Journal of the Association of Arab Universities for Basic and Applied Sciences (2014) **15**, 94–99

University of Bahrain

Journal of the Association of Arab Universities for Basic and Applied Sciences

www.elsevier.com/locate/jaaubas
www.sciencedirect.com

REVIEW ARTICLE

New ridge parameters for ridge regression

A.V. Dorugade *

Y C Mahavidyalaya Halkarni, Tal-Chandgad, Kolhapur 416552, Maharashtra, India

Received 8 October 2012; revised 9 February 2013; accepted 31 March 2013
Available online 10 May 2013

KEYWORDS

Ridge regression;
Ridge parameter;
Multicollinearity

Abstract Hoerl and Kennard (1970a) introduced the ridge regression estimator as an alternative to the ordinary least squares (OLS) estimator in the presence of multicollinearity. In ridge regression, ridge parameter plays an important role in parameter estimation. In this article, a new method for estimating ridge parameters in both situations of ordinary ridge regression (ORR) and generalized ridge regression (GRR) is proposed. The simulation study evaluates the performance of the proposed estimator based on the mean squared error (MSE) criterion and indicates that under certain conditions the proposed estimators perform well compared to OLS and other well-known estimators reviewed in this article.

1. Introduction

In the presence of multicollinearity OLS estimator yields regression coefficients whose absolute values are too large and whose signs may actually reverse with negligible changes in the data (Buonaccorsi, 1996). Whenever the multicollinearity presents in the data, the OLS estimator performs 'poorly'. The method of ridge regression, proposed by Hoerl and Kennard (1970a) is one of the most widely used tools to the problem of multicollinearity. In a ridge regression an additional parameter, the ridge parameter (k) plays an vital role to control the bias of the regression toward the mean of the response variable. In addition, they proposed the generalized ridge regression (GRR) procedure that allows separate ridge parameter for each regressor. Using GRR, it is easier to find optimal values of ridge parameter, i.e., values for which the MSE of the ridge estimator is minimum. In addition, if the optimal values for biasing constants differ significantly from each other then this estimator has the potential to save a greater amount of MSE than the OLS estimator (Stephen and Christopher, 2001). In both ORR and GRR as 'k' increases from zero and continues up to infinity, the regression estimates tend toward zero. Though these estimators result in biased, for certain value of k, they yield a minimum mean squared error (MMSE) compared to the OLS estimator (see Hoerl and Kennard, 1970a). Ridge parameter 'k' proposed by Hoerl et al. (1975) performs fairly well.

Much of the discussions on ridge regression concern the problem of finding good empirical value of k. Recently, many researchers have suggested various methods for choosing ridge parameter in ridge regression. These methods have been suggested by Hoerl and Kennard (1970a), Hoerl et al. (1975), McDonald and Galarneau (1975), Hocking et al. (1976), Lawless and Wang (1976), Gunst and Mason (1977), Lawless (1978), Nomura (1988), Heath and co-workers (1979), Nordberg (1982), Saleh and Kibria (1993), Haq and Kibria (1996), Kibria (2003), Pasha and Shah (2004), Khalaf and

* Tel./fax: +91 02320 239076.
E-mail address: adorugade@rediffmail.com.

Peer review under responsibility of University of Bahrain.

Production and hosting by Elsevier

1815-3852 © 2013 Production and hosting by Elsevier B.V. on behalf of University of Bahrain.
http://dx.doi.org/10.1016/j.jaubas.2013.03.005

Shukur (2005), Norliza et al. (2006), Alkhamisi and Shukur (2007), Mardikyan and Cetin (2008), Dorugade and Kashid (2010) and Al-Hassan (2010) to mention a few. The objective of the article is to investigate some of the existing popular techniques that are available in the literature and to make a comparison among them based on mean square properties. Moreover, we suggested some methods for estimating ridge parameters in ORR and GRR which produce ridge estimators that yield minimum MSE than other estimators. The organization of the article is as follows.

In this article, we introduce alternative ordinary and generalized ridge estimators and study their performance by means of simulation techniques. Comparisons are made with other ridge-type estimators evaluated elsewhere, and the estimators to be included in this study are described in Section 2. In Section 3, we propose some new methods for estimating the ridge parameter. In Section 4, we illustrate the simulation technique that we have adopted in the study and related results of the simulations appear in the tables and figures. In Section 5, we give a brief summary and conclusion.

2. Model and estimators

Consider, a widely used linear regression model

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

where Y is a $n \times 1$ vector of observations on a response variable. β is a $p \times 1$ vector of unknown regression coefficients, X is a matrix of order $(n \times p)$ of observations on 'p' predictor (or regressor) variables and ε is an $n \times 1$ vector of errors with $E(\varepsilon) = 0$ and $V(\varepsilon) = \sigma^2 I_n$. For the sake of convenience, we assume that the matrix X and 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 $\wedge$ and T be the matrices of eigen values and eigen vectors of $X'X$, respectively, satisfying $T'X'XT = \wedge =$ diagonal $(\lambda_1, \lambda_2, \ldots, \lambda_p)$, where λi being the ith eigen value 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 $Z'Z = \wedge$, and $\alpha = T'\beta$ (see Montogomery et al. (2006)) Then OLS estimator of α is given by

$$\hat{\alpha}_{OLS} = (Z'Z)^{-1}Z'Y = \wedge^{-1}Z'Y. \tag{3}$$

Therefore, OLS estimator of β is given by

$$\hat{\beta}_{OLS} = T\hat{\alpha}_{OLS}.$$

2.1. Generalized ridge estimator (GRR)

The GRR estimator of α is defined by

$$\hat{\alpha}_{GR} = (I - KA^{-1})\hat{\alpha}_{OLS}, \tag{4}$$

where $K = \text{diagonal}(k_1, k_2 \ldots k_p)$, $k_i \geqslant 0$, $i = 1,2,\ldots,p$ be the different ridge parameters for different regressors and $A = \wedge + K$.

Hence GRR estimator for β is $\hat{\beta}_{GR} = T\hat{\alpha}_{GR}$.
and mean square error of $\hat{\alpha}_{GR}$ is

$$\text{MSE}(\hat{\alpha}_{GR}) = \text{Variance}(\hat{\alpha}_{GR}) + [\text{Bias}(\hat{\alpha}_{GR})]^2$$
$$= \hat{\sigma}^2 \sum_{i=1}^{p} \lambda_i/(\lambda_i + k_i)^2 + \sum_{i=1}^{p} k_i^2 \hat{\alpha}_i^2/(\lambda_i + k_i)^2 \tag{5}$$

In case of GRR, various methods are available in the literature to determine the separate ridge parameter for each regressor. Among these, well known methods for determination of ridge parameter which are used in the further study are given below.

(1) Hoerl and Kennard (1970a) have proposed the following ridge parameter

$$k_i(\text{HK}) = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}, \quad i = 1, 2, \ldots, p \tag{6}$$

(2) Nomura (1988) proposed a ridge parameter and it is given by

$$k_i(\text{HMO}) = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\left\{1 + \left[1 + \lambda_i(\hat{\alpha}_i^2/\hat{\sigma}^2)^{1/2}\right]\right\}, \quad i = 1, 2, \ldots, p \tag{7}$$

(3) Troskie and Chalton (1996) proposed a ridge parameter and it is given by

$$k_i(\text{TC}) = \lambda_i\hat{\sigma}^2/\left(\lambda_i\hat{\alpha}_i^2 + \hat{\sigma}^2\right), \quad i = 1, 2, \ldots, p \tag{8}$$

(4) Firinguetti (1999) proposed a ridge parameter and it is given by

$$k_i(F) = \lambda_i\hat{\sigma}^2/\left[\lambda_i\hat{\alpha}_i^2 + (n-p)\hat{\sigma}^2\right], \quad i = 1, 2, \ldots, p \tag{9}$$

(5) Batah et al. (2008) proposed a ridge parameter and it is given by

$$k_i(\text{FG}) = \hat{\sigma}^2\left\{\left[\left(\hat{\alpha}_i^4\lambda_i^2/4\hat{\sigma}^2\right) + \left(6\hat{\alpha}_i^4\lambda_i/\hat{\sigma}^2\right)\right]^{1/2} - \left(\hat{\alpha}_i^2\lambda_i/2\hat{\sigma}^2\right)\right\}/\hat{\alpha}_i^2 \quad i = 1, 2, \ldots, p \tag{10}$$

where, $\hat{\alpha}_i$ is the ith element of $\hat{\alpha}_{OLS}$, $i = 1, 2, \ldots, p$ and $\hat{\sigma}^2$ is the OLS estimator of σ^2, i.e. $\hat{\sigma}^2 = \frac{Y'Y - \hat{\alpha}'Z'Y}{n-p-1}$.

2.2. Ordinary ridge estimator (ORR)

Setting $k_1 = k_2 = \ldots = k_p = k$ and $k \geqslant 0$, the Ordinary ridge regression (ORR) estimator of β is

$$\hat{\beta}_{RR} = T\hat{\alpha}_{RR} = T\left[I - kA_k^{-1}\right]\hat{\alpha}, \quad \text{where } A_k = (\wedge + kI_p) \tag{11}$$

and mean square error of $\hat{\alpha}_{RR}$ is

$$\text{MSE}(\hat{\alpha}_{RR}) = \text{Variance}(\hat{\alpha}_{RR}) + [\text{Bias}(\hat{\alpha}_{RR})]^2$$
$$= \hat{\sigma}^2\sum_{i=1}^{p}\lambda_i/(\lambda_i + k)^2 + k^2\sum_{i=1}^{p}\hat{\alpha}_i^2 \Big/ (\lambda_i + k)^2 \tag{12}$$

We observe that, when $k = 0$ in (12), MSE of OLS estimator of α is recovered. Hence

$$\text{MSE}(\hat{\alpha}_{OLS}) = \hat{\sigma}^2\sum_{i=1}^{p} 1 \Big/ \lambda_i$$

Hoerl et al. (1975) suggested that, the value of 'k' is chosen small enough, for which the mean squared error of ridge estimator, is less than the mean squared error of OLS estimator.

In case of ORR also, many researchers have suggested different ways of estimating the ridge parameter. Some of the well known methods for choosing the ridge parameter value are listed below.