Journal of the Association of Arab Universities for Basic and Applied Sciences (2016) 20, 89–99

University of Bahrain
Journal of the Association of Arab Universities for Basic and Applied Sciences
www.elsevier.com/locate/jaaubas
www.sciencedirect.com

ORIGINAL ARTICLE

Some new modifications of Kibria's and Dorugade's methods: An application to Turkish GDP data

Adnan Karaibrahimoğlu [a], Yasin Asar [b,*], Aşır Genç [c]

[a] Meram Faculty of Medicine, Medical Education and Informatics Department, Biostatistics Unit, Necmettin Erbakan University, Konya, Turkey
[b] Department of Mathematics-Computer Sciences, Faculty of Science, Necmettin Erbakan University, Konya, Turkey
[c] Department of Statistics, Faculty of Science, Selçuk University, Konya, Turkey

Received 19 May 2014; revised 11 August 2014; accepted 31 August 2014
Available online 7 October 2014

KEYWORDS

Multicollinearity;
Multiple linear regression;
Ridge regression;
Ridge estimator;
Monte Carlo simulation

Abstract In multiple linear regression analysis, multicollinearity is an important problem. Ridge regression is one of the most commonly used methods to overcome this problem. There are many proposed ridge parameters in the literature. In this paper, we propose some new modifications to choose the ridge parameter. A Monte Carlo simulation is used to evaluate parameters. Also, biases of the estimators are considered. The mean squared error is used to compare the performance of the proposed estimators with others in the literature. According to the results, all the proposed estimators are superior to ordinary least squared estimator (OLS).

1. Introduction

Consider the following standard linear regression model

$$Y = X\beta + \varepsilon \tag{1.1}$$

where Y is an $n \times 1$ vector of dependent variable, X is a design matrix of order $n \times p$ where p is the number of explanatory variables, β is a $p \times 1$ vector of coefficients and ε is the error vector of order $n \times 1$ distributed as $N(0, \sigma^2 I_n)$. Ordinary least squared (OLS) method is the most common method of estimating β and the OLS estimator of β is given as follows

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{1.2}$$

In some situations, the matrix $X'X$ has almost zero eigenvalues meaning the explanatory variables are correlated. This leads to a large variance and so large mean squared error (MSE). Thus one may not reach a reliable solution for β. This is the commonly faced problem called multicollinearity. There are various methods to solve this problem. The ridge regression is one of the most popular methods proposed by Hoerl and Kennard (1970a,b).

In ridge regression, adding a small positive number $k(k > 0)$ called ridge parameter to the diagonal elements of the matrix $X'X$, we obtain the following ridge estimator

$$\hat{\beta}_{RR} = (X'X + kI_p)^{-1}X'Y, \quad k > 0 \tag{1.3}$$

The MSEs of the OLS estimator and the ridge estimator $\hat{\beta}_{RR}$ are as follows respectively,

* Corresponding author.
E-mail addresses: yasar@konya.edu.tr, yasinasar@hotmail.com (Y. Asar).
Peer review under responsibility of University of Bahrain.

http://dx.doi.org/10.1016/j.jaubas.2014.08.005
1815-3852

$$\mathrm{MSE}(\hat{\beta}) = \sigma^2 \sum_{i=1}^{p} \frac{1}{\lambda_i} \tag{1.4}$$

$$\mathrm{MSE}\left(\hat{\beta}_{RR}\right) = Var\left(\hat{\beta}_{RR}\right) + \left[\mathrm{Bias}(\hat{\beta}_{RR})\right]^2$$
$$= \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \beta'(X'X + kI_p)^{-2}\beta \tag{1.5}$$

where λ_i 's are eigenvalues of the matrix $X'X$ and σ^2 is the error variance.

Hoerl and Kennard, 1970b showed the properties of this function in detail. They concluded that the total variance decreases and the squared bias increases as k increases. The variance function is monotonically decreasing and the squared bias function is monotonically increasing. Thus, there is the probability that some k exists such that the MSE for $\hat{\beta}_{RR}$ is less than MSE for the usual $\hat{\beta}$ (Hoerl and Kennard, 1970ab).

We know that k is estimated from the observed data. There are many papers proposing different ridge parameters in the literature. In recent papers, these parameters have been compared with the one proposed by Hoerl et al., 1975 and each other. After Hoerl and Kennard, 1970b, many researchers studied this area and proposed different estimates of the ridge parameter. Some of them are McDonald and Galarneau (1975), Lawless and Wang (1976), Saleh and Kibria (1993), Liu and Gao (2011), Kibria (2003), Khalaf and Shukur (2005), Alkhamisi et al. (2006), Adnan et al. (2006), Yan (2008), Yan and Zhao (2009), Muniz and Kibria (2009), Mansson et al. (2010), Al-Hassan (2010), Muniz et al. (2012), Asar et al. (2014) and Dorugade (2014).

The purpose of this article is to study much of the parameters in the literature and propose some new ones and also make a comparison between them by conducting a Monte Carlo experiment. The comparison criterion is based on the mean squared properties.

The article is organized as follows. In Section 2, we present the methodology of different estimators and give some new estimators. A Monte Carlo simulation has been provided in Section 3. Results of the simulation are discussed in Section 4. In Section 5, an application of the estimators is given. Finally, we give a summary and conclusion.

2. Model and estimators

Firstly we write the general model (1.1) in canonical form. Suppose that there exits an orthogonal matrix D we apply a transformation such that

$$D(X'X)D' = \Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_p) \tag{2.1}$$

where D is a $p \times p$ orthogonal matrix and $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p$. If we substitute $Z = XD$ and $\alpha = D'\beta$ in the model (1.1), then the model may be rewritten as

$$Y = Z\alpha + \varepsilon \tag{2.2}$$

where $Z'Z = \Lambda$.

Thus, the ridge estimator of α becomes $\hat{\alpha}_{RR} = (Z'Z + kI_p)^{-1}Z'Y$. It is stated in Hoerl and Kennard, 1970a that the value of k minimizing the $\mathrm{MSE}(\hat{\alpha}_{RR})$ is

$$k_i = \frac{\hat{\sigma}^2}{\alpha_i^2}. \tag{2.3}$$

As seen in the formula (2.3), k depends on the unknown parameters σ^2 and α. Hence we use the estimators $\hat{\sigma}^2$ and $\hat{\alpha}$ due to Hoerl and Kennard, 1970b and get

$$k_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}. \tag{2.4}$$

2.1. Proposed estimators

In this section, we review some of the ridge estimators suggested earlier and propose some new ones. The list of estimators with which we will compare ours is given below:

$$(1)\ k_1 = k_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2}\ \text{(Hoerl \& Kennard, 1970a)} \tag{2.5}$$

where $\hat{\alpha}_{\max}$ is the maximum element of $\hat{\alpha}$.

$$(2)\ k_2 = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p} \lambda_i \hat{\alpha}_i^2}, i = 1, 2, \ldots, p\ \text{(Lawless \& Wang, 1976)} \tag{2.6}$$

which is proposed from the Bayesian point of view.

$$(3)\ k_3 = median\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\right), i = 1, 2, \ldots, p\ \text{(Kibria, 2003)} \tag{2.7}$$

which is the median of $\hat{k}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$.

$$(4)\ k_4 = \frac{2\hat{\sigma}^2}{\lambda_{\max}(\prod \hat{\alpha}_i^2)^{1/p}}, i = 1, 2, \ldots, p\ \text{(Dorugade, 2014)} \tag{2.8}$$

which is the geometric mean of $\hat{k}_i = \frac{2\hat{\sigma}^2}{\lambda_{\max}\hat{\alpha}_i^2}$.

$$(5)\ k_5 = \frac{2p}{\lambda_{\max}} \frac{\hat{\sigma}^2}{\sum_{i=1}^{p} \hat{\alpha}_i^2}, i = 1, 2, \ldots, p\ \text{(Dorugade, 2014)} \tag{2.9}$$

which is the harmonic mean of $\hat{k}_i = \frac{2\hat{\sigma}^2}{\lambda_{\max}\hat{\alpha}_i^2}$.

A sufficient condition that $\mathrm{MSE}(\hat{\alpha}_{RR}) < \mathrm{MSE}(\hat{\alpha})$ is given by Hoerl and Kennard (1970a,b) such that $k < k_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2}$. A quick survey shows us that some of the existing ridge parameters are smaller than k_{HK}. However, if we try the estimators larger than k_{HK}, we observe that one can also have better estimators in sense of MSE.

In the figure given by Hoerl and Kennard (1970a,b), it is obvious that the first derivative of the function $\mathrm{MSE}(\hat{\alpha}_{RR})$ is negative when the value of k_{HK} is used as the biasing parameter. Therefore, any estimator satisfying $0 < k < k_{HK}$ gives us a negative derivative. However, if we examine the intersection point of the variance and the squared bias functions, we see that it is absolutely greater than k_{HK}. Thus, one can find estimators such that the first derivative of the $\mathrm{MSE}(\hat{\alpha}_{RR})$ function is positive and being greater than k_{HK}. There are greater estimators than k_{HK} in the literature, for example see Alkhamisi and Shukur, 2007 for the estimators k_{NAS} and k_{AS}.

It should also be pointed out that the optimal selection process of the parameter k in ridge regression cannot be truly provided from the theoretical point of view. Actually, this is an open problem to researchers. Thus we suggest some estimators which are modifications of $k_K = \frac{1}{p}\sum_{i=1}^{p} \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$ proposed in Kibria, 2003 $k_D = \frac{2p}{\lambda_{\max}} \frac{\hat{\sigma}^2}{\sum_{i=1}^{p} \hat{\alpha}_i^2}$ and $k_D = \frac{2p}{\lambda_{\max}} \frac{\hat{\sigma}^2}{\sum_{i=1}^{p} \hat{\alpha}_i^2}$ proposed in Dorugade, 2014. We apply some transformations and we fol-