Journal of Biometrics & Biostatistics
ISSN: 2155-6180

Ghadban and Iguernane, J Biom Biostat 2015, 6:5
DOI: 10.4172/2155-6180.1000264

Research Article Open Access

The Traditional Ordinary Least Squares Estimator under Collinearity

Ghadban AK* and Iguernane M

King Khalid University, Saudi Arabia

Abstract

In a multiple regression analysis, it is usually difficult to interpret the estimator of the individual coefficients if the explanatory variables are highly inter-correlated. Such a problem is often referred to as the multicollinearity problem. There exist several ways to solve this problem. One such way is ridge regression. Two approaches of estimating the shrinkage ridge parameter k are proposed. Comparison is made with other ridge-type estimators. To investigate the performance of our proposed methods with the traditional ordinary least squares (OLS) and the other approaches for estimating the parameters of the ridge regression model, we calculate the mean squares error (MSE) using the simulation techniques. Results of the simulation study shows that the suggested ridge regression outperforms both the OLS estimator and the other ridge-type estimators in all of the different situations evaluated in this paper.

Keywords: Linear regression model; Multicollinearity; Ridge regression estimators; Simulation study

Mathematics Subject Classification: Primary 62J07; Secondary 62J05

Introduction

Consider the standard multiple linear regression model;

$$Y = X \quad + e, \tag{1}$$

where Y is an $(n\times 1)$ vector of responses, X is an $(n\times p)$ matrix of the explanatory variables of full rank p, β is a $(p\times 1)$ vector of unknown regression coefficients, and finally, $e \sim N(0,\sigma^2 I)$ is an $(n\times 1)$ vector of error terms.

The OLS estimator is often used to estimate the regression coefficients β as:

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

The standard assumption in the linear regression analysis is that all the explanatory variables are linearly independent. When this assumption is violated, the problem of multicollinearity enters into the data and it inflates the variance of an ordinary least squares estimator of the regression coefficient. Obtaining the estimators for multicollinear data is an important problem in the literature. In fact, when the problem of multicollinearity is present in the measurement error ridden data, then an important issue is how to obtain the consistent estimators of regression coefficients. One of the most popular estimator for combating multicollinearity is the ridge estimator, originally proposed by Hoerl et al. [1]. They suggested a small positive number (k>0) to be added to the diagonal elements of the $X'X$ matrix from the multiple regression and the resulting estimators are obtained as:

$$\hat{\beta}(k) = (X'X + kI)^{-1}X'Y, \tag{3}$$

which is known as a ridge regression estimator. For a positive value of k, this estimator provides a smaller MSE compared to the OLS estimator, i.e.,

$$MSE(\hat{\beta}(k)) < MSE(\hat{\beta})\cdot$$

Most of the later efforts in this area have concentrated on estimating the value of the ridge parameter k. Many different techniques for estimating k have been proposed by different researchers, for example, Hoerl et al. [1], Hoerl et al. [2] Dempster et al. [3], Gibbons [4], Kibria [5], Khalaf et al. [6], Alkhamisi et al. [7], Khalaf [8] and Khalaf [9].

The plan of the paper is as follows: in Section 2, we present different methods for estimating the parameter of ridge regression together with our proposed estimators. A simulation study has been conducted in Section 3. The simulation results are discussed in Section 4. In Section 5 we give a brief summary and conclusions.

The Proposed Ridge Regression Parameter

In case of ordinary ridge regression, many researchers have suggested different ways of estimating the ridge parameter. Hoerl et al. [1] showed, by letting $\beta_{\max}$ denote the maximum of the β_i, that choosing;

$$\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\hat{\beta}^2_{\max}}, \tag{4}$$

implies that $MSE(\hat{\beta}(k)) < MSE(\hat{\beta})$. The ridge estimator using $\hat{k}_{HK}$ will be denoted by HK.

Hoerl et al. [2] suggested that, the value of k is chosen small enough, for which the MSE of ridge estimator is less than the MSE of OLS estimator. They showed, through simulation, that the use of the ridge with biasing parameter given by:

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}, \tag{5}$$

has a probability greater than 0.50 of producing estimator with a smaller MSE than the OLS estimator, where $\hat{\sigma}^2$ is the usual estimator of σ^2, defined by $\hat{\sigma}^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n-p-1}$. The ridge estimator using Eq. (5) will be denoted by HKB.

The purpose of this study is to modify the approaches of estimating k mentioned in Hoerl and Kennard [1] and Hoerl et al. [2] given in

***Corresponding author:** Ghadban AK, Department of Mathematics, Faculty of Science, King Khalid University, Saudi Arabia, Tel: 0172418000; E-mail: albadran50@yahoo.com

Received November 23, 2015; **Accepted** December 02, 2015; **Published** December 09, 2015

Citation: Ghadban AK, Iguernane M (2015) The Traditional Ordinary Least Squares Estimator under Collinearity. J Biom Biostat 6: 264. doi:10.4172/2155-6180.1000264

Citation: Ghadban AK, Iguernane M (2015) The Traditional Ordinary Least Squares Estimator under Collinearity. J Biom Biostat 6: 264. doi:10.4172/2155-6180.1000264

equations (4) and (5), to suggest the following two estimators:

$$\hat{k}_1 = \begin{cases} \sqrt{HK + HKB}, & \textit{if the sum} < 1 \\ \sqrt{\dfrac{HK + HKB}{2}}, & \textit{if the sum} > 1 \end{cases} \quad (6)$$

$$\hat{k}_2 = \begin{cases} \sqrt{HK + HKB}, & \textit{if the sum} < 1 \\ \sqrt{\dfrac{HK + HKB}{p}}, & \textit{if the sum} > 1 \end{cases} \quad (7)$$

where p denotes the number of parameters (excluding the intercept). The ridge estimators using $\hat{k}_1$ and $\hat{k}_2$ will be denoted by KI_1 and KI_2, respectively.

The performance of these proposed estimators will be then compared with the traditional OLS estimation and those of HK and HKB estimators in terms of MSE. This will mainly be done by means of simulations under conditions where the sample size n, the number of the explanatory variables p and the strength of correlations between the explanatory variables are varied.

The Simulation Study

This section consists of a brief description of how the data is generated together with a discussion about the different factors varied in the simulation study. Also the criteria for judging the performance of the different estimation methods are presented.

The design of the experiment

Following McDonald et al. [10], the explanatory variables are generated by

$$x_{ij} = (1-\rho^2)^{\frac{1}{2}} z_{ij} + \rho z_{ip}, \qquad i = 1,2,\ldots,n \qquad j = 1,2,\ldots,p$$

where z_{ij} are independent standard normal pseudo-random numbers, and ρ is specified so that the correlation between any two explanatory variables is given by ρ^2. Four different sets of correlation are considered, corresponding to $\rho =$ 0.60, 0.90, 0.94 and 0.98. The explanatory variables are then standardized so that $X'X$ is in correlation form.

Observations on the dependent variable are determined by

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_p x_{ip} + e_i, \qquad i = 1,2,\ldots,n$$

Where β_0 is taken to be identically zero. Four values of σ^2 are considered which are 0.8, 0.9, 0.95 and 0.99. Then the dependent variable is standardized so that $X'y$ is the vector of correlation of dependent variable with each explanatory variable. In this experiment, we choose $p = 7$ and 10 for $n = 15$, 25, 80 and 200. Then the experiment is replicated 5000 times by generating new error terms.

Judging the performance of the estimators

To investigate the performance of the different proposed ridge regression estimators and the OLS method, we calculate the MSE using the following equation:

$$MSE = \frac{\sum_{i=1}^{R} (\hat{\beta} - \beta)'_i (\hat{\beta} - \beta)_i}{R},$$

where $\hat{\beta}$ is the estimator of β obtained from the OLS or the other different ridge parameters, and R equals 5000 which corresponds to the number of replications used in the situation.

The Simulation Results

Ridge estimators are constructed with the aim of having smaller MSE than the MSE of the OLS estimator. Improvement, if any, can therefore be studied by looking at the amount of the MSE. These MSEs are reported in Tables 1 and 2. The results of our simulation study indicate that ridge estimators outperform OLS estimator in all cases and the suggested estimators KI^1 and KI^2 performed very well in this study. They appear to offer an opportunity for large reduction in MSE, especially when the sample size and the correlation between the explanatory variables are high (Table 1).

		σ^2 = 0.8				
ρ	n	OLS	HK	HKB	KI_1	KI_2
0.6	15	8.61	4.20	2.74	1.76	1.94
	25	1.064	0.975	0.783	0.769	0.811
	80	0.2517	0.2479	0.2347	0.2350	0.2386
	200	0.0937	0.0932	0.0913	0.0914	0.0920

		σ^2 = 0.8				
ρ	n	OLS	HK	HKB	KI_1	KI_2
0.9	15	35.88	12.72	7.70	2.46	2.63
	25	4.188	3.058	1.830	1.466	1.582
	80	1.012	0.937	0.718	0.709	0.751
	200	0.3793	0.3688	0.3251	0.3287	0.3408

		σ^2 = 0.8				
ρ	n	OLS	HK	HKB	KI_1	KI_2
0.94	15	65.873	21.228	14.814	3.0203	3.1189
	25	7.4153	4.6495	2.6701	1.7583	1.8506
	80	1.7701	1.5495	1.0626	0.9965	1.0601
	200	0.6507	0.6189	0.5043	0.5066	0.5330

		σ^2 = 0.8				
ρ	n	OLS	HK	HKB	KI_1	KI_2
0.98	15	218.62	63.98	46.10	4.37	4.39
	25	23.351	10.661	6.373	2.388	2.410
	80	5.451	3.788	2.167	1.5789	1.6223
	200	2.102	1.777	1.154	1.039	1.087

		σ^2 = 0.8				
ρ	n	OLS	HK	HKB	KI_1	KI_2
0.6	15	6.8406	3.5983	2.3898	1.5982	1.7575
	25	0.8384	0.7839	0.6537	0.6477	0.6789
	80	0.1962	0.1938	0.1851	0.1854	0.1878
	200	0.0741	0.0738	0.0726	0.0727	0.0731

		σ^2 = 0.8				
ρ	n	OLS	HK	HKB	KI_1	KI_2
0.9	15	27.622	10.530	6.556	2.364	2.526
	25	3.340	2.566	1.583	1.342	1.456
	80	0.7973	0.7512	0.6008	0.5988	0.6307
	200	0.2998	0.2934	0.2654	0.2686	0.2771

Citation: Ghadban AK, Iguernane M (2015) The Traditional Ordinary Least Squares Estimator under Collinearity. J Biom Biostat 6: 264. doi:10.4172/2155-6180.1000264

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.94	15	47.988	16.688	10.105	2.680	2.805
	25	5.773	3.915	2.268	1.645	1.755
	80	1.3607	1.2231	0.8749	0.8433	0.8977
	200	0.5184	0.4978	0.4182	0.4223	0.4427

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.98	15	158.08	49.069	31.913	4.064	4.102
	25	18.470	8.957	5.164	2.215	2.250
	80	4.364	3.193	1.864	1.453	1.507
	200	1.6500	1.4402	0.9722	0.9059	0.9551

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.6	15	6.440	3.459	2.228	1.539	1.690
	25	0.7611	0.7140	0.6001	0.5955	0.6229
	80	0.1765	0.1745	0.1676	0.1678	0.1698
	200	0.0659	0.0657	0.0646	0.0647	0.0651

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.9	15	26.715	10.069	6.155	2.219	2.383
	25	3.049	2.386	1.489	1.293	1.407
	80	0.7101	0.6728	0.5464	0.5474	0.5761
	200	0.2645	0.2594	0.2367	0.2395	0.2466

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.94	15	44.747	14.785	9.448	2.656	2.783
	25	5.184	3.587	2.097	1.567	1.673
	80	1.2181	1.1079	0.8093	0.7888	0.8403
	200	0.4629	0.4463	0.3803	0.3841	0.4012

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.98	15	150.20	44.75	30.23	3.92	3.97
	25	16.983	8.301	4.901	2.175	2.213
	80	3.896	2.923	1.726	1.384	1.438
	200	1.4606	1.2923	0.8920	0.8437	0.8918

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.6	15	5.848	3.082	1.985	1.436	1.590
	25	0.7038	0.6658	0.5655	0.5622	0.5860
	80	0.1639	0.1622	0.1559	0.1562	0.1581
	200	0.0620	0.0618	0.0609	0.0610	0.0613

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.9	15	23.002	9.022	5.371	2.153	2.339
	25	2.750	2.196	1.383	1.221	1.326
	80	0.6726	0.6394	0.5243	0.5256	0.5521
	200	0.2439	0.2395	0.2198	0.2225	0.2288

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.94	15	38.996	13.866	8.865	2.532	2.675
	25	4.918	3.444	2.021	1.536	1.648
	80	1.1318	1.0341	0.7638	0.7490	0.7986
	200	0.4228	0.4089	0.3521	0.3563	0.3718

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.98	15	132.45	41.75	27.32	3.72	3.77
	25	15.685	7.873	4.621	2.162	2.206
	80	3.702	2.825	1.686	1.362	1.415
	200	1.3698	1.2222	0.8535	0.8128	0.8583

Table 1: Estimated MSE when p = 7.

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.6	15	6.5313	4.2852	2.4810	2.0218	2.2595
	25	1.9391	1.7640	1.2815	1.2552	1.3536
	80	0.3799	0.3751	0.3504	0.3513	0.3578
	200	0.1412	0.1406	0.1370	0.1372	0.1382

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.9	15	27.1747	12.4955	6.2813	3.0539	3.2615
	25	7.7115	5.5316	2.8908	2.2409	2.4154
	80	1.5642	1.4634	1.0591	1.0417	1.1066
	200	0.5785	0.5647	0.4834	0.4884	0.5090

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.94	15	46.8813	20.0779	9.6678	3.5778	3.7064
	25	13.3331	8.2681	4.1480	2.6650	2.7821
	80	2.7026	2.4034	1.5277	1.4297	1.5126
	200	1.0056	0.9615	0.7435	0.7435	0.7848

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.98	15	152.1745	53.8533	28.0708	5.4931	5.5123
	25	42.4797	19.5072	10.2101	3.8327	3.8562
	80	8.6617	6.2425	3.2356	2.3936	2.4296
	200	3.2246	2.7729	1.6640	1.5011	1.5472

σ^2 = 0.8

ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.6	15	5.1583	3.5488	2.0914	1.7885	2.0080
	25	1.5015	1.3916	1.0632	1.0496	1.1219
	80	0.3079	0.3048	0.2883	0.2887	0.2929
	200	0.1129	0.1125	0.1102	0.1103	0.1110

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.9	15	21.5841	10.4375	5.2976	2.8178	3.0594
	25	6.0798	4.5944	2.4765	2.0355	2.2102
	80	1.2484	1.1829	0.8924	0.8857	0.9376
	200	0.4509	0.4427	0.3907	0.3949	0.4091

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.94	15	36.5916	15.8731	8.0752	3.3272	3.4763
	25	10.5081	6.9777	3.5163	2.4477	2.5894
	80	2.1424	1.9455	1.2911	1.2373	1.3161
	200	0.7777	0.7505	0.6040	0.6075	0.6392

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.98	15	121.9148	44.9126	23.0672	5.0840	5.1219
	25	34.4905	16.7414	8.4861	3.5828	3.6167
	80	6.9268	5.2124	2.7602	2.1686	2.2207
	200	2.5272	2.2342	1.3946	1.2931	1.3449

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.6	15	4.8443	3.3635	1.9976	1.7168	1.9370
	25	1.3692	1.2774	0.9866	0.9745	1.0372
	80	0.2727	0.2702	0.2568	0.2572	0.2608
	200	0.1016	0.1014	0.0997	0.0998	0.1003

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.9	15	19.1460	9.6208	4.7720	2.6614	2.9010
	25	5.4160	4.2002	2.2972	1.9271	2.0963
	80	1.1199	1.0674	0.8218	0.8220	0.8708
	200	0.4083	0.4015	0.3579	0.3621	0.3745

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.94	15	32.4676	14.0097	7.1604	3.1164	3.2841
	25	9.5260	6.4257	3.2854	2.3640	2.5141
	80	1.9154	1.7555	1.1939	1.1542	1.2284
	200	0.6905	0.6690	0.5486	0.5540	0.5824

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.98	15	106.6568	38.7463	20.5424	4.8432	4.8767
	25	30.9499	15.4485	7.8942	3.4899	3.5305
	80	6.1767	4.7439	2.5543	2.0581	2.1169
	200	2.2575	2.0205	1.2931	1.2172	1.2730

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.6	15	4.2840	3.1024	1.8590	1.6290	1.8248
	25	1.2607	1.1830	0.9290	0.9234	0.9818
	80	0.2531	0.2510	0.2396	0.2400	0.2431
	200	0.0927	0.0924	0.0909	0.0909	0.0914

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.9	15	17.4206	8.8620	4.4007	2.5474	2.7966
	25	5.0021	3.9189	2.1558	1.8498	2.0298
	80	1.0351	0.9899	0.7723	0.7738	0.8181
	200	0.3760	0.3700	0.3317	0.3356	0.3469

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.94	15	29.8622	13.6424	6.7319	3.1074	3.2918
	25	8.6360	5.9772	3.0399	2.2691	2.4267
	80	1.7623	1.6261	1.1247	1.0960	1.1692
	200	0.6306	0.6130	0.5108	0.5165	0.5419

$\sigma^2 = 0.8$						
ρ	*n*	*OLS*	*HK*	*HKB*	KI_1	KI_2
0.98	15	100.3572	36.5617	19.2035	4.6061	4.6499
	25	28.1448	14.4566	7.2695	3.3876	3.4369
	80	5.6109	4.3818	2.3663	1.9535	2.0138
	200	2.1242	1.9125	1.2351	1.1687	1.2243

Table 2: Estimated MSE when p = *10.*

If we focus on the values of σ^2, ρ and the sample size *n*, we find that among the ridge estimators considered, KI^1 is the best followed by HKB and KI^2.

In comparing Tables 1 and 2 which involve p = 7 and p = 10, respectively, we find that the MSEs are lowest for Table 1. This is to say that the ridge estimators are more helpful when high multicollinearity exists and the number of explanatory is not large.

Conclusions

In this article, we introduce two alternatives ridge estimators and study their performance using simulation techniques. Comparisons are made with other ridge type estimators evaluated elsewhere. The results from the simulation study show that the sample size, the correlation between the independent variables and the number of explanatory variables are important factors for the performance of the different estimation methods. In most of the cases, the MSE decreases when the first two factors increase.

The result also shows that, with respect to MSE criteria, the proposed ridge regression methods out performs both the OLS estimator and the estimators of Hoerl et al. [1] and Hoerl et al. [2] in all cases investigated. The use of the proposed estimators is recommended since it reduces the MSE substantially in all of the different situations investigated in this paper.

References

1. Hoerl AE, Kennard RW (1970) Ridge Regression: Biased Estimation for non-orthogonal Problems. Technometrics 12: 55-67.
2. Hoerl AE, Kennard RW, Baldwin KF (1975) Ridge Regression: Some Simulation. Communications in Statistics - Theory and Methods 4: 105-124.
3. Dempster AP, Schatzoff M, Wermuth N (1977) A simulation study of alternatives to ordinary least squares. J Amer Statist Assoc 72: 77-91.
4. Gibbons DG (1981) A simulation study of some ridge estimators. Journal of the American Statistical Association 76: 131-139.
5. Kibria BMG (2003) Performance of some New Ridge Regression Estimators.

Communication in Statistics-Theory and Methods 32: 419-435.

6. Khalaf G, Shukur G (2005) Choosing Ridge Parameters for Regression Problems. Communication in Statistics - Theory and Methods 34: 1177-1182.

7. Alkhamisi M, Shukur G (2008) Developing Ridge Parameters for SUR Model. Communication in Statistics-Theory and Methods 37: 544-564.

8. Khalaf G (2011) Ridge Regression: An Evaluation to some New Modifications. International Journal of Statistics and Analysis 1: 325-342.

9. Khalaf G (2013) A Comparison between Biased and Unbiased Estimators. JMASM 12: 293-303.

10. McDonald GC, Galarneau DI (1975) A Monte Carlo evaluation of some ridge-type estimators. J Amer Statist Assoc 70: 407-416.