Some Improved Classification-Based Ridge Parameter Of Hoerl And Kennard Estimation Techniques

[1]Adewale F. Lukmanand [1]Kayode Ayinde
[1] Department of Statistics, Ladoke Akintola University of Technology, P.M.B. 4000, Ogbomoso, Oyo State, Nigeria.Email address: wale3005@yahoo.com, kayinde@lautech.edu.ng

Abstract

In a linear regression model, it is often assumed that the explanatory variables are independent. This assumption is often violated and Ridge Regression estimator introduced by [2]has been identified to be more efficient than ordinary least square (OLS) in handling it. However, it requires a ridge parameter, K, of which many have been proposed. In this study, estimators based on Hoerl and Kennard were classified into different forms and various types and some modifications were proposed to improveit. Investigation were done by conducting 1000 Monte-Carlo experiments under five (5) levels of multicollinearity, three (3) levels of error variance and five levels of sample size. For the purpose of comparing the performance of the improved ridge parameter with the existing ones, the number of times the MSE of the improved ridge parameter is less than the existing ones is counted over the levels of multicollinearity (5) and error variance (3). Also, a maximum of fifteen (15) counts is expected. Results show that the improved ridge parameters proposed in this study are better than the existing ones.

Keywords: Linear Regression Model, Multicollinearity, Ridge Parameter Estimation Techniques, Relative Efficiency

---

1.0 Introduction

A general linear regression model is defined in matrix form as:

$$Y = X\beta + U, \quad (1)$$

where X is an $n \times p$ matrix with full rank, Y is a $n \times 1$ vector of dependent variable, β is a $p \times 1$ vector of unknown parameters, and U is the error term such that :

$E(\mathbf{U}) = \mathbf{0}$ and $\mathbf{E}($ EMBED Equation.DSMT4 $)=\sigma^2 \mathbf{I_n}$. The Ordinary Least Square (OLS) estimator commonly used to estimate the regression parameter β in (1) is defined as

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

Gauss Markov theorem states that the OLS estimator in the class of unbiased estimators has minimum variance, that is, they are best linear unbiased estimator [1]. The theorem holds as long as the assumptions of classical linear regression model are satisfied. However, if one or more of these assumptions do not hold, OLS is no longer the best linear unbiased estimator (BLUE). A pertinent example is whentwo or more explanatory variables are linearly related. Consequently, the performance of OLS estimator is unsatisfactory when the explanatory variables are related.The regression coefficients is determinate but cannot be estimated with great precision and sometimes have wrong signs [1]. Several methods have been suggested in literature to solve this problem. [2] introduced the method of ridge regression whichis generally acceptable as alternative to the OLS estimator to handle the problem of multicollinearity. They suggested the addition of ridge parameter K to the diagonal of $\mathbf{X'X}$ matrix in (2). Therefore, ridgeestimator is defined as:

$$\hat{\beta}_R = (X'X + K)^{-1}X'Y, \quad (3)$$

where K is a diagonal matrix of non-negative constants that is K≥0. Though thisestimator is biased but it gives a smaller mean squared error when compared to the

OLS estimator for a positive value of K [2]. The use of the estimator depends largely on the ridge parameter, K. Several methods for estimating this ridge parameter have been proposed as follows: [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15]. The purpose of this study is to apply the modificationin [16, 17], and proposed another modificationto improve the various types and different forms of [2]as classified by [15]. A Simulation study is conducted and the performances of the estimators examined via mean square error (MSE).

2.1 Review of Methods of Estimating the Ridge Parameter

[2] defined the ridge parameter as:

$$K_{HK_i} = \frac{\sigma^2}{\alpha_i^2} \quad (4)$$

They suggested estimating the ridge parameter by taking the maximum (Fixed Maximum) of α_i^2 such that the estimator of K is:

$$\hat{K}_{HK}^{FM} = \frac{\hat{\sigma}^2}{Max(\alpha_i^{\,2})} \quad (5)$$

[18] proposed a different estimator of K by taking the Harmonic Mean of the ridge parameter K_{HK_i}. This estimator is given as:

$$\hat{K}_{HK}^{HM} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p} \alpha_i^2} \quad (6)$$

[15] reviewed that the several methods of estimating the ridge parameters earlier mentioned and observed that the existing ridge parameters followed some different forms such as Fixed Maximum, Varying Maximum, Arithmetic Mean, Harmonic Mean, Geometric Mean and Median) and various types such as Original, Reciprocal, Square Root and Reciprocal of Square Root. This is further illustrated in Table 1

Table 1: Summary of Different Forms and Various Types for $\hat{K}_{HK_i} = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$

	Various Types of K			
	O	R	SR	RSR
	$\hat{K}_{H}^{F}$ [2]	$\hat{K}_{HK}^{FM}$ [15]	$\hat{K}_{HK}^{FMS}$ [15]	$\hat{K}_{HK}^{FMR}$ [15]
	$\hat{K}_{H}^{V}$ [15]	$\hat{K}_{HK}^{VM}$ [15]	$\hat{K}_{HK}^{VMS}$ [12]	$\hat{K}_{HK}^{VMR}$ [12]
	$\hat{K}_{HI}^{AI}$ [9]	$\hat{K}_{HK}^{R}$ [15]	$\hat{K}_{HK}^{AMS}$ [15]	$\hat{K}_{HK}^{AMR}$ [15]
	$\hat{K}_{HI}^{HI}$ [18]	$\hat{K}_{HK}^{HM}$ [15]	$\hat{K}_{HK}^{HMS}$ [15]	$\hat{K}_{HK}^{HMR}$ [15]
	$\hat{K}_{HI}^{GI}$ [9]	$\hat{K}_{HK}^{GM}$ [15]	$\hat{K}_{HK}^{GMS}$ [12]	$\hat{K}_{HK}^{GMR}$ [12]
	$\hat{K}_{HI}^{M}$ [9]	$\hat{K}_{HK}^{MR}$ [15]	$\hat{K}_{HK}^{MSR}$ [12]	$\hat{K}_{HK}^{MRS}$ [12]

Furthermore, [16] present new methods of estimating the ridge parameter K as:

$$\hat{K}_{AS} = \frac{\hat{\sigma}^2}{Max\left(\hat{\alpha}_i^{\,2}\right)} + \frac{1}{\lambda_{max}} \quad (7)$$

[17] suggested the improvement of ridge parameter by introducing variance inflation factor, which is defined as:

$$\hat{K}_{DK} = \text{Max}\left(0, \frac{p\sigma^2}{\sum_{i=1}^{p}\alpha_i^2} - \frac{1}{n\left(VIF_j\right)_{max}}\right) \quad (8)$$

where $VIF_j = \frac{1}{1-R_j^2}$; j=1,2,…,p is the variance inflation factor of j^{th} regressor.

Also, the quantity $\frac{1}{n\lambda_{max}}$ is suggested in this study to improve the ridge estimator.

2.2.Proposed Ridge Parameter

Following [16 and 17], the following quantities are used to improve ridge parameter in this

study: $\frac{1}{\lambda_{max}}$, $\frac{1}{n\lambda_{max}}$ and $\frac{1}{nVIF_{max}}$ and consideredin their different forms and various types

asclassified by [15].The improved version of Hoerl and Kennard using the quantity $\frac{1}{\lambda_{max}}$ is summarized in Table 2.

Table 2: Summary of the Different Forms and Various Types for $\hat{K}_{HK_{Ai}} = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_{max}}$

Different Forms	Various Types of K			
	O	R	SR	RSR
FM	$\hat{K}^{FM}_{HK}$ [16]	$\hat{K}^{FMR}_{HK}$ Proposed	$\hat{K}^{FMSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{FMRSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed
VM	$\hat{K}^{VM}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{VMR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{VMSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{VMRSR}_{HK}$ [12]
AM	$\hat{K}^{AM}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{R}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{AMSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{AMRSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed
HM	$\hat{K}^{HM}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{HMR}_{HK}$ Proposed	$\hat{K}^{HMSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{HMRSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed
GM	$\hat{K}^{GM}_{HK}$ Proposed	$\hat{K}^{GMR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{GMSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{GMRSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed
	$\hat{K}^{MO}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{MR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{MSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed	$\hat{K}^{MRSR}_{HK} + \frac{1}{\lambda_{max}}$ Proposed

3.1 Simulation study

Simulation procedure used by [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15] was also used to generate the explanatory variables in this study: This is given as:

$$X_{ij} = (1 - \rho^2)^{\frac{1}{2}} Z_{ij} + \rho Z_{ip} \quad (9)$$

i=1, 2, 3,…, n. j=1, 2,…p.

where Z_{ij} is independent standard normal distribution with mean zero and unit variance, ρ is the correlation between any two explanatory variables and P is the number of explanatory variables. The value of ρ is taken as 0.8, 0.9, 0.95, 0.99, 0.999 respectively. Thus, the correlations between the variable is the same. In this study, the number of explanatory variable (p) is taken to be three (3) and seven (7).

The considered regression model is of the form:

$$\mathbf{Y_t = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + U_t} \quad (10)$$

where t=1, 2…, n; p=3, 7.

The error term $\mathbf{U_t}$ was generated to be normally distributed with mean zero and variance σ^2, $\mathbf{U_t \sim N(0, \sigma^2)}$. In this study, σ were taken to be 0.5, 1 and 5.

β_0 was taken to be identically zero. When p=3, the values of β were chosen to be: β_1=0.8, β_2=0.1, β_3=0.6. When p=7, the values of β were chosen to be: β_1=0.4, β_2=0.1, β_3=0.6, β_4=0.2, β_5=0.25, β_6=0.3, β_7=0.53. The parameter values were chosen such that $\beta'\beta$ =1 which is a common restriction in simulation studies of this type [12].The sample sizes were varied between 10, 20, 30, 40 and 50. Three different values of σ: 0.5, 1 and 5 were also used. At a specified value of n, p and σ, the fixed Xs are first generated; followed by the U, and the values of Y are then obtained using the regression model. The experiment is repeated 1000 times. The performance of this model is evaluated using mean square error (MSE). The MSE for ridge and OLS are calculated using the equation defineds here $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues of $\mathbf{X'X}$, $\hat{\mathbf{K}}$ is the estimator of the ridge parameter K, α_i is the i^{th} element of the vector $\hat{\alpha} = \mathbf{Q'\hat{\beta}}$ where Q is an orthogonal matrix.

For the purpose of comparing the performance of the improved ridge parameter with the existing ones, the number of times the MSE of the improved ridge parameter is less than the existing ones is counted over the levels of multicollinearity (5) and error variance (3). Also, a maximum of fifteen (15) counts is expected.

4.1 Results and Discussion

The number of times the improved ridge parameters estimator is better than the existing ridge parameter with multicollinearity (5 levels) and error variances (3 levels) effect partial out is summarized in Table 3 and 4. From Table 3 and 4, the best five improved techniques by introducing the quantity $\frac{1}{\lambda_{\mathbf{Max}}}$ are [**FMR**, **VMR**], [**FMRSR**, **VMRSR]** and FMO.Four of the proposed perform better than FMO. The best five improved techniques by introducing the quantity $\frac{1}{n\lambda_{\mathbf{Max}}}$ are **FMO**, [**FMR**, **VMR**], [**FMRSR** and **VMRSR**]. Consequently, by introducing the quantity $\frac{1}{n\mathrm{VIF}_{\mathbf{Max}}}$, results show that the best five improved techniques are **AMO**, **GMO**, **VMO**,

VMSR and **MO**.These perform better than HMO which was proposed by [17].

The relative efficiency of the ridge parameter based on $\hat{\mathbf{K}}_{HKA}$ and $\hat{\mathbf{K}}_{HK}$ at different forms and various types are given in Appendix 1 and 2 respectively. Generally, the results show that the improved ridge parameters perform better than the existing ones especially with the ones identified as the best five. However, the performance of the quantity proposed in this study is also good especially with Fixed Maximum Original

5.1 Conclusion

In this study, some improved ridge parameters are classified into different forms and various types. The performances of these estimators are evaluated through Monte-Carlo Simulation where levels of multicollinearity, sample sizes and error variances have been varied.The number of times the MSE of the improved ridge parameter is less than the existing ones is counted over the levels of multicollinearity (5) and error variance (3). Also, a maximum of fifteen (15) counts is expected. Having counted over all the levels of multicollinearity, sample sizes and error variances, the best five with the highest counts is selected. The improved ridge parameters proposed in this study are better than the existing ones.

Different Forms	Various Types	Methods	P=3					
			10			20		
			$\hat{K}_{HKA}$	$\hat{K}_{HKB}$	$\hat{K}_{HKC}$	$\hat{K}_{HKA}$	$\hat{K}_{HKB}$	$\hat{K}_{HKC}$
Fixed Maximum	Original	FMO	9	12	0	10	13	0
	Reciprocal	FMR	15	15	0	12	11	0
	Square root	FMSR	3	3	12	6	6	7
	Reciprocal of Square root	FMRSR	15	15	5	10	10	2
Varying Maximum	Original	VMO	0	0	15	0	0	13
	Reciprocal	VMR	15	15	0	12	11	0
	Square root	VMSR	0	0	15	0	0	13
	Reciprocal of Square root	VMRSR	15	15	5	10	10	2
Arithmetic Mean	Original	AMO	0	0	15	0	0	13
	Reciprocal	AMR	15	15	0	9	9	0
	Square root	AMSR	0	0	15	2	0	10
	Reciprocal of Square root	AMRSR	15	15	0	9	9	3
Harmonic Mean	Original	HMO	2	2	13	5	2	8
	Reciprocal	HMR	15	15	0	10	10	0
	Square root	HMSR	0	0	12	3	3	8
	Reciprocal of Square root	HMRSR	15	15	3	10	10	4
Geometric Mean	Original	GMO	0	0	15	0	0	13
	Reciprocal	GMR	15	15	0	10	10	0
	Square root	GMSR	0	0	15	3	2	10
	Reciprocal of Square root	GMRSR	15	15	5	10	10	4
Median	Original	MO	0	0	15	2	1	11
	Reciprocal	MR	15	15	2	10	10	1
	Square root	MSR	0	0	13	3	0	8
	Reciprocal of Square root	MRSR	14	14	5	10	10	5

Table 4: Number of times the improved ridge parameters estimators are better than the existing ridge parameter $\hat{K}_{HK}=\frac{(\hat{\sigma}^2}{\hat{\alpha}_i^2}$ with multicollinearity (5 levels) and error variances (3 levels) effect partial out when the number of regressor is seven

Different Forms	Various Types	Methods	P=7					
			10			20		
			$\hat{K}_{HKA}$	$\hat{K}_{HKB}$	$\hat{K}_{HKC}$	$\hat{K}_{HKA}$	$\hat{K}_{HKB}$	$\hat{K}_{HKC}$
Fixed Maximum	Original	FMO	0	5	0	11	14	0
	Reciprocal	FMR	15	15	0	15	15	4
	Square root	FMSR	0	0	15	2	2	11
	Reciprocal of Square root	FMRSR	15	15	9	15	15	9
Varying Maximum	Original	VMO	0	0	15	0	0	15
	Reciprocal	VMR	15	15	0	15	15	4
	Square root	VMSR	0	0	15	0	0	15
	Reciprocal of Square root	VMRSR	15	15	9	15	15	9
Arithmetic Mean	Original	AMO	0	0	15	0	0	15
	Reciprocal	AMR	15	15	1	10	10	1
	Square root	AMSR	0	0	15	0	0	15
	Reciprocal of Square root	AMRSR	15	15	7	10	10	4
Harmonic Mean	Original	HMO	0	0	15	0	0	15
	Reciprocal	HMR	15	15	0	14	14	6
	Square root	HMSR	0	0	15	0	0	14
	Reciprocal of Square root	HMRSR	15	15	11	14	14	7
Geometric Mean	Original	GMO	0	0	15	0	0	15
	Reciprocal	GMR	15	15	7	11	11	6
	Square root	GMSR	0	0	15	0	0	15
	Reciprocal of Square root	GMRSR	15	15	12	13	12	8
Median	Original	MO	0	0	15	0	0	15
	Reciprocal	MR	2	3	9	7	7	6
	Square root	MSR	0	0	15	0	0	15
	Reciprocal of Square root	MRSR	0	0	13	7	7	8

Conclusion

In this study, some improved ridge parameters are classified into different forms and various types. The performances of these estimators are evaluated through Monte-Carlo Simulation where levels of multicollinearity, sample sizes and error variances have been varied.The number of times the MSE of the improved ridge parameter is less than the existing ones is counted over the levels of multicollinearity (5) and error variance (3). Also, a maximum of fifteen (15) counts is expected. Having counted over all the levels of multicollinearity,mple sizes and error variances, the best five with the highest counts is selected. The improved ridge parameters proposed in this study are better than the existing ones/

References

1. Gujarati, D.N.(1995). Basic Econometrics, McGraw-Hill, New York.
2. Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: biased estimation for non-orthogonal problems. Technometrics, 12, 55-67.
3. McDonald, G. C. and Galarneau, D. I. (1975). A Monte Carlo evaluation of some ridge-type estimators. Journal of the American Statistical Association, 70, 407-416.
4. Lawless, J. F. and Wang, P. (1976). A simulation study of ridge and other regression estimators. Communications in StatisticsA, 5, 307-323.
5. Hocking, R., Speed, F. M. and Lynn, M. J. (1976). A class of biased estimators in linear regression. Technometrics 18 (4), 425-437.
6. Wichern, D. and Churchill, G. (1978). A Comparison of Ridge Estimators. Technometrics, 20,301–311.
7. Gibbons, D. G. (1981). A simulation study of some ridge estimators. Journal of the American Statistical Association, 76, 131-139.
8. Nordberg, L. (1982). A procedure for determination of a good ridge parameter in linear regression. Communications in Statistics A11:285–309.
9. Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. Communications in Statistics-Simulation and Computation, 32, 419-435.
10. Khalaf, G. and Shukur, G. (2005). Choosing ridge parameters for regression problems. Communications in Statistics- Theory and Methods, 34, 1177-1182.
11. Alkhamisi, M., Khalaf, G. and Shukur, G. (2006). Some modifications for choosing ridge parameters.Communications in Statistics- Theory and Methods, 35(11), 2005-2020.
12. Muniz, G. and Kibria, B. M. G. (2009). On some ridge regression estimators: An empirical comparison. Communications in Statistics-Simulation and Computation*,* 38, 621-630.
13. Mansson, K., Shukur, G. and Kibria, B. M. G. (2010). A simulation study of some ridge regression estimators under different distributional assumptions.Communications in Statistics-Simulations and Computations, 39(8), 1639 –1670.
14. Dorugade, A. V. (2014). On comparison of some ridge parameters in Ridge Regression. Sri Lankan Journal of Applied Statistics, 15(1), 31-46.
15. Lukman, A. F. and Ayinde, K. (2015). Review and classification of the Ridge Parameter Estimation Techniques. Hacettepe Journal of Mathematics and Statistics. Accepted for Publication.
16. Alkhamisi, M. and Shukur, G. (2007). A Monte Carlo study of recent ridge parameters. Commun. Statist. Simulation and Computation, 36(3), 535-547.
17. Dorugade, A. V. and Kashid, D. N. (2010). Alternative method for choosing ridge parameter for regression. International Journal of Applied MathematicalSciences, *4*(9), 447-456.

    Hoerl, A. E., Kennard, R. W. and Baldwin, K. F. (1975). Ridge regression: Some simulation. Communications in Statistics, 4 (2), 105–123.