International Journal of Statistics and Applications 2016, 6(2): 35-39
DOI: 10.5923/j.statistics.20160602.01

Estimation of Linear Regression Model with Correlated Regressors in the Presence of Autocorrelation

Ismail B.[1], Manjula Suvarna[2,*]

[1]Department of Statistics, Mangalore University, Mangalagangothri, Mangalore, India
[2]Department of Community Medicine, A.J. Institute of Medical Sciences and Research Centre, Mangalore, India

Abstract When using the linear statistical model, researchers face variety of problems due to non experimental nature i.e uncertainity about the nature of the error process, model mis- specifications, dependent regressors etc. The phenomenon of correlated errors in linear regression models involving time series data is called autocorrelation. Violation of the assumption of independent regressors leads to multicollinearity. Hence, Ordinary ridge estimates are imprecise to be of much use in case of autocorrelated regression model with the multicollinearity problem. **Objective**: To develop a new estimator for the regression parameter in the presence of multicollinearity and autocorrelation. To choose an appropriate ridge parameter for the proposed estimator using Monte Carlo simulation. **Materials and Methods**: Monte Carlo simulation study is carried out using the Statistical programming language MATLAB version 7.0 to evaluate the performance of the proposed estimator based on the Mean squared error (MSE) criterion. **Findings**: Determined the regions where a particular method for estimating ridge parameter performs better among different existing methods. This estimate of ridge parameter is used in the proposed estimator. The proposed estimator performs better than the existing estimator under the MSE criterion.

Keywords Correlated regressors, Autocorrelation, Mean Squared error, Monte Carlo Simulation

1. Introduction

In a linear regression model there are situations where the regressors may be correlated and the error terms may be autocorrelated. This phenomenon is known as autocorrelated model with multi collinearity. It is well known that when there is multicollinearity, the ordinary least square (OLS) estimator for regression coefficients or the predictor based on these estimates may give poor results [1]. For overcoming the problem of multicollinearity several methods are available such as Principal component regression, Ridge regression and Partial least squares. These methods are useful when the errors are non autocorrelated. But in the presence of autocorrelated model with multicollinearity, appropriate modifications needs to be incorporated in the estimation. Accordingly a new estimator called generalized ridge estimator is proposed and its performance is compared with Ordinary ridge estimator. Ridge estimator involves unknown ridge parameter. In the literature several methods have been discussed for the choice of ridge parameter. Simulation study has been carried out to find the appropriate method for estimating the ridge parameter which gives minimum MSE for the proposed ridge estimator.

* Corresponding author:
manjula_anil2006@yahoo.co.in (Manjula Suvarna)
Published online at http://journal.sapub.org/statistics

2. Materials and Methods

Consider the multiple linear regression model

$$Y=X\beta+\varepsilon \quad (1)$$

where Y is an observational vector of dimension n x 1, X is a n x p data matrix of regressors, β is a p x 1 vector of regression coefficient and ε is a n x 1 disturbance vector. Under the assumption that X is full rank, the errors are non autocorrelated and X and ε are independently distributed, the Ordinary Least Square (OLS) estimator of β is

$\hat{\beta}=(X'X)^{-1}X'Y$ with the covariance matrix of $\hat{\beta}$ is obtained as $\text{cov}(\hat{\beta})=\sigma^2(X'X)^{-1}$.

It is not necessary that assumptions mentioned above hold good in real life situation. The regressors may be nearly correlated and the responses may also be correlated. In such instances the OLS estimator mentioned above do not possess the optimum statistical property. Hence there is a need to develop a new estimator which takes care of this situation.

Ridge Regression: The violation of the assumption of independent regressors leads to multicollinearity. If X is less than full rank then such a situation is known as perfect multicollinearity. In this case OLS estimator does not exist. This situation is very rare in practice. In most of the real life situations, some regressors are nearly related to the remainining regressors. This is known as near

multicollinearity. In case of near multicollinearity, rank of the regressor matrix X is equal to k and hence OLS estimator exist, but they are too imprecise to be of much use [2]. With strongly interrelated pairs of regressors, X'X is illconditioned and the variance of the OLS estimator becomes large. With multicollinearity, the estimated OLS coefficients may be statistically insignificant (too large, too small and even have wrong sign). Hence interpretation given to the regression coefficients may no longer be valid. It may be preferable to consider biased estimators of β, if their variances are sufficiently smaller than those of OLS estimators. One such biased estimator is the "ridge estimator". The ridge estimator (ordinary ridge estimator) of β is

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

where the constant k >0 is known as "ridge" parameter . As the constant k increases from zero and continues up to infinity, the regression estimates tend towards zero. Though these estimators result in biased estimates, for certain positive values of k, this estimator yields minimum mean squared error (MMSE) compared to OLS. Several methods for estimating k has been proposed by Hoerl and Kennard [3], Hoerl et al., [4], Mc Donald and Galarneau [5], Hocking et al., [6], Saleh and Kibria [7], Khalf and Shukur [8]. From example Hoerl and Kennard (1970), the value of k that minimises the MSE is $\hat{k} = \frac{\hat{\sigma}^2}{\hat{\alpha}^2_{\max}}$, where $\hat{\sigma}^2$ represents the error variance of model (1), $\hat{\alpha}_{\max}$ is the maximum among elements of α defined as $\alpha = D'\beta$ with D being an orthogonal matrix.

Autocorrelation: Autocorrelation is said to exist when the successive observations in linear regression model are correlated. The existence of autocorrelated errors has been rationalized in a variety of ways, as noted by Maddala [9].

The successive dependence of the error term is represented by

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t \qquad t=2,3,\ldots n \tag{3}$$

u_t are independent and identically distributed random variable with mean zero and variance $\sigma^2{}_u$ [10]. When the error satisfies the relation (3), the observations follow first order autocorrelation. The variance covariance matrix of Y is

$D(Y) = D(\varepsilon) = \sigma^2\Omega \neq \sigma^2 I$

Where $\sigma^2 = \sigma^2{}_u \frac{1}{1-\rho^2}$ and

$$\Omega = \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \rho \\ \rho^{n-1} & \rho^{n-2} & \cdots\rho & 1 \end{bmatrix}$$

Since the covariance matrix of ε is nonspherical (i.e not a scalar multiple of the identity matrix), OLS, though unbiased, is inefficient relative to generalised least squares by Aitken's theorem. The generalized least squares estimator of β in (1) is [10]

$$\hat{\beta}_{GLS} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y \tag{4}$$

If the parameter ρ in (3) is known then we can write

$$\Omega^{-1} = \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \tag{5}$$

3. Generalised Ridge Type Estimator

Consider a general linear regressive model (1) with errors satisfying relation (3) and the regressors exhibiting near multicollinearity. As seen earlier, in case of autocorrelation $D(\varepsilon) = \sigma^2\Omega \neq \sigma^2 I$. Hence autocorrelation is a particular case of heteroscedasticity. In the case of heteroscedasticity, GLS is an appropriate method of estimation as given in (4). Further, when there is multicollinearity, often used method is the ridge regression as mentioned in (2). Combining these two methods, we propose for the autocorrelated model with multicollinearity a generalized ridge type estimator represented as $\hat{\beta}_{GR} = (X'\,\Omega^{-1}\,X + kI)^{-1}\,X'\,\Omega^{-1}\,Y$ where Ω^{-1} is as defined in (5).

Hence the model under consideration contains the unknown parameters k, ρ, σ^2 and β.

In the following [11] we present some existing methods for estimating ridge parameter k

1. Hoerl and Kennard method

$$\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\max(\hat{\alpha}_i^2)} \tag{6}$$

2. Hoerl, Kennard and Baldwin method

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\sum \hat{\alpha}_i^2} \tag{7}$$

3. Hocking, Speed and Lynn method

$$\hat{k}_{HO} = \hat{\sigma}^2 \frac{\sum_{i=1}^{p}(\lambda_i \hat{\alpha}_i)^2}{(\sum_{i=1}^{p}\lambda_i \hat{\alpha}_i^2)^2} \tag{8}$$

For the proposed estimator, λ_i's are the eigen values of ($X'\Omega^{-1}X$).

4. The Monte Carlo Simulation Study

A simulation study is carried out to find out the appropriate estimate for the ridge parameter among (6), (7), (8) mentioned above which gives minimum MSE for the proposed estimator.

The data is simulated in accordance with the multiple linear regression model given in (1) with the number of regressors p=3 and ε satisfying first order auto regressive scheme as mentioned in (3). The dependent variable Y is generated using the relation Y=Xβ +ε with $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ X=[X_0 , X_1 , X_2 , X_3], β=[β_0 β_1 β_2 β_3]' =[4 2.5 1.8 0.6]', X_0 =[1 1 1….1]'.

To generate normally distributed random variables X_1, X_2, X_3 with specified intercorrelations we use the following equations [12, 13].

$$X_1 = \mu_1 + \sigma_1 Z_1$$

$$X_2 = \mu_2 + \lambda_{12}\sigma_2 Z_1 + \sqrt{m_{22}}Z_2$$

$$X_3 = \mu_3 + \lambda_{13}\sigma_3 Z_1 + \frac{m_{23}}{\sqrt{m_{22}}}Z_2 + \sqrt{n_{23}}Z_3$$

where

$$m_{22} = \sigma^2{}_2[1-\lambda_{12}{}^2], m_{23} = \sigma_2\sigma_3[\lambda_{23} - \lambda_{12}\lambda_{13}],$$

$$m_{33} = \sigma^2[1-\lambda_{13}{}^2]$$

$$n_{33} = m_{33} - \frac{m_{23}{}^2}{m_{22}} \text{ and } Z_i \sim N(0,1) \text{ for i=1, 2,3}$$

λ_{12} =corr(X_1 ,X_2) λ_{13} =corr(X_1, X_3) λ_{23} =corr(X_2, X_3)

In equation (3), u_t are independent and identically distributed normal random variables with mean 0 and variance σ_u^2. The autocorrelation coefficient ρ in (3) is ranging from -0.9 to -0.1 and the regression parameters are fixed as $\beta_0 = 4, \beta_1 = 2.5, \beta_2 = 1.8, \beta_3 = 0.6$.

The parameters of the model in equation (4) are fixed as $\beta_0 = 4, \beta_1 = 2.5$, $\beta_2 = 1.8$ and $\beta_3 = 0.6$. Taking $\lambda(X_{12}) = \lambda(X_{13}) = \lambda(X_{23}) = \lambda$ sixteen different levels of intercorrelation (multicollinearity) among the regressors are taken as -0.2, -0.3, -0.4, -0.5, -0.6, -0.7, -0.8, -0.9, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9.

With the above setup a sample of 100 observations are generated and replicated 1000 times. For each choice of the k, the MSE for the generalized ridge estimator is computed. The estimator of the ridge parameter (k) which gives minimum MSE is recorded for different choice of the parameters β_0 , β_1 , β_2 , β_3 , ρ , λ and the results are presented in Table 1.

5. Results and Discussions of the Simulation Study

The first column of Table 1 contains 9 levels of autocorrelation and the first row represents different levels of intercorrelation between the regressors. The other elements in Table 1 represents the choice of the ridge parameter which gives minimum MSE for the proposed generalized ridge estimator. For example, ρ = -0.3 and λ =+0.2, Hoerl and Kennard (HK) estimator of the ridge parameter gives minimum MSE. Similarly when ρ = -0.7 and λ = -0.2, Hoerl Kennard and Baldwin (B) estimator of the ridge parameter gives minimum MSE.

From the results in Table 1, it is clear that when ρ and λ are negative and very high, the estimator for k proposed by Baldwin et al possesses minimum MSE. For the same values of λ , when autocorrelation is low, Hoerl and Kennard (HK) estimator is superior to Hoerl Kennard and Baldwin(B)estimator. Also as multicollinearity is positive and increases with the autocorrelation being low, then it is observed that Hoerl and Kennard (HK) estimator performs better than the other estimators.

When the intercorrelation among regressors is high and autocorrelation is also high, the ridge parameter proposed by Hoerl Kennard and Baldwin (B) is superior compared to the other estimates. Hence using (B) estimator, the proposed generalized ridge estimator (GR) is compared with Ordinary ridge estimator (RR) through MSE. Table 2 gives the MSE of GR estimator and RR estimator for different choice of λ and ρ .

The results in Table 2, depicts that developed Generalised ridge estimator (GR) has minimum MSE compared to Ordinary ridge (RR) estimator. Therefore in the presence of autocorrelation with multicollinearity the proposed ridge estimator is superior to ordinary ridge estimator. Hence use of ordinary ridge estimator leads to larger MSE if autocorrelation is ignored.

6. Conclusions

There are a number of articles where multicollinearity and autocorrelation are dealt separately. However limited studies are available which describes these two problems together. Hence in this article, an attempt has been made to address these 2 issues. It is observed through simulation that the use of ordinary ridge estimator leads to larger MSE if autocorrelation is ignored.

Therefore while conducting research in the field of Social Sciences or Epidemiological studies, there is a critical need to check data for the existence of multicollinearity between the regressors as well as the presence of autocorrelation. This

will avoid misinterpretation of the results and will also ensure that the emerging problems involving the inter relationships between a number of variables of interest may be addressed appropriately and effectively.

Table 1. Performance of the Proposed Estimator when p=3 for Different Choice of Ridge parameter, Different Levels of Correlation Between the Regressors (λ) and the Autocorrelation (ρ) Ranging from -0.9 to -0.1

ρ	λ															
	-0.9	**-0.8**	**-0.7**	**-0.6**	**-0.5**	**-0.4**	**-0.3**	**-0.2**	**+0.2**	**+0.3**	**+0.4**	**+0.5**	**+0.6**	**+0.7**	**+0.8**	**+0.9**
-0.9	B	B	B	B	HO	HO	HO	B	B	HO	HO	HO	HO	HO	HO	HO
-0.8	B	B	B	B	HO	B	B	B	B	B	B	B	B	B	B	HO
-0.7	B	B	B	B	HO	B	B	B	B	B	HK	B	B	B	B	B
-0.6	B	B	B	B	HK	B	B	HK	HK	HK	HK	HK	HK	B	B	B
-0.5	B	B	B	B	HK	HK	HK	HK	HK	B	HK	HK	HK	HK	B	B
-0.4	B	HK	HK	B	B	B	HK	HK	HK	HK	HK	HK	HK	HK	B	HK
-0.3	B	HK	HK	HK	B	B	B	HK	HK	HK	HK	HK	HK	HK	B	HK
-0.2	B	HK	HK	HK	B	HK	HK	B	HK	HK	HK	HK	HK	HK	B	HK
-0.1	B	HK	HK	HK	B	HK	HK	HK	HK	HK	HK	HK	HK	HK	B	HK

B= Hoerl, Kennard and Baldwin method; HK= Hoerl and Kennard method; HO= Hocking, Speed and Lynn method

Table 2. Comparison of the Proposed Generalised Ridge Estimator (GR) with Ordinary Ridge Estimator (RR) for Different Levels of Autocorrelation and Inter Correlations between the Regressors

	λ = -0.9		**λ = -0.8**		**λ = -0.7**		**λ = -0.6**	
	GR	RR	GR	RR	GR	RR	GR	RR
ρ =-0.9	20.8214	28.4813	14.5029	27.6239	15.3571	28.4237	20.3553	31.3304
ρ =-0.8	11.1806	24.1588	7.3473	23.8725	7.3130	24.7635	10.7937	27.1429
ρ =-0.7	7.0867	22.997	4.8752	23.0145	4.9123	23.8989	7.2610	25.5912
ρ =-0.6	5.3110	22.0598	3.8412	22.2421	3.8393	23.1996	5.3256	24.6439
ρ =-0.5	4.4693	21.6943	3.2457	22.537	3.1545	23.045	4.6707	24.264
ρ =-0.4	3.9414	21.4783	2.8319	21.8993	2.7390	22.8049	3.9805	24.0389
ρ =-0.3	3.6534	21.2125	2.5023	21.7106	2.5573	22.8045	3.6053	23.8935
ρ =-0.2	3.3871	21.0056	2.4467	21.9826	2.3658	22.631	3.3031	23.7467
ρ =-0.1	3.2545	21.0418	2.2946	21.686	2.2259	22.5208	3.1772	23.8829

	λ = -0.5		**λ = -0.4**		**λ = -0.3**		**λ = -0.2**	
	GR	RR	GR	RR	GR	RR	GR	RR
ρ =-0.9	199.8822	49.1531	20.0921	32.7013	11.6972	28.5534	7.7164	26.1152
ρ =-0.8	72.0298	41.4491	10.0910	27.9772	5.8495	25.5883	4.1563	24.282
ρ =-0.7	36.1971	38.7429	6.9372	26.7755	3.9133	24.4849	2.7901	23.3956
ρ =-0.6	22.6866	37.0179	5.0300	26.1235	3.2329	24.0875	2.2464	23.2887
ρ =-0.5	19.7516	36.8716	4.2545	25.6171	2.5722	23.9203	1.8807	23.124
ρ =-0.4	15.4482	35.9579	3.7252	25.3408	2.2708	23.7254	1.6807	22.8259
ρ =-0.3	14.1846	35.6086	3.3227	25.2213	2.0405	23.5864	1.5831	22.9292
ρ =-0.2	13.7764	35.7806	3.0278	25.2916	1.9169	23.598	1.3755	22.7732
ρ =-0.1	13.1096	35.3365	3.000	24.9384	1.8532	23.5199	1.3628	22.6183

	λ = +0.9		λ = +0.8		λ = +0.7		λ = +0.6	
	GR	RR	GR	RR	GR	RR	GR	RR
ρ =-0.9	17.0649	10.2150	9.2187	6.6896	7.0452	6.0621	6.0527	5.9508
ρ =-0.8	9.1292	6.4531	5.1840	4.4913	3.7368	3.9850	3.0471	4.1352
ρ =-0.7	6.1900	5.5187	3.5447	3.7676	2.6166	3.5976	2.2589	3.7465
ρ =-0.6	4.7260	4.5195	2.7440	3.4583	1.9908	3.2707	1.8175	3.5057
ρ =-0.5	3.8366	4.0887	2.2362	3.1312	1.6981	3.0398	1.3938	3.0303
ρ =-0.4	3.3728	3.9213	2.0145	2.965	1.4777	2.9085	1.2906	3.1830
ρ =-0.3	3.0877	3.8317	1.7626	2.8131	1.3390	2.8399	1.1633	3.0306
ρ =-0.2	2.7796	3.7462	1.5580	2.8310	1.2766	2.9517	1.1229	3.0700
ρ =-0.1	2.6801	3.5996	1.5684	2.8537	1.2086	2.8849	0.9973	3.0282

	λ = +0.5		λ = +0.4		λ = +0.3		λ = +0.2	
	GR	RR	GR	RR	GR	RR	GR	RR
ρ =-0.9	5.3924	5.8605	5.3058	5.8578	5.0683	6.3270	5.1080	7.6828
ρ =-0.8	2.7493	4.3247	2.6378	4.4834	2.5090	5.1597	2.6232	6.1052
ρ =-0.7	1.9740	3.6630	1.7006	3.9858	1.7465	4.5226	1.8103	5.7873
ρ =-0.6	1.5584	3.6442	1.4348	3.9782	1.4740	4.5312	1.4172	5.6275
ρ =-0.5	1.3255	3.5115	1.1688	3.9152	1.1348	4.1688	1.1615	5.4054
ρ =-0.4	1.2173	3.3381	1.0653	3.7828	1.0802	4.2253	1.0654	5.3570
ρ =-0.3	1.0328	3.3645	0.9576	3.4897	0.9507	4.1495	0.9479	5.2651
ρ =-0.2	1.0310	3.3315	0.9106	3.6045	0.9078	4.0325	0.9053	5.2602
ρ =-0.1	0.9036	3.1394	0.8536	3.6475	0.8484	4.0621	0.8659	5.4263

ACKNOWLEDGEMENTS

The authors are grateful to the reviewers for their valuable suggestions.

REFERENCES

[1] Gunst, R..F and Manson, R. L, 1979, Some considerations in the evaluation of alternate prediction equations, Technometrics, 21:55-63.

[2] Peter Schmidt, 1976,. Econometrics, New York:Marcel dekker Inc.

[3] Hoerl, A. E. and Kennard, R.W, 1970, Ridge regression: biased estimation for nonorthogonal problems, Technometrics, 12:9-82.

[4] Hoerl, A.E., Kennard, R. W. and Baldwin, K. F., 1975, Ridge regression: Some simulation. Communications in Statistics, 4:105-123.

[5] Mc Donald, G.C and Galarneau, D. I ,1975, A Monte carlo evaluation of some ridge type estimators. Journal of American statistical Association ,70(350):407-412.

[6] Hocking, R.R., F.M. Speed and M.J. Lynn, 1976, A class of biased estimators in linear regression. Technometrics, 18:425-437.

[7] Saleh, A.K. and Kibria, B. M., 1993, Performances of some new preliminary test ridge regression estimators and their properties. Communications in Statistics- Theory and Methods, 22:2747-2764.

[8] Khalaf, G. and Shukur, G., 2005, Choosing ridge parameter for regression problems. Communications in Statistics-Theory and Methods, 34:1177-1182.

[9] Maddala, G, 1977, Econometrics, New York: McGraw-Hill.

[10] Thomas B. Fomby, R. Carter Hill, Stanley R. Johnson, 1984, Advanced Econometric Methods, New York:Springer Verlag.

[11] Al-Hassan, Yazid M., 2010, Performance of a new Ridge Regression Estimator. Journal of Association of Arab Universities for Basic and Applied Sciences, 9:43-50.

[12] K. Ayinde and O. S. Adegboye, 2010, Equations for generating normally distributed random variables with specified intercorrelation. Journal of Mathematical Sciences, 21:83-203.

[13] Kayode Ayinde, Emmanuel O. Apata, Oluwayemisi O. Alaba, 2012, Estimators of Linear Regression model and prediction under some assumptions violation. Journal of Statistics, 2:534-546.