A new robust ridge parameter estimator based on search method for linear regression model

Atila Göktaş , Özge Akkuş & Aykut Kuvat

JOURNAL OF APPLIED STATISTICS
https://doi.org/10.1080/02664763.2020.1803814

A new robust ridge parameter estimator based on search method for linear regression model

Atila Göktaş, Özge Akkuş and Aykut Kuvat

Department of Statistics, Muğla Sıtkı Koçman University, Muğla, Turkey

ABSTRACT

A large and wide variety of ridge parameter estimators proposed for linear regression models exist in the literature. Actually proposing new ridge parameter estimator lately proving its efficiency on few cases seems endless. However, so far there is no ridge parameter estimator that can serve best for any sample size or any degree of collinearity among regressors. In this study we propose a new robust ridge parameter estimator that serves best for any case assuring that is free of sample size, number of regressors and degree of collinearity. This is in fact realized by choosing three best from enormous number of ridge parameter estimators performing well in different cases in developing the new ridge parameter estimator in a way of search method providing the smallest mean square error values of regression parameters. After that a simulation study is conducted to show that the proposed parameter is robust. In conclusion, it is found that this ridge parameter estimator is promising in any case. Moreover, a recent data set is used as an example for illustration to show that the proposed ridge parameter estimator is performing better.

ARTICLE HISTORY

Received 23 March 2020
Accepted 26 July 2020

KEYWORDS

Ridge regression; multicollinearity; ridge parameters; robust ridge parameter

1. Introduction

When multicollinearity exists in a linear regression model, using t test statistics for testing the coefficients of the independent variables becomes a serious problem [5]. One of the methods that can be used to troubleshoot multicollinearity is the Ridge Regression (RR) method developed by Hoerl and Kennard in the 1970s. It is mainly used to reduce the degree of collinearity by adding a positive and fairly small value (k) to the diagonal elements of the covariance matrix. The essential problem we face is how small or what size value is optimum to overcome multicollinearity. The objective in RR is to select the ridge parameter (k), which makes the Mean Square Error (MSE) minimum. When an optimal k is selected in RR, MSE will be minimum together with the variance. From past to present, there have been many studies conducted to suggest methods on the selection of k. These suggestions most of which are listed below are not fairly effective methods to determine the best k in all cases.

CONTACT Özge Akkuş ozge.akkus@mu.edu.tr; ozgeakku@gmail.com Department of Statistics, Muğla Sıtkı Koçman University, Muğla 48000, Turkey

Supplemental data for this article can be accessed here. https://doi.org/10.1080/02664763.2020.1803814

Hoerl, and Kennard [8] suggested in their extended study that a separate k value could be selected for each regression. However, they also stated that there is no guarantee that this will give better results than the k trace in any case. Hoerl, and Kennard [9] stated that there is no single value of k that is the ridge parameter estimator and that the results would be better than OLS if the optimal k could be determined. They suggested the ridge trace for the selection of k. Marquardt and Snee [18] stated that when the independent variables are highly correlated, RR produces coefficients better than OLS. Hoerl et al. [10] suggested an algorithm for selecting the parameter k with superior features than OLS. McDonald, and Galarneau [20] proposed two analytic methods of determining k parameter and evaluated their performances in terms of MSE values by Monte Carlo simulations. Lawless, and Wang [14] made a simulation study of ridge and other regression estimators. Golub et al. [7] proposed to select the k that minimizes the cross-validation statistics. Andersson [3] stated that the chosen values of k will be one where the mean square error is less than for the OLS. Kibria [13] proposed a few new ridge parameter estimators based on a generalized ridge regression approach. Alkhamisi et al. [2] and Khalaf, and Shukur [12] proposed two new ridge parameter estimators based on the median and largest eigenvalue in the linear regression (Equations (5) and (6)). Sakallıoğlu and Kaçıranlar [22] presented a new approach in determining the k parameter by augmenting a new equation to the classical linear regression model. Muniz and Kibria [21] proposed a new ridge parameter estimator based on the orthogonal eigenvectors matrix in linear regression (Equation (4)). Mansson et al. [16] conducted a simulation study to compare the performance of some ridge estimators based on both MSE values and Prediction Sum of Square (PRESS) values. A new method for estimating ridge parameter is proposed by Al-Hassan [1] and a simulation study has been made to evaluate the performance of the proposed estimator based on MSE values. Dorugade [4] proposed a new ridge parameter estimator in Ordinary Ridge Regression and also in Generalized Ridge Regression. Khalaf and Iguernane [11] proposed a new estimator of ridge parameter and evaluated by simulation techniques in term of MSE value. Göktaş and Sevinç [7] proposed two new k parameters and conducted multiple simulation studies. Lukman and Olatunji [15] proposed a new ridge parameter estimator, a function of the standard error of regression resulting with an independent estimation of the regression coefficients in Ordinary Ridge Regression. Göktaş and Sevinç [6] compared the effectiveness of 37 different k ridge parameter estimators presented mostly in the above studies in addition to the estimators they proposed through a simulation study designed according to different sample sizes, different correlation coefficients and different numbers of variables.

A regression model with the smallest MSE value in all circumstances cannot be predicted using the ridge parameter estimators proposed in the literature. In our study, a new ridge parameter estimator is developed which gives the smallest MSE in each case regardless of the sample size, the number of variables or the correlation coefficient. Through a simulation study designed under different scenarios, the effectiveness of the proposed ridge parameter estimator was investigated. Unlike other studies, while proposing a new robust ridge parameter estimator, in the first stage, search method using the priori information obtained as a result of search to reach a new robust estimator was used. Three best k parameters determined in the simulation study by Golub et al. [6] were mainly used. These parameters, expressed as k_1, k_2, k_3, are given in Equations (4)–(6) respectively.

Let the linear regression model used in the calculation of these parameters be defined as follows;

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

where Y in the Equation (1) represents the $(n \times 1)$ dimensional dependent variable vector. X represents the $(n \times p)$ dimensional explanatory variables vector, β represents the $(p \times 1)$ dimensional unknown regression coefficients, ε denotes $(n \times 1)$ dimensional zero mean and constant σ^2 variance error term. The OLS estimator of β is given as follows:

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{2}$$

The OLS estimation of α parameters vector that are used in the calculation of ridge parameter estimators presented from Equations (4) to (7) can be written in the following form,

$$\hat{\alpha} = T'\hat{\beta} \tag{3}$$

where the matrix T is the orthogonal eigenvectors obtained from the covariance matrix of $X'X$.

$$k_1 = \left[\prod_{i=1}^{p}\left(\frac{\hat{\alpha}_{i,ols}}{\hat{\sigma}_{ols}}\right)\right]^{\frac{1}{p}}; \quad i = 1, 2, \ldots, p \tag{4}$$

k_1 parameter in Equation (4) was proposed by Muniz and Kibria [21], where p represents the number of explanatory variables, $\hat{\alpha}_i$ represents the ith function of OLS parameter estimator of the corresponding variable and $\hat{\sigma}_{ols}$ represents the square root of the MSE of the model. The next ridge parameter estimator given in the following form in Equation (5) was proposed by Alkhamisi et al. [2] and Khalaf and Shukur [12],

$$k_2 = \left[\frac{\lambda_{max}\hat{\sigma}^2_{ols}}{(n-p)\hat{\sigma}^2_{ols} + \lambda_{max}\hat{\alpha}^2_i}\right]_{max}; \quad i = 1, 2, \ldots, p \tag{5}$$

where λ_{max} represents the largest eigenvalue of $X'X$ matrix, n represents the sample size and $\hat{\alpha}_i$ represents the ith value of the vector $\hat{\alpha}$. The next ridge parameter estimator presented in Equation (6) was also proposed by Alkhamisi et al. [2] and Khalaf and Shukur [12] considering the median for p different number of $\hat{\alpha}_i$ parameters;

$$k_3 = \left[\frac{\lambda_i\hat{\sigma}^2_{ols}}{(n-p)\hat{\sigma}^2_{ols} + \lambda_i\hat{\alpha}^2_i}\right]_{Median}; \quad i = 1, 2, \ldots, p \tag{6}$$

where λ_i is the ith eigenvalue of the $X'X$ matrix.

Göktaş and Sevinç [7] determined that the above-given three ridge parameter estimators out of numerous number of ridge parameter estimators produced the smallest MSE value in different cases by separately evaluating 30, 50, 80, 100, 250, 500 sample sizes; 0.3, 0.5, 0.9 correlation coefficients; and 3 and 5 number of variables. For instance, with $n = 500$, $\rho = 0.9$ and $p = 3$ constraints, the k_1 parameter; with $n = 250$, $\rho = 0.5$ and $p = 7$ constraints, k_2 parameter; and with $n = 100$, $\rho = 0.3$ and $p = 7$ constraints, the k_3 parameter yield the minimum MSE and so on.

2. Methodology

In the case of multicollinearity in a linear regression, the standard errors of the regression coefficients of the significant explanatory variables increase and the t-test of the coefficients turns out to be insignificant. As a result, regression coefficients may partially emerge as different from what is expected while working with data having multicollinearity. Meanwhile the standardized regression coefficients calculated lose their stability.

One of the methods used to eliminate multicollinearity is the RR method. This method was first proposed in 1970 by Hoerl and Kennard in one of their first studies presenting a detailed discussion of the unbiased estimation problem in a multiple regression complying with the full-rank general hypothesis model.

The studies Hoerl and Kennard [8] and [9] suggested to use ridge trace graph to show the inconsistency in the estimated coefficients in the case of multicollinearity to the highest degree, to obtain coefficients having smaller variance compared to OLS estimates.

Since RR is a biased method eliminating the multicollinearity, a small k value called ridge parameter estimator is added to the diagonal elements of the $X'X$ matrix resulting with parameter estimation of the regression model is as follows:

$$\hat{\beta}_{RR} = (X'X + kI)^{-1}X'Y \tag{7}$$

The purpose of adding the k value in Equation (7) is to significantly reduce the inflated variances of the estimators due to the multicollinearity. If $k = 0$, then the results coincide with OLS estimations. In this regard, the ridge estimation can be called a linear transformation of OLS [19,22].

2.1. The relationship of ridge estimation with OLS

In OLS, $\hat{\beta}$ estimation is defined as follows.

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{8}$$

Here; it is possible to rewrite the Equation (8) as follows.

$$X'X\hat{\beta} = X'Y \tag{9}$$

The ridge estimation can be expressed as,

$$\hat{\beta}^* = (X'X + kI)^{-1}X'Y \tag{10}$$

From Equation (9) when the $X'Y$ term is substituted in Equation (10), the following result is obtained.

$$\hat{\beta}^* = (X'X + kI)^{-1}X'X\hat{\beta} \tag{11}$$

As the inverse of $(X'X)^{-1}$ matrix is equal to $X'X$ itself, Equation (11) can be rewritten as follows;

$$\hat{\beta}^* = (X'X + kI)^{-1}[(X'X)^{-1}]^{-1}\hat{\beta} \tag{12}$$

As neither matrix is singular, Equation (12) can be written as;

$$\hat{\beta}^* = [(X'X)^{-1}(X'X + kI)]^{-1}\hat{\beta} \tag{13}$$

From here,

$$\hat{\beta}^* = [(X'X)^{-1}X'X + k(X'X)^{-1}]^{-1}\hat{\beta} \tag{14}$$

is obtained. After the operations are conducted,

$$\hat{\beta}^* = [I + k(X'X)^{-1}]^{-1}\hat{\beta} \tag{15}$$

is obtained. If a new expression Z is defined as

$$Z = [I + k(X'X)^{-1}]^{-1} \tag{16}$$

then, Equation (15) can be rewritten in the following format depending on Z.

$$\hat{\beta}^* = Z\hat{\beta} \tag{17}$$

The relationship in Equation (17) demonstrates that ridge estimation is a transformation of OLS estimation having weight Z [17].

2.2. Mean square error of the parameters in ridge regression

Since MSE for ridge estimation is a measure of the quadratic distance ($L_\imath^2(k)$) between $\hat{\beta}^*$ and β as follows,

$$L_\imath^2(k) = (\hat{\beta}^* - \beta)'(\hat{\beta}^* - \beta) \tag{18}$$

the expected value of $L_\imath^2(k)$ can be used as the MSE approximation as follows.

$$E[L_\imath^2(k)] = \sigma^2 \sum_{j=1}^{P} \frac{\lambda_j}{(\lambda_j + k)^2} + k^2\beta'[tr(X'X + kI)]^{-2}\beta \tag{19}$$

When $k = 0$, ridge estimation becomes identical to OLS and the expected MSE in Equation (19) reduces to the following.

$$E[L_\imath^2(0)] = \sigma^2 \sum_{j=1}^{p} \lambda_j^{-1} \tag{20}$$

On the basis of these solutions, the following result is obtained.

$$E[L_\imath^2(k)] < E[L_\imath^2(0)] \tag{21}$$

Hoerl and Kennard [8] stated that it is always possible to find a value of k where the above Equation (21) holds. Therefore, in the case of multicollinearity, ridge estimation always yields a smaller MSE value than the OLS estimation. Another form of that value can be estimated where the second term of Equation (19) is simplified as follows [13].

$$MSE(\hat{\beta}) = \hat{\sigma}^2 \sum_{j=1}^{P} \frac{\lambda_j}{(\lambda_j + \hat{k})^2} + \hat{k}^2 \sum_{j=1}^{P} \frac{\alpha_i^2}{(\lambda_j + \hat{k})^2} \tag{22}$$

3. The proposed method

In this study, a new robust ridge parameter estimator that provides the smallest MSE for regression parameters in any condition has been proposed. To achieve this, the study by Göktaş and Sevinç [6], which investigated 37 different ridge parameter estimators proposed in the literature, is examined. Three different ridge parameter estimators that were determined to be best in different cases and were given with Equations (4)–(6) are taken into consideration. The values obtained from these three different parameters were determined to be between 0 and 6. Therefore, the new ridge parameter estimator, developed to be robust, was allowed to be between 0 and 10. In fact, there is a single ridge parameter estimator that gives the smallest MSE value for the parameters. This ridge parameter estimator can be obtained by searching all values from 0 with 0.001 increments until it reaches 10. The MSE value for the regression parameters is calculated at each searching value. The k value that gives the smallest MSE value as a result of the search is considered the best (See Table 1). The multicollinearity data used in the study were obtained by simulation. For the ridge parameter estimator, which gives the smallest MSE value, the ridge parameter estimator search method was performed with the data produced 15 times from each combination with 12 different sample sizes, 9 different collinearity levels, and 3 different explanatory variables. For a total of 4860 different data sets, ridge parameter estimator search method was performed. The best ridge parameters obtained as a result of the search were determined according to the MSE values in Table 1. The three different ridge parameter estimator values, which are considered the best in the study of Göktaş and Sevinç [6] were calculated for the same data sets. Some of the results are given in Table 2.

In the second phase of our study, taking the robust ridge parameter estimator values as the dependent variable, we investigated whether there is a linear relationship among k_1, k_2, k_3 parameter estimators, sample size and degree of collinearity. During the data

Table 1. A part of MSE results obtained from different k parameters in the first stage of search.

MSE obtained from k_{search}	MSE obtained from k_1	MSE obtained from k_2	MSE obtained from k_3	n	ρ	p	MSE_{Min}
1.58885	1.78617	1.62856	1.70099	20	0.1	7	**1.58885**
0.83587	0.93335	0.85314	0.91272	30	0.5	5	**0.83587**
0.83173	0.87046	0.83777	0.85531	20	0.1	5	**0.83173**
0.78336	0.94679	0.84725	0.95447	30	0.8	5	**0.78336**
0.77551	0.82583	0.79413	0.82472	30	0.6	5	**0.77551**
0.77003	0.86659	0.77479	0.83512	30	0.7	5	**0.77003**
0.74640	0.78267	0.84752	0.78594	20	0.3	5	**0.74640**
0.72584	0.77332	0.89200	0.78537	30	0.2	5	**0.72584**
0.63921	0.66205	0.66385	0.65543	20	0.2	7	**0.63921**
0.63164	0.75114	0.66670	0.76513	30	0.8	5	**0.63164**
0.62888	0.81631	0.72672	0.80739	50	0.8	5	**0.62888**
0.60528	0.66526	0.64829	0.66550	20	0.3	5	**0.60528**
0.60526	1.52120	0.77422	1.38637	50	0.9	5	**0.60526**
0.60474	0.80413	0.98415	0.68937	20	0.2	5	**0.60474**
0.60136	0.60164	0.64462	0.60338	20	0.6	5	**0.60136**
0.59712	0.67272	0.78915	0.67583	20	0.3	5	**0.59712**
0.59648	0.77079	0.59781	0.77192	20	0.8	5	**0.59648**
0.59412	0.80572	0.97339	0.74613	20	0.2	5	**0.59412**
0.59167	0.61972	0.67816	0.60847	20	0.5	5	**0.59167**
0.58908	0.63477	0.59030	0.58925	20	0.1	7	**0.58908**
0.57841	0.88521	0.93643	0.78455	20	0.2	7	**0.57841**
0.55804	0.62042	0.59728	0.60372	30	0.5	5	**0.55804**

Table 2. Part of simulation results used to obtain the k value.

k_{search}	k_1	k_2	k_3	n	ρ	p
3.185	0.57811	5.12608	1.04781	20	0.1	7
4.373	0.63783	4.46262	0.49421	20	0.7	5
1.223	1.64366	2.87248	1.07251	20	0.1	5
3.679	0.99958	2.71986	1.52776	20	0.5	5
5.718	0.87451	3.21015	1.24678	30	0.5	5
3.186	0.73315	2.13891	1.21666	20	0.1	5
1.271	1.27004	3.00570	1.34433	20	0.3	5
0.530	0.98321	2.51109	0.79434	20	0.5	5
0.827	1.28944	2.03219	1.11020	20	0.3	7
3.969	0.91686	3.16477	1.36405	30	0.7	5
3.249	0.79279	2.07189	1.24338	20	0.7	5
1.918	0.76926	3.05607	1.02267	20	0.6	5
0.863	1.09421	1.86871	1.52918	20	0.2	5
0.962	1.67479	1.70558	1.55231	20	0.2	5
4.283	0.97878	2.09682	0.87024	30	0.8	5
6.290	0.76191	6.03092	1.01802	20	0.6	5
0.001	1.90329	1.97443	1.27600	20	0.3	5
2.589	1.12567	3.89738	1.29138	20	0.4	5
0.266	1.02545	1.76363	0.83427	20	0.5	5
1.333	2.37265	1.48746	1.27654	20	0.1	7
4.355	1.28039	1.09299	1.72801	20	0.2	7
4.731	0.85571	1.53241	1.32141	30	0.5	5

reproduction process of the exploratory variables with multicollinearity, Equation (23) was utilized,

$$X_{ij} = Z_{ij}\sqrt{(1-\rho^2)} + \rho Z_{ip}; \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, p \tag{23}$$

where, ρ represents the degree of correlation; Z_{ij} represents the i. observation value of j. variable derived from the standard normal distribution; and Z_{ip} represents the i. observation value i. of p. variable derived from the standard normal distribution. Error term being a variable derived from the standard normal distribution, the dependent variable is derived as follows;

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i \tag{24}$$

In the data derivation process, the following combinations were taken into consideration for the design of the simulation, where p represents variable numbers, ρ represents the correlation coefficient and n represents the sample size.

$$p = 3, 5, 7$$
$$\rho = 0.1,\ 0.2,\ 0.3,\ 0.4\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9$$
$$n = 20,\ 30,\ 50,\ 60,\ 80,\ 100,\ 120,\ 150,\ 180,\ 200,\ 250,\ 500$$

In our study, in the process of producing the dependent variable, for simplicity all of the regression parameters were taken as ‘1’. This is because changing the fixed coefficients will have no considerable effect on the order of MSE values from different ridge parameters. When the regression model for the derived data is estimated with any ridge parameter, the MSE value related to the parameter is calculated as follows, where r represents the number of repeated simulations.

$$\mathrm{MSE}_{\hat{\beta}_{ridge}} = \sum_{i=1}^{r} \sum_{j=0}^{p} (1 - \hat{\beta}_{i,j_{ridge}})^2 / [r(p+1)] \tag{25}$$

Three different k parameters for each derived data set and a part of the MSE results for the k parameter obtained as a result of the search are given in Table 1. Since the table was originally composed of 4860 lines, only part of it was shared.

As can be seen from Table 1, in the first stage, a robust parameter discovery study was performed until k_{search} gave the smallest MSE value and a new k value was obtained for each repetition or data reproduction type. Part of the simulation results used to obtain the k value is given in Table 2.

Based on the assumption that the k_{search} value acquired as a result of search method can be a linear function of k_1, k_2, k_3, n, ρ and p variables, by considering k_{search} to be the dependent variable and k_1, k_2, k_3, n, ρ and p to be the explanatory variables, the linear regression model was established. In the resulting regression model, a new model was obtained by subtracting the p explanatory variable from the linear regression model since it was not found to be significant. On the basis of the estimated model, it is thought that $\hat{k}_{search} = k_{robust}$. k_{robust} ridge parameter estimator value is given in Equation (26).

$$k_{robust} = 0.6149k_1 - 0.1589k_2 + 0.093k_3 + 0.00203n + 1.013\rho + 0.7484p \tag{26}$$

After the k_{robust} ridge parameter estimator is suggested, significance of the model and variables needs to be tested. The hypothesis established to test the significance of the multiple linear regression model and the results of ANOVA are given in the table below.

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$$

H_1: At least one of the explanatory variables makes a significant contribution to the model.

As can be understood from the ANOVA table presented in Table 3, as $p_{value} = 0.000 < 0.05$ for the regression model established, H_0 hypothesis is rejected and it can be said that the estimated model is significant. When p_{values} for individual variables or individual coefficients are examined, it is seen that as each probability value is lower than 0.05, they seem to be significant. As both the estimated model and the estimated coefficients of the explanatory variables are statistically significant, it can be argued that k_1, k_2, k_3, n, ρ and p have significant effects on the proposed k_{robust} parameter. In order to evaluate the performance of the proposed k_{robust} ridge parameter estimator, the design of the simulation is reestablished with $p = 3, 5, 7$ variable numbers, $\rho = 0.1, \ldots, 0.9$ coefficients and $n = 20, 30, 50, 60, 80, 100, 120, 150, 180, 200, 250, 500$ sample sizes. In the current study, 10,000 is determined as the number of repetitions. By taking the mean of the 10,000 MSE values calculated for each case, mean MSE value was obtained for each ridge parameter estimator. The purpose of the second stage of the study is to test whether the proposed k_{robust} parameter estimator can yield smaller MSE value compared to k_1, k_2, k_3 parameters. In the second stage simulation study, the MSE results obtained for the ridge parameter estimators are given in Tables 4–17.

When the results obtained from the second stage of the study were examined, it was determined that the k_{robust} parameter estimator that we proposed gives a smaller MSE

Table 3. ANOVA results.

Source	Coef.	df	Adj SS	Adj MS	F-value	p-value
Regression		6	162578	27096.4	2163.42	0.000
k_1	0.614900	1	1517.6	1517.6	121.17	0.000
k_2	−0.158900	1	68.82	68.82	5.505	0.019
k_3	0.093000	1	50.26	50.26	4.021	0.045
n	0.002026	1	287.2	287.2	22.93	0.000
ρ	1.013000	1	246.2	246.2	19.66	0.000
p	0.748400	1	4496.3	4496.3	358.99	0.000
Error		4854	60675	12.5		
Total		4860	229919.38			

Table 4. MSE results obtained from $p = 3$, $\rho = 0.1$ and $\rho = 0.2$, different sample sizes and different ridge parameter estimators.

	$p = 3$, $\rho = 0.1$					$p = 3$, $\rho = 0.2$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.2247	0.2339	0.2222	0.2150	0.2150	0.2216	0.2279	0.2162	0.2081	0.2081
$n = 30$	0.1443	0.1487	0.1430	0.1289	0.1289	0.1419	0.1455	0.1398	0.1262	0.1262
$n = 50$	0.0821	0.0837	0.0818	0.0727	0.0727	0.0815	0.0829	0.0807	0.0723	0.0723
$n = 60$	0.0674	0.0685	0.0672	0.0599	0.0599	0.0677	0.0687	0.0671	0.0606	0.0606
$n = 80$	0.0504	0.0510	0.0503	0.0457	0.0457	0.0502	0.0508	0.0500	0.0459	0.0459
$n = 100$	0.0397	0.0401	0.0396	0.0364	0.0364	0.0400	0.0404	0.0398	0.0370	0.0370
$n = 120$	0.0328	0.0331	0.0328	0.0304	0.0304	0.0327	0.0329	0.0326	0.0305	0.0305
$n = 150$	0.0263	0.0265	0.0263	0.0247	0.0247	0.0261	0.0262	0.0260	0.0246	0.0246
$n = 180$	0.0218	0.0219	0.0218	0.0207	0.0207	0.0213	0.0215	0.0213	0.0203	0.0203
$n = 200$	0.0196	0.0197	0.0196	0.0187	0.0187	0.0197	0.0198	0.0197	0.0188	0.0188
$n = 250$	0.0154	0.0154	0.0153	0.0148	0.0148	0.0155	0.0156	0.0155	0.0149	0.0149
$n = 500$	0.0078	0.0078	0.0078	0.0077	0.0077	0.0078	0.0079	0.0078	0.0077	0.0077

Table 5. MSE results obtained from $p = 3$, $\rho = 0.3$ and $\rho = 0.4$, different sample sizes and different ridge parameter estimators.

	$p = 3$, $\rho = 0.3$					$p = 3$, $\rho = 0.4$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.2590	0.2297	0.2186	0.2100	0.2100	0.2313	0.2307	0.2219	0.2098	0.2098
$n = 30$	0.1449	0.1478	0.1419	0.1276	0.1276	0.1525	0.1544	0.1485	0.1328	0.1328
$n = 50$	0.0822	0.0835	0.0812	0.0730	0.0730	0.0871	0.0882	0.0857	0.0767	0.0767
$n = 60$	0.0696	0.0705	0.0688	0.0623	0.0623	0.0722	0.0731	0.0713	0.0645	0.0645
$n = 80$	0.0515	0.0520	0.0510	0.0468	0.0468	0.0530	0.0535	0.0525	0.0483	0.0483
$n = 100$	0.1513	0.1408	0.1440	0.0982	0.0982	0.0429	0.0432	0.0426	0.0396	0.0396
$n = 120$	0.0342	0.0344	0.0340	0.0319	0.0319	0.0356	0.0358	0.0354	0.0333	0.0333
$n = 150$	0.0266	0.0266	0.0265	0.0251	0.0251	0.0285	0.0286	0.0283	0.0269	0.0269
$n = 180$	0.0223	0.0224	0.0223	0.0213	0.0213	0.0235	0.0236	0.0234	0.0224	0.0224
$n = 200$	0.0197	0.0198	0.0197	0.0188	0.0188	0.0213	0.0214	0.0213	0.0204	0.0204
$n = 250$	0.0161	0.0162	0.0161	0.0155	0.0155	0.0168	0.0169	0.0168	0.0162	0.0162
$n = 500$	0.0080	0.0080	0.0080	0.0079	0.0079	0.0083	0.0084	0.0083	0.0082	0.0082

value than the other parameter estimators in almost every case except for 16 cases out of 324 different cases. Thus, it can be argued that k_{robust} parameter estimator is successful in approximately 95% of 324 different cases. When the cases in which k_{robust} does not yield the smallest MSE are examined, it is understood that there are very small differences between the MSE values given by the parameters (especially from $p = 5$ and $\rho = 0.5$ onwards at

Table 6. MSE results obtained from $p = 3$, $\rho = 0.5$ and $\rho = 0.6$, different sample sizes and different ridge parameter estimators.

	$p = 3$, $\rho = 0.5$					$p = 3$, $\rho = 0.6$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.2517	0.2429	0.2386	0.2253	0.2253	0.2760	0.2561	0.2590	0.2392	0.2392
$n = 30$	0.1625	0.1623	0.1573	0.1392	0.1392	0.1815	0.1772	0.1746	0.1519	0.1519
$n = 50$	0.0940	0.0949	0.0923	0.0819	0.0819	0.1063	0.1064	0.1038	0.0905	0.0905
$n = 60$	0.0791	0.0798	0.0779	0.0700	0.0700	0.0893	0.0896	0.0875	0.0770	0.0770
$n = 80$	0.0584	0.0589	0.0577	0.0528	0.0528	0.0658	0.0662	0.0649	0.0588	0.0588
$n = 100$	0.0468	0.0472	0.0464	0.0431	0.0431	0.0528	0.0530	0.0521	0.0478	0.0478
$n = 120$	0.0385	0.0388	0.0383	0.0359	0.0359	0.0436	0.0438	0.0432	0.0402	0.0402
$n = 150$	0.0309	0.0311	0.0307	0.0291	0.0291	0.0351	0.0352	0.0348	0.0328	0.0328
$n = 180$	0.0256	0.0258	0.0255	0.0244	0.0244	0.0290	0.0291	0.0288	0.0274	0.0274
$n = 200$	0.0231	0.0232	0.0230	0.0221	0.0221	0.0264	0.0265	0.0263	0.0251	0.0251
$n = 250$	0.0185	0.0185	0.0184	0.0178	0.0178	0.0210	0.0210	0.0209	0.0201	0.0201
$n = 500$	0.0091	0.0091	0.0091	0.0089	0.0089	0.0106	0.0106	0.0105	0.0103	0.0103

Table 7. MSE results obtained from $p = 3$, $\rho = 0.7$ and $\rho = 0.8$, different sample sizes and different ridge parameter estimators.

	$p = 3$, $\rho = 0.7$					$p = 3$, $\rho = 0.8$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.3190	0.2723	0.2913	0.2584	0.2584	0.4010	0.2925	0.3478	0.2779	0.2779
$n = 30$	0.2115	0.1974	0.2000	0.1658	0.1658	0.2765	0.2305	0.2523	0.1890	0.1890
$n = 50$	0.1279	0.1258	0.1239	0.1040	0.1040	0.1694	0.1588	0.1612	0.1233	0.1233
$n = 60$	0.1062	0.1054	0.1035	0.0883	0.0883	0.1425	0.1367	0.1369	0.1072	0.1072
$n = 80$	0.0805	0.0805	0.0791	0.0694	0.0694	0.1085	0.1063	0.1053	0.0858	0.0858
$n = 100$	0.0634	0.0636	0.0625	0.0562	0.0562	0.0871	0.0862	0.0852	0.0720	0.0720
$n = 120$	0.0531	0.0533	0.0524	0.0476	0.0476	0.0729	0.0725	0.0715	0.0614	0.0614
$n = 150$	0.0424	0.0425	0.0420	0.0388	0.0388	0.0585	0.0584	0.0576	0.0510	0.0510
$n = 180$	0.0359	0.0361	0.0357	0.0334	0.0334	0.0484	0.0484	0.0478	0.0429	0.0429
$n = 200$	0.0321	0.0322	0.0318	0.0299	0.0299	0.0432	0.0433	0.0427	0.0388	0.0388
$n = 250$	0.0254	0.0254	0.0252	0.0240	0.0240	0.0354	0.0354	0.0350	0.0322	0.0322
$n = 500$	0.0128	0.0128	0.0128	0.0125	0.0125	0.0177	0.0177	0.0176	0.0168	0.0168

Table 8. MSE results obtained from$p = 3$, $\rho = 0.9$ and $p = 5$, $\rho = 0.1$, different sample sizes and different ridge parameter estimators.

	$p = 3$, $\rho = 0.9$					$p = 5$, $\rho = 0.1$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.5500	0.3083	0.4329	**0.2860**	**0.2860**	0.3873	0.3988	0.3815	**0.3728**	**0.3728**
$n = 30$	0.4244	0.2582	0.3489	**0.1986**	**0.1986**	0.2359	0.2406	0.2336	**0.2156**	**0.2156**
$n = 50$	0.2795	0.2147	0.2494	**0.1448**	**0.1448**	0.1317	0.1336	0.1310	**0.1201**	**0.1201**
$n = 60$	0.2387	0.1954	0.2178	**0.1303**	**0.1303**	0.1062	0.1075	0.1057	**0.0977**	**0.0977**
$n = 80$	0.1894	0.1669	0.1766	**0.1109**	**0.1109**	0.0787	0.0794	0.0786	**0.0735**	**0.0735**
$n = 100$	0.1527	0.1411	0.1449	**0.0976**	**0.0976**	0.0613	0.0618	0.0612	**0.0577**	**0.0577**
$n = 120$	0.1295	0.1228	0.1241	**0.0880**	**0.0880**	0.0512	0.0514	0.0511	**0.0486**	**0.0486**
$n = 150$	0.1052	0.1019	0.1018	**0.0754**	**0.0754**	0.0407	0.0409	0.0406	**0.0390**	**0.0390**
$n = 180$	0.0879	0.0860	0.0855	**0.0657**	**0.0657**	0.0337	0.0339	0.0337	**0.0325**	**0.0325**
$n = 200$	0.0797	0.0784	0.0777	**0.0612**	**0.0612**	0.0304	0.0305	0.0304	**0.0295**	**0.0295**
$n = 250$	0.0643	0.0637	0.0630	**0.0512**	**0.0512**	0.0243	0.0244	0.0243	**0.0237**	**0.0237**
$n = 500$	0.0336	0.0336	0.0333	**0.0294**	**0.0294**	0.0120	0.0120	0.0120	**0.0120**	**0.0120**

Table 9. MSE results obtained from $p = 5$, $\rho = 0.2$ and $\rho = 0.3$, different sample sizes and different ridge parameter estimators.

	$p = 5$, $\rho = 0.2$					$p = 5$, $\rho = 0.3$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.3903	0.3932	0.3815	0.3723	0.3723	0.3961	0.3888	0.3836	0.3742	0.3742
$n = 30$	0.2393	0.2422	0.2356	0.2172	0.2172	0.2452	0.2461	0.2402	0.2212	0.2212
$n = 50$	0.1304	0.1319	0.1292	0.1196	0.1196	0.1385	0.1396	0.1369	0.1270	0.1270
$n = 60$	0.1075	0.1086	0.1066	0.0992	0.0992	0.1121	0.1130	0.1110	0.1036	0.1036
$n = 80$	0.0795	0.0801	0.0791	0.0746	0.0746	0.0835	0.0840	0.0829	0.0785	0.0785
$n = 100$	0.0639	0.0643	0.0636	0.0605	0.0605	0.0655	0.0659	0.0651	0.0622	0.0622
$n = 120$	0.0522	0.0525	0.0520	0.0498	0.0498	0.0542	0.0544	0.0539	0.0518	0.0518
$n = 150$	0.0414	0.0416	0.0413	0.0398	0.0398	0.0434	0.0435	0.0432	0.0418	0.0418
$n = 180$	0.0342	0.0343	0.0341	0.0330	0.0330	0.0357	0.0358	0.0356	0.0346	0.0346
$n = 200$	0.0309	0.0311	0.0309	0.0299	0.0299	0.0324	0.0325	0.0323	0.0314	0.0314
$n = 250$	0.0247	0.0247	0.0246	0.0240	0.0240	0.0254	0.0254	0.0253	0.0247	0.0247
$n = 500$	0.0120	0.0120	0.0120	0.0118	0.0118	0.0125	0.0125	0.0125	0.0124	0.0124

Table 10. MSE results obtained from $p = 5$, $\rho = 0.4$ and $\rho = 0.5$, different sample sizes and different ridge parameter estimators.

	$p = 5$, $\rho = 0.4$					$p = 5$, $\rho = 0.5$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.4193	0.3971	0.4041	0.3917	0.3917	0.4505	0.4076	0.4324	0.4098	0.4076
$n = 30$	0.2620	0.2585	0.2557	0.2335	0.2335	0.2875	0.2766	0.2799	0.2534	0.2534
$n = 50$	0.1476	0.1480	0.1456	0.1343	0.1343	0.1645	0.1636	0.1619	0.1478	0.1478
$n = 60$	0.1212	0.1218	0.1199	0.1116	0.1116	0.1365	0.1363	0.1346	0.1238	0.1238
$n = 80$	0.0904	0.0909	0.0896	0.0847	0.0847	0.0990	0.0992	0.0979	0.0917	0.0917
$n = 100$	0.0702	0.0705	0.0697	0.0664	0.0664	0.0789	0.0791	0.0783	0.0743	0.0743
$n = 120$	0.0589	0.0592	0.0586	0.0563	0.0563	0.0649	0.0651	0.0645	0.0617	0.0617
$n = 150$	0.0466	0.0468	0.0464	0.0449	0.0449	0.0520	0.0522	0.0517	0.0498	0.0498
$n = 180$	0.0391	0.0392	0.0389	0.0378	0.0378	0.0432	0.0433	0.0430	0.0416	0.0416
$n = 200$	0.0346	0.0347	0.0344	0.0336	0.0336	0.0388	0.0389	0.0386	0.0375	0.0375
$n = 250$	0.0276	0.0276	0.0275	0.0269	0.0269	0.0308	0.0309	0.0307	0.0300	0.0300
$n = 500$	0.0138	0.0138	0.0138	0.0136	0.0136	0.0152	0.0152	0.0152	0.0150	0.0150

Table 11. MSE results obtained from$p = 5$, $\rho = 0.6$ and $\rho = 0.7$, different sample sizes and different ridge parameter estimators.

	$p = 5$, $\rho = 0.6$					$p = 5$, $\rho = 0.7$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.5120	0.4296	0.4867	0.4469	0.4296	0.5861	0.4385	0.5496	0.4757	0.4385
$n = 30$	0.3297	0.3034	0.3191	0.2803	0.2803	0.3945	0.3324	0.3757	0.3110	0.3110
$n = 50$	0.1890	0.1849	0.1852	0.1657	0.1657	0.2333	0.2206	0.2275	0.1951	0.1951
$n = 60$	0.1565	0.1547	0.1540	0.1397	0.1397	0.1934	0.1864	0.1894	0.1654	0.1654
$n = 80$	0.1167	0.1161	0.1153	0.1061	0.1061	0.1441	0.1413	0.1418	0.1267	0.1267
$n = 100$	0.0925	0.0924	0.0916	0.0856	0.0856	0.1140	0.1128	0.1125	0.1027	0.1027
$n = 120$	0.0753	0.0754	0.0747	0.0707	0.0707	0.0950	0.0944	0.0939	0.0864	0.0864
$n = 150$	0.0603	0.0604	0.0599	0.0571	0.0571	0.0755	0.0753	0.0749	0.0702	0.0702
$n = 180$	0.0504	0.0504	0.0501	0.0481	0.0481	0.0626	0.0625	0.0622	0.0588	0.0588
$n = 200$	0.0457	0.0457	0.0454	0.0438	0.0438	0.0572	0.0571	0.0568	0.0541	0.0541
$n = 250$	0.0361	0.0361	0.0359	0.0348	0.0348	0.0458	0.0458	0.0455	0.0436	0.0436
$n = 500$	0.0179	0.0179	0.0179	0.0175	0.0175	0.0225	0.0225	0.0224	0.0219	0.0219

Table 12. MSE results obtained from $p = 5$, $\rho = 0.8$ and $\rho = 0.9$, different sample sizes and different ridge parameter estimators.

	$p = 5$, $\rho = 0.8$					$p = 5$, $\rho = 0.9$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.7185	0.4432	0.6565	0.4925	0.4432	0.9578	0.4379	0.8395	0.4881	0.4379
$n = 30$	0.5068	0.3627	0.4737	0.3457	0.3457	0.7654	0.3723	0.6752	0.3559	0.3559
$n = 50$	0.3160	0.2703	0.3030	0.2332	0.2332	0.5204	0.3272	0.4765	0.2684	0.2684
$n = 60$	0.2629	0.2351	0.2541	0.2004	0.2004	0.4468	0.3095	0.4150	0.2461	0.2461
$n = 80$	0.1980	0.1865	0.1933	0.1602	0.1602	0.3535	0.2752	0.3344	0.2135	0.2135
$n = 100$	0.1616	0.1556	0.1585	0.1352	0.1352	0.2901	0.2429	0.2775	0.1881	0.1881
$n = 120$	0.1339	0.1303	0.1317	0.1145	0.1145	0.2434	0.2133	0.2347	0.1668	0.1668
$n = 150$	0.1082	0.1064	0.1067	0.0948	0.0948	0.2008	0.1834	0.1952	0.1454	0.1454
$n = 180$	0.0890	0.0881	0.0880	0.0803	0.0803	0.1671	0.1567	0.1632	0.1263	0.1263
$n = 200$	0.0816	0.0809	0.0808	0.0738	0.0738	0.1517	0.1441	0.1487	0.1177	0.1177
$n = 250$	0.0648	0.0645	0.0643	0.0595	0.0595	0.1248	0.1204	0.1227	0.1001	0.1001
$n = 500$	0.0322	0.0322	0.0321	0.0309	0.0309	0.0631	0.0625	0.0626	0.0550	0.0550

Table 13. MSE results obtained from $p = 7$, $\rho = 0.1$ and $\rho = 0.2$, different sample sizes and different ridge parameter estimators

	$p = 7$, $\rho = 0.1$					$p = 7$, $\rho = 0.2$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.5186	0.5393	0.5030	0.5570	0.5030	0.4645	0.4894	0.4564	0.5230	0.4564
$n = 30$	0.3192	0.3327	0.3057	0.2806	0.2806	0.3294	0.3637	0.3182	0.2607	0.2607
$n = 50$	0.2011	0.2129	0.1892	0.1528	0.1528	0.2617	0.2880	0.2485	0.1542	0.1542
$n = 60$	0.1739	0.1843	0.1631	0.1286	0.1286	0.2472	0.2701	0.2349	0.1431	0.1431
$n = 80$	0.1433	0.1516	0.1340	0.1033	0.1033	0.2279	0.2453	0.2171	0.1351	0.1351
$n = 100$	0.1266	0.1335	0.1184	0.0904	0.0904	0.2187	0.2327	0.2091	0.1353	0.1353
$n = 120$	0.1146	0.1204	0.1075	0.0823	0.0823	0.2138	0.2256	0.2053	0.1380	0.1380
$n = 150$	0.1045	0.1093	0.0985	0.0756	0.0756	0.2070	0.2166	0.1997	0.1408	0.1408
$n = 180$	0.0968	0.1008	0.0916	0.0708	0.0708	0.2025	0.2105	0.1963	0.1438	0.1438
$n = 200$	0.0936	0.0972	0.0887	0.0689	0.0689	0.1997	0.2069	0.1940	0.1449	0.1449
$n = 250$	0.0871	0.0900	0.0830	0.0651	0.0651	0.1965	0.2023	0.1917	0.1489	0.1489
$n = 500$	0.0746	0.0762	0.0724	0.0591	0.0591	0.1886	0.1915	0.1860	0.1576	0.1576

Table 14. MSE results obtained from $p = 7$, $\rho = 0.3$ and $\rho = 0.4$, different sample sizes and different ridge parameter estimators.

	$p = 7$, $\rho = 0.3$					$p = 7$, $\rho = 0.4$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.4089	0.4532	0.4205	0.4623	0.4089	0.3586	0.4152	0.3902	0.3956	0.3586
$n = 30$	0.3325	0.3849	0.3380	0.2393	0.2393	0.3129	0.3741	0.3377	0.2266	0.2266
$n = 50$	0.3014	0.3372	0.2999	0.1708	0.1708	0.3030	0.3424	0.3151	0.1856	0.1856
$n = 60$	0.2957	0.3257	0.2929	0.1698	0.1698	0.2998	0.3321	0.3089	0.1897	0.1897
$n = 80$	0.2872	0.3097	0.2839	0.1762	0.1762	0.2987	0.3227	0.3047	0.2025	0.2025
$n = 100$	0.2833	0.3013	0.2800	0.1853	0.1853	0.2970	0.3158	0.3013	0.2138	0.2138
$n = 120$	0.2792	0.2941	0.2761	0.1927	0.1927	0.2956	0.3111	0.2990	0.2224	0.2224
$n = 150$	0.2760	0.2878	0.2732	0.2020	0.2020	0.2952	0.3074	0.2977	0.2327	0.2327
$n = 180$	0.2753	0.2851	0.2728	0.2098	0.2098	0.2938	0.3038	0.2957	0.2394	0.2394
$n = 200$	0.2745	0.2832	0.2722	0.2139	0.2139	0.2937	0.3027	0.2954	0.2433	0.2433
$n = 250$	0.2724	0.2793	0.2704	0.2208	0.2208	0.2934	0.3005	0.2947	0.2507	0.2507
$n = 500$	0.2689	0.2723	0.2678	0.2365	0.2365	0.2924	0.2959	0.2930	0.2657	0.2657

Table 15. MSE results obtained from $p = 7$, $\rho = 0.5$ and $\rho = 0.6$, different sample sizes and different ridge parameter estimators.

	$p = 7$, $\rho = 0.5$					$p = 7$, $\rho = 0.6$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.3135	0.3705	0.3522	0.3437	0.3135	0.2709	0.3198	0.3078	0.3010	0.2709
$n = 30$	0.2781	0.3398	0.3151	0.2152	0.2152	0.2401	0.2982	0.2820	0.2030	0.2030
$n = 50$	0.2765	0.3156	0.2980	0.1897	0.1897	0.2409	0.2790	0.2679	0.1834	0.1834
$n = 60$	0.2769	0.3095	0.2945	0.1940	0.1940	0.2430	0.2744	0.2650	0.1873	0.1873
$n = 80$	0.2777	0.3014	0.2901	0.2063	0.2063	0.2460	0.2691	0.2621	0.1972	0.1972
$n = 100$	0.2790	0.2976	0.2886	0.2170	0.2170	0.2469	0.2651	0.2595	0.2041	0.2041
$n = 120$	0.2786	0.2939	0.2864	0.2241	0.2241	0.2481	0.2630	0.2584	0.2101	0.2101
$n = 150$	0.2796	0.2916	0.2857	0.2331	0.2331	0.2493	0.2611	0.2574	0.2169	0.2169
$n = 180$	0.2791	0.2890	0.2841	0.2386	0.2386	0.2495	0.2592	0.2561	0.2209	0.2209
$n = 200$	0.2796	0.2884	0.2839	0.2420	0.2420	0.2501	0.2588	0.2561	0.2237	0.2237
$n = 250$	0.2794	0.2864	0.2829	0.2475	0.2475	0.2504	0.2573	0.2551	0.2278	0.2278
$n = 500$	0.2800	0.2834	0.2817	0.2600	0.2600	0.2514	0.2548	0.2537	0.2371	0.2371

Table 16. MSE results obtained from $p = 7$, $\rho = 0.7$ and $\rho = 0.8$, different sample sizes and different ridge parameter estimators.

	$p = 7$, $\rho = 0.7$					$p = 7$, $\rho = 0.8$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.2430	0.2762	0.2688	0.2676	0.2430	0.2222	0.2329	0.2282	0.2330	0.2222
$n = 30$	0.2062	0.2562	0.2456	0.1921	0.1921	0.1787	0.2181	0.2112	0.1772	0.1772
$n = 50$	0.2054	0.2419	0.2346	0.1726	0.1726	0.1726	0.2072	0.2022	0.1592	0.1592
$n = 60$	0.2079	0.2385	0.2324	0.1749	0.1749	0.1746	0.2045	0.2004	0.1601	0.1601
$n = 80$	0.2102	0.2330	0.2285	0.1801	0.1801	0.1782	0.2015	0.1985	0.1633	0.1633
$n = 100$	0.2130	0.2311	0.2276	0.1858	0.1858	0.1805	0.1993	0.1969	0.1666	0.1666
$n = 120$	0.2146	0.2296	0.2266	0.1903	0.1903	0.1823	0.1980	0.1961	0.1693	0.1693
$n = 150$	0.2156	0.2274	0.2251	0.1945	0.1945	0.1841	0.1966	0.1951	0.1725	0.1725
$n = 180$	0.2167	0.2265	0.2246	0.1979	0.1979	0.1851	0.1955	0.1942	0.1746	0.1746
$n = 200$	0.2172	0.2260	0.2243	0.1998	0.1998	0.1859	0.1952	0.1940	0.1759	0.1759
$n = 250$	0.2181	0.2250	0.2237	0.2030	0.2030	0.1870	0.1944	0.1935	0.1782	0.1782
$n = 500$	0.2195	0.2229	0.2222	0.2097	0.2097	0.1890	0.1927	0.1922	0.1829	0.1829

Table 17. MSE results obtained from $p = 7$ and $\rho = 0.9$, different sample sizes and different ridge parameter estimators.

	$p = 7$, $\rho = 0.9$				
	k_1	k_2	k_3	k_{robust}	MSE_{min}
$n = 20$	0.2080	0.1904	0.1876	0.1868	0.1868
$n = 30$	0.1582	0.1823	0.1777	0.1564	0.1564
$n = 50$	0.1442	0.1765	0.1730	0.1455	0.1442
$n = 60$	0.1447	0.1747	0.1718	0.1455	0.1447
$n = 80$	0.1474	0.1728	0.1706	0.1474	0.1474
$n = 100$	0.1497	0.1711	0.1694	0.1492	0.1492
$n = 120$	0.1519	0.1703	0.1689	0.1509	0.1509
$n = 150$	0.1541	0.1692	0.1681	0.1528	0.1528
$n = 180$	0.1560	0.1686	0.1677	0.1544	0.1544
$n = 200$	0.1568	0.1682	0.1674	0.1550	0.1550
$n = 250$	0.1585	0.1677	0.1671	0.1566	0.1566
$n = 500$	0.1618	0.1664	0.1661	0.1596	0.1596

20 sample size). Because of the low probability of encountering sample size $n = 20$ in large-scale studies, the deviation can be ignored. As a result, it can be said that the k_{robust} parameter estimator yields the smallest MSE in almost all cases.

4. A numerical example

In this section for illustration we apply an example of the recently published [23] data set to compare the performance of the proposed robust ridge estimator with the estimators found to be best in the study of Göktaş and Sevinç [6]. The market historical data set of real estate valuation is collected in 2018 from Sindian District New Taipei City, Taiwan. The following linear regression model has been considered.

$$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \beta_6 X_{i6} + \varepsilon_i \tag{27}$$

where Y is the house price of unit area (10,000 New Taiwan Dollar/Ping, where ping is a local unit, 1 Ping = 3.3 meter squared), X_1 is the transaction date, X_2 is the house age (unit: year), X_3 is the distance to the nearest MRT station (unit: meter), X_4 is the number of convenience stores in the living circle on foot (integer), X_5 is the geographic coordinate, latitude (unit: degree), X_6 is the geographic coordinate, longitude (unit: degree). It is observed from correlation Table 18 that the some of the regressors (X_3, X_4, X_5 and X_6) are highly inter-correlated. This, in fact, implies the existence of multi-collinearity in the data set. So, the data set can be used to compare the proposed robust ridge estimator with the others and we are dealing with the dataset in different aspect. The study of Yeh, and Hsu [23] avoids collinearity and use three different methods to present best functional relationship between house pricing with other factors. Therefore, in this regard it is advantageous to use a prediction method eliminating the effect of multicollinearity like a ridge method with best ridge parameter estimator representative proposed in the study.

Table 18. Pearson correlation coefficient matrix of the regressors.

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.000	0.018	0.061	0.010	0.035	−0.041
X_2	0.018	1.000	0.026	0.050	0.054	−0.049
X_3	0.061	0.026	1.000	**−0.603**	**−0.591**	**−0.806**
X_4	0.010	0.050	**−0.603**	1.000	**0.444**	**0.449**
X_5	0.035	0.054	**−0.591**	**0.444**	1.000	**0.413**
X_6	−0.041	−0.049	**−0.806**	**0.449**	**0.413**	1.000

Table 19. The MSE and the estimated regression coefficients of the estimators.

Estimators	OLS	k_1	k_2	k_3	k_{robust}	VIF
Ridge parameter estimator	0	0.252842	1.231537	0.735769	**6.169428**	
MSE	0.0128002	0.012736	0.012514	0.012621	**0.011960**	
$\hat{\beta}_1$	0.106714	0.106618	0.106249	0.106435	**0.105868**	**1.1556**
$\hat{\beta}_2$	−0.225812	−0.225669	−0.225116	−0.225396	**−0.224536**	**1.2148**
$\hat{\beta}_3$	−0.416252	−0.415410	−0.412213	−0.413820	**−0.408956**	**32.4496**
$\hat{\beta}_4$	0.245345	0.245366	0.245438	0.245404	**0.245497**	**14.4316**
$\hat{\beta}_5$	0.205647	0.205739	0.206084	0.205912	**0.206423**	**10.8921**
$\hat{\beta}_6$	−0.014020	−0.013377	−0.010937	−0.012164	**−0.008452**	**30.4963**

As seen in Table 19, the estimation of each ridge parameter varies from 0.25 to 6.17. Each of the estimated ridge parameters has little effect on the estimated regression parameters. From the MSE results, the proposed robust estimator did perform better in comparison with the other three best ridge parameter estimators. That means the proposed ridge parameter estimator performs parallel to the results obtained from the simulation studies (for verification see Tables 12 and 16). This example simply shows the applicability of the new proposed ridge estimator and its priority in practice.

5. The concluding remarks

In the current study, a new robust k parameter estimator has been proposed according to the search method by selecting the parameters that perform best among the many ridge parameter estimators proposed in the literature so far. This proposed new k_{robust} parameter estimator was compared to other parameter estimators known to be the best. As a result of the comparison, k_{robust} parameter was found to give smaller MSE results compared to others except in the few cases with small samples. Even in these exceptional cases, the MSE value of the k_{robust} does not differ significantly from the smallest MSE value presented by other ridge parameter estimators. As the sample size increased, the MSE values obtained from the k_{robust} were found to be smaller in every case. As a result, more effective estimates can be obtained by using the k_{robust} parameter estimator we have proposed for the RR method in cases where there is multicollinearity and parameter estimation comes to the fore.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Y.M. Al-Hassan, *Performance of a new ridge estimator*. J. Assoc. Arab Univ. Sci. 9 (2010), pp. 23–26.
[2] M. Alkhamisi, G. Khalaf, and G. Shukur, *Some modifications for choosing ridge parameters*. Commun. Stat.- Theor. M. 35 (2006), pp. 2005–2020.
[3] B. Andersson, *Scandinavian evidence on growth and age structure*, ESPE 1997 Conference at Uppsala University, Sweden, 1998.
[4] A.V. Dorugade, *New ridge parameters for ridge regression*. J. Assoc. Arab Univ. Sci. 15 (2014), pp. 94–99.
[5] D.E. Ferrar, and R.R. Glauber, *Multicollinearity in regression analysis: The problem revisited*. Rev. Econ. Stat. 49 (1967), pp. 92–107.
[6] A. Göktaş, and V. Sevinç, *Two new ridge parameters and a guide for selecting an appropriate ridge parameter in linear regression*. Gazi Univ. J. Sci. 29 (2016), pp. 201–211.
[7] G.H. Golub, M. Heath, and G. Wahba, *Generalized cross-validation as a method for choosing a good ridge parameter*. Technometrics. 21 (1979), pp. 215–223. doi:10.1080/00401706.1979.10489751.
[8] A.E. Hoerl, and R.W. Kennard, *Ridge regression: biased estimation for non-orthogonal problems*. Technometrics. 12 (1970a), pp. 55–67.
[9] A.E. Hoerl, and R.W. Kennard, *Ridge regression: applications to non-orthogonal problems*. Technometrics. 12 (1970b), pp. 69–82.
[10] A.E. Hoerl, R.W. Kennard, and K.F. Baldwin, *Ridge regression: Some simulations*. Commun. Stat. 4 (1975), pp. 105–123.

[11] G. Khalaf, and M. Iguernane, *Multicollinearity and a ridge parameter estimation approach*. J. Mod. Appl. Stat. Methods. 15 (2016), pp. 400–410.
[12] G. Khalaf, and G. Shukur, *Choosing ridge parameter for regression problems*. Commun. Stat.-Theor. M. 34 (2005), pp. 1177–1182.
[13] B.M.G. Kibria, *Performance of some new ridge regression estimators*. Commun. Stat.-Simul. C 32 (2003), pp. 419–435.
[14] J.F. Lawless, and P. Wang, *A simulation study of ridge and other regression estimators*. Commun. Stat.-Theor. M. 5 (1976), pp. 307–323.
[15] A.F. Lukman, and A. Olatunji, *Newly proposed estimator for ridge parameter: An application to the Nigerian economy*. Pak. J. Stat. 34 (2018), pp. 91–98.
[16] K. Mansson, G. Shukur, and B.M.G. Kibria, *A simulation study of some ridge regression estimators under different distributional assumptions*. Commun. Stat.-Simul. C. 39 (2010), pp. 1639–1670.
[17] D.W. Marquardt, *Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation*. Technometrics. 12 (1970), pp. 591–612.
[18] D.W. Marquardt, and R.D. Snee, *Ridge regression in practice*. Am. Stat. 29 (1975), pp. 3–20.
[19] R.L. Mason, R.F. Gunst, and J.T. Webster, *Regression analysis and problems of multicollinearity*. Commun. Stat. 4 (1975), pp. 277–292.
[20] G.C. McDonald, and D.I. Galarneau, *A Monte Carlo evaluation of some ridge-type estimators*. J. Am. Stat. Assoc. 70 (1975), pp. 407–416.
[21] G. Muniz, and B.M.G. Kibria, *On some ridge regression estimators: An empirical comparisons*. Commun. Stat.-Simul. C. 38 (2009), pp. 621–630.
[22] S. Sakallıoğlu, and S. Kaçıranlar, *A new biased estimator based on ridge estimation*. Stat. Pap. 49 (2008), pp. 669–689.
[23] I.C. Yeh, and T.K. Hsu, *Building real estate valuation models with comparative approach through case-based reasoning*. Appl. Soft. Comput. 65 (2018), pp. 260–271.