smaller than $\text{mse}(\hat{\boldsymbol{\gamma}}) = 3.42946$. By contrast, taking $c = 1.3$ where $c > \max\{(a_3)_j\}$, it was found that $\text{mse}(\hat{\boldsymbol{\gamma}}_r(c)) = 5.62615 > \text{mse}(\hat{\boldsymbol{\gamma}}) = 3.42946$. This agrees with Theorem 3.3.

For Theorem 3.4, it was found that $\hat{\sigma}^2 < \hat{\gamma}_j^2 \lambda_j$ for $j = \{1, 2, 3, 4\}$. Since the condition for Theorem 3.4(b) is satisfied, Theorem 3.4(b) was used to illustrate Theorem 3.4. Choosing $c = 1.00003$ where $1 < c < \min\{(a_3)_j\}$, $\text{mse}(\hat{\boldsymbol{\gamma}}_r(c)) = 1.536356$ and $\text{mse}(\hat{\boldsymbol{\gamma}}_r) = 1.536359$ were obtained. This shows that new estimator has a smaller mean squared error for $1 < c < \min\{(a_3)_j\}$. By contrast, taking $c = 1.3$ where $c > \max\{(a_3)_j\}$, it was found that $\text{mse}(\hat{\boldsymbol{\gamma}}_r(c)) = 5.62615 > \text{mse}(\hat{\boldsymbol{\gamma}}_r) = 1.536358$. This is in line with the Theorem 3.4. This numerical comparison is shown to be in line with the theoretical comparison.

Conclusion

The new biased estimator was obtained by combining the Principal Component Regression Estimator and the special case of Liu-type estimator. When certain conditions are satisfied, the new estimator has been shown to have smaller mean squared error compared to the special case of Liu-type estimator, the Ordinary Least Squares Estimator and the Principal Component Regression Estimator. The numerical comparison was also shown to be in line with the theoretical comparison.

In conclusion, the proposed new estimator was shown be an improvement in terms of reduction in mean squared error. Thus, the new estimator could be considered as an alternative to the unbiased Ordinary Least Squares Estimator when multicollinearity is detected in a linear regression model.

Acknowledgement

This research is supported by the Fundamental Research Grant Scheme.

References

Baye, M. R., & Parker, D. F. (1984). Combining ridge and Principal component regression: A money demand illustration. *Communications in Statistics-Theory and Methods*, *13*(*2*), 197-205.

Greenberg, E. (1975). Minimum variance properties of principal component regression. *Journal of the American Statistical Association*, *70*, 194-197.

Hawkins, D. M. (1973). On the investigation of alternative regressions by principal component analysis. *Applied Statistics*, *22*, 275-286.

Kaciranlar, S., & Sakallioglu, S. (2001). Combining the Liu estimator and the principal component regression estimator. *Communications in Statistics-Theory and Methods*, *30*(*12*), 2699-2705.

Liu, K. (2003). Using Liu-type estimator to combat collinearity. *Communications in Statistics-Theory and Methods*, *32*(*5*), 1009-1020.

Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation. *Technometrics*, *12*, 591-612.

Massy, W. F. (1965). Principal components regression in exploratory statistical research. *Journal of the American Statistical Association*, *60*, 234-246.

Ng, S. F., Low, H. C., & Quah, S. H. (2007). A study on the Liu-type estimator. *Proceedings of the Third IMT-GT Regional Conference on Mathematics, Statistics and Applications: Strengthening Regional Cooperation through the Mathematical Sciences.* Penang: Universiti Sains Malaysia.

Rawlings, J. O., Pantula, S. G. & Dickey, D. A. (1998). *Applied Regression Analysis-A Research Tool.* New Jersey: Springer-Verlag.

Ryan, T. P. (1997). *Modern Regression Methods*. NY: John Wiley & Sons, Inc.