DOI: 10.20472/IAC.2018.035.034

FELA ÖZBEY

Çukurova University, Turkey

EVALUATION ESTIMATION PERFORMANCES OF LIU TYPE AND TWO-PARAMETER RIDGE ESTIMATORS USING MONTE CARLO EXPERIMENTS

Abstract:

Multiple linear regression model is a widely used statistical technique in social and life sciences. Ordinary Least Squares (OLS) estimator is the Best Linear Unbiased Estimator (BLUE) for the unknown population parameters of this model. Unfortunately, sometimes two or more of the regressors may be moderately or highly correlated causing multicollinearity problem. Various biased estimators are proposed to refine the ill-conditioning of X’X matrix and shrink the variance under the multicollinearity. The most popular of them is Ridge estimator. But it may worsen the fit when solving the ill-conditioning problem. Two-Parameter Ridge (2PR) and Liu Type (LT) estimators are proposed to overcome the fitting degeneration of Ridge estimator by using a tuning parameter. In this study, holding the parameter refining the ill-conditioning of X’X matrix fixed, the success of the tuning parameters of these estimators is investigated. Minimizers of Predicted Sum of Squares (PRESS) and Generalized Cross Validation (GCV) statistics are used as estimates of tuning parameters. Optimum parameter estimates are compared via their Scalar Mean Squared Errors (SMSE). It is observed that the SMSEs of estimates obtained by LT and 2PR estimators decreases when estimates of parameter refining the ill-conditioning of X’X matrix increases, and in all cases estimates obtained by the 2PR estimator are much more efficient than estimates obtained by LT and OLS estimators.

Keywords:

Biased Estimators, Estimation, Monte Carlo Simulations, Multicollinearity.

JEL Classification: C13, C52, C63

1 Introduction

Multiple linear regression model is a widely used statistical technique in social and life sciences. OLS estimator is the BLUE for the unknown population parameters of this model. Unfortunately, in the presence of multicollinearity, ill-conditioning of X'X matrix enlarges the variance of parameters estimates of OLS and makes them unstable and unreliable. Various biased estimators are proposed to overcome this problem. The most popular is the Ridge estimator proposed by Hoerl and Kennard (1970). But this estimator sometimes worsens the fit when solving the ill-conditioning problem. LT (Liu, 2003) and 2PR (Lipovetsky & Conklin, 2005) estimators are proposed to overcome the fitting degeneration of Ridge estimator by using a tuning parameter. Liu (2003) and Lipovetsky (2006) suggest choosing biasing parameter *k* so that to reduce the ill-conditioning problem to the desired level, and then to evaluate the tuning parameter by optimizing a criterion of optimum fit.

In this study, following suggestions mentioned above, holding the parameter refining the ill-conditioning of X'X matrix fixed, the success of the tuning parameters of these estimators is investigated. Most of the studies use the minimizer of Mean Squared Error (MSE) as the biasing parameter, replacing the unknown population parameters by their unbiased estimates. But in the presence of the multicollinearity, these estimates are very unstable and may be far from the population parameters. To avoid this inconvenience, biasing parameter which solves the ill-conditioning of X'X matrix is estimated as to reduce the conditional index (CI), and minimizers of PRESS and GCV statistics are used as estimates of tuning parameters.

In the literature, there are numerous theoretical studies comparing estimators under various criteria. Unfortunately, in theoretical studies comparisons are for any values of biasing parameters, not for optimal ones. For example, see Farebrother (1976), Draper and Van Nostrand (1979), Toker and Kaçıranlar (2013). In order to compare the best estimates (under given criteria), Monte Carlo experiments may be preferred to be used. A simulation study comparing prediction performances of LT and 2PR is performed in Özbey (2013) and some real data analyses are performed in Özbey (2012) selecting *k* such that to reduce CI to 10, as it suggested and applied in Liu (2003). It is established that predictions obtained using 2PR estimator are better than LT estimator.

In this study, in order to compare the best estimates (under given criteria), Monte Carlo experiments are preferred to be used.

2 Estimators and Statistics Used

2.1. Estimators

Ridge estimator proposed as an alternative to OLS in the presence of multicollinearity is defined as:

$$\hat{\beta}_R = (X'X + kI)^{-1} X'y \quad k > 0. \tag{1}$$

Here *k* is biasing parameter used to refine the ill-conditioning of X'X matrix.

LT estimator, which is proposed to improve the performance of the Ridge estimator, is defined as:

$$\begin{aligned}\hat{\beta}_{LT} &= (X'X + kI)^{-1}(X'y - d\hat{\beta}_{OLS}) \\ &= (X'X + kI)^{-1}(X'X - dI)\hat{\beta}_{OLS} \\ k &> 0.\end{aligned} \tag{2}$$

Here *d* is the tuning parameter.

2PR estimator, which is another estimator proposed to improve the performance of the Ridge estimator, is defined as:

$$\hat{\beta}_{2PR} = q(X'X + kI)^{-1} X'y \quad k > 0. \tag{3}$$

Here *q* is the tuning parameter.

2.2. Statistics

CI of (X'X+*k*I) matrix is defined as:

$$CI = \sqrt{\frac{\lambda_{\max} + k}{\lambda_{\min} + k}}. \tag{4}$$

Here $\lambda_{\max}$ and $\lambda_{\min}$ are the largest eigenvalue and the lowest eigenvalue of X'X matrix, correspondingly.

PRESS statistics proposed by Allen (1971) is defined as:

$$PRESS = \sum_{i=1}^{n} \left(\frac{\hat{e}_i}{1 - h_{ii}} \right)^2. \tag{5}$$

ISBN ISBN 978-80-87927-45-8, IISES

By definition, PRESS statistics of LT estimator is:

$$PRESS_{LT} = \sum_{i=1}^{n}\left(\frac{y_i - x_i'(X'X+kI)^{-1}(X'X-dI)(X'X)^{-1}X'y}{1-x_i'(X'X+kI)^{-1}(X'X-dI)(X'X)^{-1}x_i}\right)^2, \tag{6}$$

and PRESS statistics of the 2PR estimator is:

$$PRESS_{2PR} = \sum_{i=1}^{n}\left(\frac{y_i - qx_i'(X'X+kI)^{-1}X'y}{1-qx_i'(X'X+kI)^{-1}x_i}\right)^2. \tag{7}$$

GCV statistics proposed in Golub et al. (1979) is defined as:

$$GCV = \frac{n\|(I-H)y\|^2}{(Tr(I-H))^2} \tag{8}$$

which turns to:

$$GCV_{LT} = \frac{n\sum_{i=1}^{n}(y - X(X'X+kI)^{-1}(X'X-dI)(X'X)^{-1}X'y)^2}{(\sum_{i=1}^{n}(1-x_i(X'X+kI)^{-1}(X'X-dI)(X'X)^{-1}x_i'))^2} \tag{9}$$

and

$$GCV_{2PR} = \frac{n\sum_{i=1}^{n}(y - qX(X'X+kI)^{-1}X'y)^2}{(\sum_{i=1}^{n}(1-qx_i(X'X+kI)^{-1}x_i'))^2} \tag{10}$$

for LT and 2PR estimators correspondingly.

3 Monte Carlo Simulations

3.1. Simulations Setup

In the experiment, regressors are simulated as in McDonald and Galarneau (1975), Kibria (2003), Liu (2004) and Güler and Kaçıranlar (2009):

$$x_{ij} = (1-\rho)^{1/2} z_{ij} + \sqrt{\rho} z_{im+1}, \quad i = 1,2,\ldots,n, \quad j = 1,2,\ldots,m \tag{11}$$

z_{ij}s are independent standard normal pseudo-random numbers. In this way, the regressors are simulated as to be collinear. The square root of rho is selected as 0.9999 to assure *k* to be positive. The dependent variable is simulated as:

$$y = X\beta + \varepsilon. \tag{12}$$

β is the normalized eigenvector corresponding to the largest eigenvalue of the X'X matrix, see Newhouse and Oman (1971). $\varepsilon \sim N(0,1)$ is a vector consists of pseudo-random numbers. Seed is chosen as 45324762. The number of regressors simulated is 3, and the number of observations is 50. The vector of parameters to be estimated is $\beta' = \begin{bmatrix} 0.5761 & 0.5774 & 0.5786 \end{bmatrix}$.

SMSE of each estimator is evaluated as:

$$SMSE(\hat{\beta}) = \frac{1}{MCR} \sum_{i=1}^{MCR} (\hat{\beta}_i - \beta)'(\hat{\beta}_i - \beta), \tag{13}$$

where $\hat{\beta}_i$ is the estimated value of β at the *i*th replication. MCR is the number of Monte Carlo replications.

Three different values of *k* are evaluated as to reduce the CI to 10, 5, and 3. Replicating simulations and estimations 100, 500, 1000, and 10,000 times, and by using minimizers of both PRESS and GCV statistics as estimates of tuning parameters, the following results are obtained.

3.1 Simulations Results

Table 1: Results for k=1.4705 by minimizing PRESS and GCV statistics.

	0.5761	***0.5774***	***0.5786***	***PRESS***	***SMSE($\hat{\beta}$)***	***0.5761***	***0.5774***	***0.5786***	***GCV***	***SMSE($\hat{\beta}$)***
MCR=100										
OLS	1.6264	-1.5082	-0.1046	5251	20404	1.6264	-1.5082	-0.1046	4905.6	20404
LT	1.0707	-0.9902	-0.0682	***4592.1***	81700	0.9798	-0.9055	-0.0622	90.2993	83507
2PR	-0.0246	0.0098	-0.0018	4630.7	***395.00***	-0.0160	0.0064	-0.0012	***90.1832***	***387.43***
MCR=500										
OLS	-0.0597	-0.0708	0.1314	5230.1	19692	-0.0597	-0.0708	0.1314	4900.2	19692
LT	0.1067	0.1268	-0.2315	***4587.2***	75910	0.1360	0.1616	-0.2954	90.7641	73486
2PR	0.0002	-0.0001	-0.0044	4635.1	***431.47***	0.0003	-0.0002	-0.0052	***90.2657***	***449.7799***
MCR=1,000										
OLS	0.0171	0.0850	-0.1018	5226.9	20623	0.0171	0.0850	-0.1018	4894.8	20623
LT	-0.0006	-0.0038	0.0046	***4618.3***	72367	-0.0038	-0.0196	0.0234	91.5840	67285

ISBN ISBN 978-80-87927-45-8, IISES

2PR	-0.0215	-0.0506	0.0606	4652.8	***486.30***	-0.0215	-0.0506	0.0606	***90.8153***	***526.1356***
MCR=10,000										
OLS	-0.0035	0.0085	-0.0047	5228.7	20986	-0.0035	0.0085	-0.0047	104.1356	20986
LT	0.0039	-0.0088	0.0052	***4637.2***	70622	0.0045	-0.0101	0.0059	91.8418	67503
2PR	0.00008	0.00015	0.00008	4675.4	***440.89***	0.00008	0.00014	0.00008	***91.4781***	***471.3824***

Table 2: Results for k=6.092 by minimizing PRESS and GCV statistics.

	0.5761	***0.5774***	***0.5786***	***PRESS***	**SMSE($\hat{\beta}$)**	***0.5761***	***0.5774***	***0.5786***	***GCV***	**SMSE($\hat{\beta}$)**
MCR=100										
OLS	1.6264	-1.5082	-0.1046	5251	20404	1.6264	-1.5082	-0.1046	4905.6	20404
LT	1.0623	-0.9828	-0.0673	***4609.1***	50901	0.9700	-0.8970	-0.0612	90.7430	49002
2PR	-0.0105	-0.0017	-0.0047	4633.1	***28.041***	-0.0068	-0.0011	-0.0031	***90.1902***	***26.8209***
MCR=500										
OLS	-0.0597	-0.0708	0.1314	5230.1	19692	-0.0597	-0.0708	0.1314	4900.2	19692
LT	0.1097	0.1300	-0.2379	***4604.8***	47336	0.1389	0.1647	-0.3016	91.1932	44566
2PR	-0.0010	-0.0011	-0.0022	4635.9	***29.654***	-0.0012	-0.0013	-0.0026	***90.2695***	***30.4148***
MCR=1,000										
OLS	0.0171	0.0850	-0.1018	5226.9	20623	0.0171	0.0850	-0.1018	4894.8	20623
LT	-0.0013	-0.0069	0.0083	***4634.6***	43540	-0.0042	-0.0211	0.0253	91.9489	40684
2PR	-0.0079	-0.0150	0.0120	4652	***35.176***	-0.0079	-0.0150	0.0120	***90.8109***	***35.1206***
MCR=10,000										
OLS	-0.035	0.085	-0.047	5228.7	20986	-0.035	0.085	-0.047	104.136	20986
LT	0.0039	-0.0087	0.0051	***4653.2***	41305	0.0045	-0.0102	0.0060	92.2008	39566
2PR	0.0001	0.0001	0.0001	4674.4	***31.7175***	0.0001	0.0001	0.0001	***91.4703***	***31.8284***

Results given in Table 1 are obtained by evaluating *k* as to reduce the CI to 10 (i.e., *k*=1.4705). Under these conditions, compared to OLS estimator, both LT and 2PR estimators reduce PRESS and GCV statistics but the SMSEs of LT estimator are even 3-4 times greater than SMSEs of OLS estimator. Generally, both LT and 2PR estimators perform better when tuning parameters estimates are minimizers of PRESS statistics.

In Table 2 results are given obtained by selecting *k*=6.092 (i.e., reducing CI to 5). It is observed that, though PRESS and GCV statistics do not change very much and though SMSEs of the LT estimator are still higher than the SMSEs of the OLS estimator, reducing CI from 10 to 5 contributes to the improvement of SMSEs of both LT and 2PR estimators.

Take into account the improvement of SMSEs; CI is lowered to 3 in the hope of reducing SMSEs of LT estimator below SMSEs of OLS estimator. Results are given in Table 3.

Table 3: Results for k=18.2927 by minimizing PRESS and GCV statistics.

	0.5761	***0.5774***	***0.5786***	***PRESS***	***SMSE($\hat{\beta}$)***	***0.5761***	***0.5774***	***0.5786***	***GCV***	***SMSE($\hat{\beta}$)***
MCR=100										
OLS	1.6264	-1.5082	-0.1046	5251	20404	1.6264	-1.5082	-0.1046	4905.6	20404
LT	1.0429	-0.9650	-0.0660	4670.9	23230	0.9478	-0.8765	-0.0597	92.3700	24467
2PR	-0.0074	-0.0042	-0.0053	***4636.3***	***7.1435***	-0.0048	-0.0028	-0.0035	***90.2404***	***6.4939***
MCR=500										
OLS	-0.0597	-0.0708	0.1314	5230.1	19692	-0.0597	-0.0708	0.1314	4900.2	19692
LT	0.1156	0.1370	-0.2510	4660.7	24381	0.1451	0.1719	-0.3152	92.6543	23626
2PR	-0.0013	-0.0013	-0.0017	***4638.8***	***6.8537***	-0.0015	-0.0016	-0.0020	***90.3174***	***6.7972***
MCR=1,000										
OLS	0.0171	0.0850	-0.1018	5226.9	20623	0.0171	0.0850	-0.1018	4894.8	20623
LT	-0.0028	-0.0139	0.0168	4681.9	24038	-0.0049	-0.0248	0.0298	93.1760	23064
2PR	-0.0048	-0.0071	0.0018	***4654.0***	***7.6473***	-0.0048	-0.0071	0.0018	***90.8497***	***7.4613***
MCR=10,000										
OLS	-0.035	0.085	-0.047	5228.7	20986	-0.035	0.085	-0.047	104.136	20986
LT	0.0035	-0.0078	0.0046	*4700.4*	22391	0.0042	-0.0096	0.0056	*93.4667*	21593
2PR	0.00010	0.00011	0.00010	***4676.1***	***7.2283***	0.0001	0.0001	0.0001	***91.5017***	***7.0399***

Results in Table 3 are for k=18.2927. Again, it is observed that PRESS and GCV statistics do not change very much, and SMSEs of the LT estimator are still higher than the SMSEs of the OLS estimator; but reducing CI to 3 improves SMSEs of both estimators, and also approximates SMSEs of LT estimator to SMSEs of OLS estimator.

4 Conclusion

In the presence of multicollinearity, ill-conditioning of X'X matrix enlarges the variance of parameters estimates of OLS and makes them unstable and unreliable. The most popular biased estimator proposed to overcome this problem is the Ridge estimator. But this estimator sometimes may worsen the fit when solving the ill-conditioning problem. LT and 2PR estimators are proposed to overcome the fitting degeneration of Ridge estimator by using a tuning parameter.

In this study, to compare the best estimates under given criteria, Monte Carlo experiments are performed. Following Liu (2003), biasing parameter *k* is chosen as to reduce the CI to the desired level. After that, following Liu (2003) and Lipovetsky (2006), the tuning parameters of these estimators are evaluated for fixed values of *k*. Tuning parameters estimates are selected as minimizers of PRESS and GCV statistics.

Three different values of *k* are evaluated as to reduce CI to 10, 5, and 3. Simulations and estimations are replicated 100, 500, 1000, and 10,000 times. Following results are obtained:

1) Reducing CI improves SMSEs of parameter estimates of both LT and 2PR estimators.
2) Under given experiments above LT estimator fails to provide lower SMSEs than the OLS estimator.
3) Parameter estimates obtained by LT and 2PR estimators tend to zero while the number of replications increases.
4) 2PR estimator generates estimates with lower SMSEs than LT and OLS estimators.

Acknowledgment: This research was supported by Scientific Research Projects Unit of Çukurova University Grant SBA-2016-5204.

References

Allen, D. M. (1971). *The prediction sum of squares as a criterion for selecting predictor variables*: University of Kentucky.

Draper, N. R., & Van Nostrand, R. C. (1979). Ridge regression and James-Stein estimation: review and comments. *Technometrics, 21*(4), 451-466.

Farebrother, R. (1976). Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society. Series B (Methodological), 38*(3), 248-250.

Golub, G. H., Heath, M., & Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics, 21*(2), 215-223.

Güler, H., & Kaçıranlar, S. (2009). A comparison of mixed and ridge estimators of linear models. *Communications in Statistics-Simulation and Computation®, 38*(2), 368-401.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55-67.

Kibria, B. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics-Simulation and Computation, 32*(2), 419-435.

Lipovetsky, S. (2006). Two-parameter ridge regression and its convergence to the eventual pairwise model. *Mathematical and Computer Modelling, 44*(3-4), 304-318. doi:10.1016/j.mcm.2006.01.017

Lipovetsky, S., & Conklin, W. M. (2005). Ridge regression in two-parameter solution. *Applied Stochastic Models in Business and Industry, 21*(6), 525-540.

Liu, K. (2003). Using Liu-type estimator to combat collinearity. *Communications in Statistics-Theory and Methods, 32*(5), 1009-1020.

Liu, K. (2004). More on Liu-type estimator in linear regression. *Communications in Statistics-Theory and Methods, 33*(11), 2723-2733.

McDonald, G. C., & Galarneau, D. I. (1975). A Monte Carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association, 70*(350), 407-416.

Newhouse, J. P., & Oman, S. D. (1971). An evaluation of ridge estimators.

Özbey, F. (2012). Çoklu iç ilişki olması durumunda lineer regresyon modellerinin kullanımı ile kestirim ve öngörü. (Unpublished doctoral dissertation). Çukurova University, Adana, Turkey.

Özbey, F. (2013). *Liu Tipi ve 2 Parametreli Ridge Tahmin Edicilerinin Kestirim Performanslarının Karşılaştırılması*. Abstracts book of the 14th International Symposium on Econometrics, Operations Research and Statistics (ISEOS). Sarajevo, Bosnia and Herzegovina.

Toker, S., & Kaçıranlar, S. (2013). On the performance of two parameter ridge estimator under the mean square error criterion. *Applied Mathematics and Computation, 219*(9), 4718-4728.