Thailand Statistician
January 2021; 19(1): 115-124
http://statassoc.or.th
Contributed paper

A New Ridge Estimator for the Negative Binomial Regression Model

Nada Nazar Alobaidi, Rehad Emad Shamany, and Zakariya Yahya Algamal*
Department of Statistics and Informatics, University of Mosul, Mosul, Iraq.
*Corresponding author; e-mail: zakariya.algamal@uomosul.edu.iq

Received: 30 July 2020
Revised: 9 November 2020
Accepted: 21 November 2020

Abstract

The ridge estimator has been consistently demonstrated to be an attractive shrinkage method to reduce the effects of multicollinearity. The negative binomial regression model is a well-known model in application when the response variable is count data. However, it is known that multicollinearity negatively affects the variance of maximum likelihood estimator of the negative binomial regression coefficients. To address this problem, a negative binomial ridge estimator has been proposed by numerous researchers. In this paper, a new negative binomial ridge estimator (NNBRE) is proposed and derived. The idea behind the NNBRE is to get diagonal matrix with small values of diagonal elements that leading to decrease the shrinkage parameter and, therefore, the resultant estimator can be better with small amount of bias. Our Monte Carlo simulation results suggest that the NNBRE estimator can bring significant improvement relative to other existing estimators. In addition, the real application results demonstrate that the NNBRE estimator outperforms both negative binomial ridge regression and maximum likelihood estimators in terms of predictive performance.

Keywords: Multicollinearity, ridge estimator, count data, shrinkage, Monte Carlo simulation.

1. Introduction

The negative binomial regression model (NBRM) is one basic model for count data analysis. "This model has found a widespread use in several real data fields, such as health, social, economic and physical sciences when the response variable comes in the form of non-negative integers or counts. The NBRM has three basic assumptions: on the conditional distribution of the dependent variable, on the specification of the mean parameter and on the independence of the distribution for all observations (Algamal 2012, Cameron and Trivedi 2013, De Jong and Heller 2008). In dealing with the NBRM, it is assumed that there is no correlation among the explanatory variables. In practice, however, this assumption often not holds, which leads to the problem of multicollinearity. In the presence of multicollinearity, when estimating the regression coefficients for NBRM using the maximum likelihood (ML) method, the estimated coefficients are usually become unstable with a high variance, and therefore low statistical significance (Kibria et al. 2015). Numerous remedial methods have been proposed to overcome the problem of multicollinearity. The ridge regression method (Hoerl and

Kennard 1970) has been consistently demonstrated to be an attractive and alternative to the ML estimation method.

Ridge regression is a shrinkage method that shrinks all regression coefficients toward zero to reduce the large variance (Asar and Genç 2015). This done by adding a positive amount to the diagonal of $\mathbf{X}^T\mathbf{X}$. As a result, the ridge estimator is biased but it guaranties a smaller mean squared error than the ML estimator.

In linear regression, the ridge estimator is defined as

$$\hat{\beta}_{Ridge} = (X^T X + kI)^{-1} X^T y, \tag{1}$$

where $\mathbf{y}$ is an $n\times 1$ vector of observations of the response variable, $X = (x_1,...,x_p)$ is an $n\times p$ known design matrix of explanatory variables, $\beta = (\beta_1,...,\beta_p)$ is a $p\times 1$ vector of unknown regression coefficients, $\mathbf{I}$ is the identity matrix with dimension $p\times p$, and $k \geq 0$ represents the ridge parameter (shrinkage parameter). The ridge parameter, k, controls the shrinkage of $\boldsymbol{\beta}$ toward zero. The OLS estimator can be considered as a special estimator from (1) with $k=0$. For larger value of k, the $\hat{\boldsymbol{\beta}}_{Ridge}$ estimator yields greater shrinkage approaching zero (Hoerl and Kennard 1970, Algamal and Lee 2015, Rashad and Algamal 2019)".

2. The Negative Binomial Ridge Estimator

The negative binomial regression model is commonly used for analyzing count data when the dependent variable y_i is distributed as $NB(\pi,\theta)$ where a sequence of identical and independent Bernoulli trials, all with success probability π, are observed until successes θ are observed. "Here, the expectation is $\mu = \theta(1-\pi)/\pi$ and the variance $\mu = \theta(1-\pi)/\pi^2$.

The probability function of the negative binomial distribution is given by

$$f(y;\pi,\theta) = \frac{(y_i+\theta-1)!}{y_i!(\theta-1)!}\pi_i^{\theta}(1-\pi_i)^{y_i}. \tag{2}$$

The re-parameterization of the negative binomial with parameter π and θ can be rewritten with μ and $\alpha = \frac{1}{\theta}$, which is an over dispersion parameter, as

$$f(y;\pi,\theta) = \frac{\left(y_i+\frac{1}{\alpha}-1\right)!}{y_i!(\frac{1}{\alpha}-1)!}\left(\frac{1}{1+\alpha\mu_i}\right)^{\frac{1}{\alpha}}\left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i}, \tag{3}$$

where $\mu_i = (1-\pi_i)/\alpha\pi_i$ and $\pi_i = 1/(1+\alpha\mu_i)$. The log-likelihood of the negative binomial model is given by

$$\mathrm{L}(\mu;\mathrm{y},\alpha) = \sum_{i=1}^{n}\left\{\begin{array}{l} y_i\log(\alpha\mu_i) - \left(y_i+\frac{1}{\alpha}\right)\log(1+\alpha\mu_i) + \log\Gamma\left(y_i+\frac{1}{\alpha}\right) \\ -\log\Gamma\left(y_i+1\right) - \log\Gamma\left(\frac{1}{\alpha}\right) \end{array}\right\}. \tag{4}$$

Let X_i be the i^{th} row of i^{th} row of X which is an $n\times p$ data matrix with p independent variables and let be an $n\times p$ vector of coefficients. β is usually estimated by the maximum likelihood estimation (MLE) which is found by maximizing the log-likelihood given by

$$\ell(\beta,\theta)=\sum_{i=1}^{n}\left\{\begin{array}{l}\sum_{t=1}^{y_i-1}\ln(t+\theta))-\ln\left(y_i!\right)-\left(y_i+\theta\right)\ln\left(1+\frac{1}{\theta}\exp(\mathrm{x}_i^T\beta)\right)\\+y_i\ln\left(\frac{1}{\theta}\right)+y_i(\mathrm{x}_i^T\beta)\end{array}\right\}, \quad (5)$$

where $\mu_i=\exp(\mathrm{X}_i^T\beta)$ and $\log\left[\frac{\Gamma(\theta+\mathrm{y}_i)}{\Gamma(\theta)}\right]=\sum_{t=0}^{yi-1}(t+\theta).$

The vector of coefficients using the MLE is then estimated by solving the likelihood equation,

$$S(\beta)=\frac{\partial\ell(\beta,\theta)}{\partial\beta}=\sum_{i=1}^{n}\frac{y_i-\mu_i}{1+\left(\frac{1}{\theta}\right)\mu_i}x_i=0. \quad (6)$$

Because (6) is nonlinear for β, the solution of likelihood equation is found by the iterative weighted least square (IWLS) algorithm as

$$\hat{\beta}_{ML}=(X^T\hat{W}X)^{-1}X^T\hat{W}\,\hat{s}, \quad (7)$$

where $\hat{W}=diag\left[\hat{\mu}_i/(1+(1/\theta)\hat{\mu}_i\right]$ and $\hat{s}$ is a vector where i^{th} element equals to

$$\hat{s}=\log(\hat{\mu}_i)+\frac{y_i-\hat{\mu}_i}{\hat{\mu}_i}.$$

The MLE estimator of β is asymptotically normally distributed with a covariance matrix $Cov(\hat{\beta}_{ML})=(X^T\hat{W}X)^{-1}$. The MSE based on the asymptotic covariance matrix equals

$$E(\hat{\beta}_{ML}-\beta)^T(\hat{\beta}_{ML}-\beta)=tr\left[(X^T\hat{W}X)^{-1}\right]=\sum_{j=1}^{p}\frac{1}{\lambda_{jML}}, \quad (8)$$

where λ_{jML} is the i^{th} eigenvalue of the $X^T\hat{W}X$ matrix. Note that the weighted matrix of cross products, $X^T\hat{W}X$, is ill-conditioned which leads to instability and high variance of the MLE estimator when the independent variables are highly correlated. In that situation, it is difficult to interpret the estimated parameters since the vector of estimated coefficients become too long.

Due to the presence of multicollinearity in the negative binomial regression analysis, the negative binomial ridge estimator (NBRE) is proposed by Månsson (2012) as follows

$$\hat{\beta}_{NBRE}=(X^T\hat{W}X+kI)^{-1}X^T\hat{W}X\hat{\beta}_{ML}=(X^T\hat{W}X+kI)^{-1}X^T\hat{W}X. \quad (9)$$

Note that this type shrinkage estimator minimizes the increase in the weighted sum of squared error. Hence, the shrinkage parameter, k, may take on values between zero and infinity. The ML estimator can be considered as a special estimator from (9) with $k=0$. Regardless of k value, the MSE of the $\hat{\beta}_{NBRE}$ is smaller than that of $\hat{\beta}_{ML}$ because the MSE of $\hat{\beta}_{NBRE}$ is equal to (Kibria et al. 2015)

$$\mathrm{MSE}(\hat{\beta}_{NBRE})=\sum_{j=1}^{p}\frac{\lambda_j}{(\lambda_j+k)^2}+k^2\sum_{j=1}^{p}\frac{\alpha_j}{(\lambda_j+k)^2}, \quad (10)$$

where α_j is defined as the j^{th} element of $\gamma\hat{\beta}_{ML}$ and γ is the eigenvector of the $X^T\hat{W}X$ matrix. Comparing with the MSE of (5), $\mathrm{MSE}(\hat{\beta}_{NBRE})$ is always small for $k>0$.

3. The Proposed Estimator

In this section, the new estimator is introduced and derived. Let $Q=(q_1,q_2,...,q_p)$ and $\Lambda = diag(\lambda_1,\lambda_2,...,\lambda_p),$ respectively, be the matrices of eigenvectors and eigenvalues of the $X^T\hat{W}X$ matrix, such that $M^T X^T \hat{W} X Q = L^T \hat{W} L = \Lambda,$ where $L = XQ.$ Consequently, the Poisson regression estimator of (5), $\hat{\beta}_{ML},$ can be written as

$$\hat{\gamma}_{ML} = \Lambda^{-1} L^T \hat{W}\hat{v}, \tag{11}$$

$$\hat{\beta}_{ML} = Q\hat{\gamma}_{ML}.$$

Accordingly, the negative binomial ridge estimator, $\hat{\beta}_{NBRE}$, is rewritten as

$$\hat{\gamma}_{NBRE} = (\Lambda + K)^{-1} L^T \hat{W} v = (I - KD^{-1})\hat{\gamma}_{ML}, \tag{12}$$

where $D = \Lambda + K$ and $K = diag(k_1,k_2,...,k_p);\ k_i {}^3 0, i = 1,2,...,p.$ Equation (12) represents the generalized ridge negative binomial regression estimator.

In generalized ridge estimator, the Jackknifing approach was used (Nyquist 1988, Khurana et al. 2014, Singh et al. 1986). Batah et al. (2008) proposed a modified Jackknifed ridge regression estimator in linear regression model. Related to negative binomial regression model, Türkan and Özel (2017) proposed a modified Jackknifed negative binomial ridge estimator depending on the study of Singh et al. (1986). Several studies dealt with the negative binomial regression model in the presence of multicollinearity (Huang and Yang 2012, KaÇiranlar and Dawoud 2018, Kandemir et al. 2019, Månsson 2013).

In this paper, the new estimator (JNBRE) is derived by following the study of Batah et al. (2008). Let the Jackknife estimator (JE), in negative binomial regression, is defined as

$$\hat{\gamma}_{JE} = (I - K^2 D^{-2})\hat{\gamma}_{ML}, \tag{13}$$

and the modified Jackknife estimator (MJE) of Batah et al. (2008), in negative binomial regression model, is defined as

$$\hat{\gamma}_{MJE} = (I - KD^{-1})(I - K^2 D^{-2})\hat{\gamma}_{ML}. \tag{14}$$

Consequently, our new estimator is an improvement of (14) by multiplying it with the amount $[(I - K^3 D^{-3})/(I - K^2 D^{-2})].$ The idea behind this is to get diagonal matrix with small values of diagonal elements which leading to decrease the shrinkage parameter, and, therefore, the resultant estimator can be better with small amount of bias. The new estimator is defined as

$$\hat{\gamma}_{JNBRE} = (I - KD^{-1})(I - K^2 D^{-2})\frac{(I - K^3 D^{-3})}{(I - K^2 D^{-2})}\hat{\gamma}_{ML}, \tag{15}$$

and

$$\hat{\beta}_{JNBRE} = Q^T \hat{\gamma}_{JNBRE}. \tag{16}$$

3.1. Bias, variance and MSE of the new estimator

The MSE of the new estimator can be obtained as

$$\text{MSE}(\hat{\gamma}_{JNBRE}) = \text{var}(\hat{\gamma}_{JNBRE}) + [\text{bias}(\hat{\gamma}_{JNBRE})]^2. \tag{17}$$

According to (16), the bias and variance of $\hat{\gamma}_{JNBRE}$ can be obtained as, respectively,

$$\begin{aligned}\text{bias}(\hat{\gamma}_{JNBRE}) &= E[\hat{\gamma}_{JNBRE}] - \gamma \\ &= (I - KD^{-1})(I - K^3 D^{-3})E[\hat{\gamma}_{ML}] - \gamma \\ &= -K\left[(KD^{-1})^{-1} - (KD^{-1})^{-1}(I - KD^{-1}) + K^2 D^{-2}(I - KD^{-1})\right]D^{-1}\gamma,\end{aligned} \tag{18}$$

$$\begin{aligned}\operatorname{var}(\hat{\gamma}_{JNBRE}) &= (I-KD^{-1})(I-K^3D^{-3})\operatorname{var}(\hat{\gamma}_{ML})(I-K^3D^{-3})^T(I-KD^{-1})^T \\ &= (I-KD^{-1})(I-K^3D^{-3})\Lambda^{-1}(I-K^3D^{-3})^T(I-KD^{-1})^T.\end{aligned} \tag{19}$$

Then,

$$\begin{aligned}\operatorname{MSE}(\hat{\gamma}_{JNBRE}) &= (I-KD^{-1})(I-K^3D^{-3})\Lambda^{-1}(I-K^3D^{-3})^T(I-KD^{-1})^T + \\ &\quad \left[-K\left[(KD^{-1})^{-1}-(KD^{-1})^{-1}(I-KD^{-1})+K^2D^{-2}(I-KD^{-1})\right]D^{-1}\gamma\right] \\ &\quad \left[-K\left[(KD^{-1})^{-1}-(KD^{-1})^{-1}(I-KD^{-1})+K^2D^{-2}(I-KD^{-1})\right]D^{-1}\gamma\right]^T \\ &= \Phi\Lambda^{-1}\Phi^T + K\Psi D^{-1}\gamma\gamma^T D^{-1}\Psi^T K,\end{aligned} \tag{20}$$

where $\Phi = (I-K^3D^{-3})^T(I-KD^{-1})$ and $\Psi = [I+KD^{-1}-KD^{-3}K]$.

3.2. Selection of parameter k

The efficiency of ridge estimator strongly depends on appropriately choosing the k parameter. To estimate the values of k for our new estimator, the most well-known used estimation methods are employed and are given below (Kibria et al. 2015)

1. Hoerl and Kennard (1970) (K1), which is defined as

$$k_j(K1) = \frac{1}{\hat{\alpha}^2_{\max}}, \quad j = 1,2,...,p, \tag{21}$$

2. Kibria et al. (2015) (K2), which is defined as

$$k_j(\mathrm{K2}) = \operatorname{median}\left\{\left[\sqrt{\frac{1}{\hat{\alpha}_j^2}}\right]^2\right\}, \quad j = 1,2,...,p, \tag{22}$$

3. Kibria et al. (2015) (K3), which is defined as

$$k_j(\mathrm{K3}) = \operatorname{median}\left\{\frac{\lambda_{\max}}{(n-p)+\lambda_{\max}\hat{\alpha}_j^2}\right\}, \quad j = 1,2,...,p. \tag{23}$$

4. Simulation Study

In this section, a Monte Carlo simulation experiment is used to examine the performance of the new estimator with different degrees of multicollinearity.

4.1. Simulation design

The response variable of n observations is generated from negative binomial regression model by

$$\mu_i = \exp(\mathbf{x}_i^T\boldsymbol{\beta}), \tag{24}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_p)$ with $\sum_{j=1}^{p}\beta_j^2 = 1$ and $\beta_1 = \beta_2 = ... = \beta_p$ (Kibria 2003, Månsson and Shukur 2011). In addition, the value of α are chosen as 1 and 2.

The explanatory variables $\mathbf{x}_i^T = (x_{i1}, x_{i2}, ..., x_{in})$ have been generated from the following formula

$$x_{ij} = (1-\rho^2)^{1/2} w_{ij} + \rho w_{ip}, \quad i = 1,2,...,n, \quad j = 1,2,...,p, \tag{25}$$

where ρ represents the correlation between the explanatory variables and w_{ij} 's are independent standard normal pseudo-random numbers. Because the sample size has direct impact on the prediction accuracy, three representative values of the sample size are considered: 30, 50 and 100. In addition, the number of the explanatory variables is considered as $p=4$ and $p=8$ because increasing the number of explanatory variables can lead to increase the MSE. Further, because we are interested in the effect of multicollinearity, in which the degrees of correlation considered more important, three values of the pairwise correlation are considered with $\rho=\{0.90, 0.95, 0.99\}$. For a combination of these different values of n, p, α and ρ the generated data is repeated 1000 times and the averaged mean squared errors (MSE) is calculated as

$$MSE(\hat{\beta})=\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\beta}-\beta)^T(\hat{\beta}-\beta), \tag{26}$$

where $\hat{\beta}$ is the estimated coefficients for the used estimator.

4.2. Simulation results

The estimated MSE of (26) for MLE, NBRE and JNBRE, for all the different selection methods of k and the combination of n, p, α and $\rho,$ are respectively summarized in Tables 1 and 2. Several observations can be made.

First, in terms of ρ values, there is increasing in the MSE values when the correlation degree increases regardless the value of $n, p, \alpha.$ However, JNBRE performs better than NBRE and MLE for all the different selection methods of $k.$ For instance, in Table 1, when $p=4,$ $n=100,$ and $\rho=0.95,$ the MSE of JNBRE was about 51.78%, 36.41% and 20.38% lower than that of NBRE for K1, K2 and K3, respectively. In addition, the MSE of JNBRE was about 96.19% lower than that of ML.

Second, regarding the number of explanatory variables, it is easily seen that there is increasing in the MSE values when the p increasing from four variables to eight variables. Although this increasing can affected the quality of an estimator, JNBRE is achieved the lowest MSE comparing with MLE and NBRE, for different n, ρ, α and different selection methods of $k.$

Third, with respect to the value of $n,$ The MSE values decreases when n increases, regardless the value of ρ, p, α and the value of k. However, JNBRE still consistently outperforms NBRE and ML by providing the lowest MSE.

Fourth, in terms of the value of the α and for a given values of ρ, p, n and the value of $k,$ JNBRE is always show smaller MSE comparing with NBRE and ML estimator.

Finally, for the different selection methods of $k,$ the performance of all methods suggesting that the JNBRE estimator is better than the other used two estimators. The K1 efficiently provides less MSE comparing with the K2 and K1 for both JNBRE and NBRE estimators. Besides, K1 is more efficient for providing less MSE than K2 or both JNBRE and NBRE estimators".

To summary, all the considered values of n, ρ, p, α , and the value of $k,$ JNBRE is superior to NBRE, clearly indicating that the new proposed estimator is more efficient.

5. Real Application

To further investigate the usefulness of our new estimator, we apply the proposed estimator to the football Spanish La Liga, season 2016-2017. "This data contains 20 teams. The response variable represents the number of won matches. The six considerable explanatory variables included the number of yellow cards $(x_1),$ the number of red cards $(x_2),$ the total number of substitutions $(x_3),$ the

Table 1 MSE values when $\alpha = 1$

p	n	ρ		K1		K2		K3	
			ML	NBRE	JNBRE	NBRE	JNBRE	NBRE	JNBRE
4	30	0.90	4.884	0.923	0.77	0.563	0.462	1.308	1.207
		0.95	5.512	1.154	1.003	1.012	0.911	1.467	1.366
		0.99	5.910	1.804	1.652	1.544	1.443	1.813	1.712
	50	0.90	3.255	0.556	0.403	0.474	0.373	0.849	0.748
		0.95	4.330	0.828	0.675	0.583	0.482	0.945	0.844
		0.99	4.522	1.145	0.992	1.473	1.372	1.158	1.057
	100	0.90	3.098	0.358	0.205	0.445	0.344	0.706	0.605
		0.95	3.308	0.482	0.329	0.472	0.371	0.722	0.621
		0.99	4.063	1.508	1.355	1.248	1.147	0.840	0.739
8	30	0.90	4.989	1.125	0.972	0.755	0.654	1.503	1.402
		0.95	5.608	1.356	1.203	1.204	1.103	1.662	1.561
		0.99	6.023	2.006	1.853	1.736	1.635	2.008	1.907
	50	0.90	3.524	0.758	0.605	0.666	0.565	1.044	0.943
		0.95	4.667	1.030	0.877	0.775	0.674	1.140	1.039
		0.99	4.992	1.347	1.194	1.665	1.564	1.353	1.252
	100	0.90	3.434	0.550	0.397	0.637	0.536	0.902	0.801
		0.95	3.709	0.674	0.522	0.664	0.563	0.917	0.816
		0.99	4.267	1.700	1.547	1.440	1.339	1.035	0.934

Table 2 MSE values when $\alpha = 2$

p	n	ρ		K1		K2		K3	
			ML	NBRE	JNBRE	NBRE	JNBRE	NBRE	JNBRE
4	30	0.90	5.024	1.063	0.91	0.703	0.602	1.448	1.347
		0.95	5.652	1.294	1.143	1.152	1.051	1.607	1.506
		0.99	6.050	1.944	1.792	1.684	1.583	1.953	1.852
	50	0.90	3.395	0.696	0.543	0.614	0.513	0.989	0.888
		0.95	4.470	0.968	0.815	0.723	0.622	1.085	0.984
		0.99	4.662	1.285	1.132	1.613	1.512	1.298	1.197
	100	0.90	3.238	0.498	0.345	0.585	0.484	0.846	0.745
		0.95	3.448	0.622	0.469	0.612	0.511	0.862	0.761
		0.99	4.203	1.648	1.495	1.388	1.287	0.98	0.879
8	30	0.90	5.129	1.265	1.112	0.895	0.794	1.643	1.542
		0.95	5.748	1.496	1.343	1.344	1.243	1.802	1.701
		0.99	6.163	2.146	1.993	1.876	1.775	2.148	2.047
	50	0.90	3.664	0.898	0.745	0.806	0.705	1.184	1.083
		0.95	4.807	1.170	1.017	0.915	0.814	1.280	1.179
		0.99	5.132	1.487	1.334	1.805	1.704	1.493	1.392
	100	0.90	3.574	0.690	0.537	0.777	0.676	1.042	0.941
		0.95	3.849	0.814	0.662	0.804	0.703	1.057	0.956
		0.99	4.407	1.840	1.687	1.580	1.479	1.175	1.074

number of matches with 2.5 goals on average (x_4), the number of matches that ended with goals (x_5), and the ratio of the goal scores to the number of matches (x_6).

First, the deviance test (Montgomery et al. 2015) is used to check whether the negative binomial regression model is fit well to this data or not. The result of the residual deviance test is equal to 8.651 with 14 degrees of freedom and the p-value is 0.822. It is indicated form this result that the negative binomial regression model fits very well to this data.

Second, to check whether there are relationships between the explanatory variables or not, Figure 1 displays the correlation matrix among the six explanatory variables. It is obviously seen that there are correlations greater than 0.82 between x_1 and x_6, x_1 and x_4, x_2 and x_4, and x_4 and x_6.

Third, to test the existence of multicollinearity, the eigenvalues of the matrix $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$ are obtained as 997.247, 321.922, 170.541, 41.386, 22.694, and 2.054. The determined condition number $\text{CN} = \sqrt{\lambda_{\max} / \lambda_{\min}}$ of the data is 22.034 indicating that the multicollinearity issue is exist.

The estimated negative binomial regression coefficients and MSE values for the MLE, NBRE, and JNBRE estimators are listed in Table 3. According to Table 3, it is clearly seen that the JNBRE estimator shrinkages the value of the estimated coefficients efficiently. Furthermore, in terms of the selection method of k, JNBRE shows the superiority results of coefficient estimation using K2". In terms of MSE, the JNBRE using K2 achieves the lowest MSE.

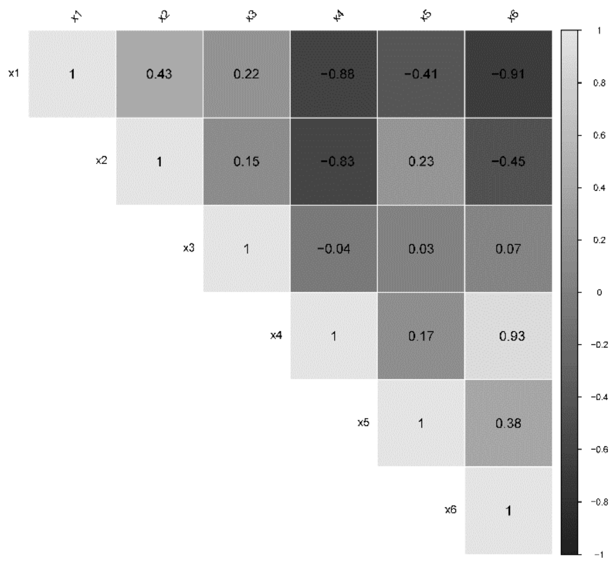

Figure 1 The correlation matrix among the six explanatory variables

6. Conclusions

In this paper, a new estimator of negative binomial ridge regression is proposed to overcome the multicollinearity problem in the negative binomial regression model. According to Monte Carlo simulation studies, the proposed estimator has better performance than maximum likelihood estimator and ordinary negative binomial ridge estimator, in terms of MSE. Additionally, a real data application is also considered to illustrate benefits of using the new estimator in the context of negative binomial regression model. The superiority of the new estimator based on the resulting MSE was observed and it was shown that the results are consistent with Monte Carlo simulation results.

Table 3 The estimated coefficients and MSE values for the MLE, NBRE, and JNBRE estimators

	K1			K2		K3	
	MLE	NBRE	JNBRE	NBRE	JNBRE	NBRE	JNBRE
$\hat{\beta}_1$	−1.219	−1.057	−0.516	−0.252	−0.615	−1.223	−0.824
$\hat{\beta}_2$	0.441	0.135	0.084	0.032	0.014	0.440	0.438
$\hat{\beta}_3$	0.575	0.127	0.016	0.096	0.012	0.576	0.393
$\hat{\beta}_4$	−3.476	−1.158	−0.134	−0.063	−0.114	−3.047	−0.619
$\hat{\beta}_5$	−2.432	−1.118	−0.008	−0.0162	−0.007	−2.419	−1.626
$\hat{\beta}_6$	5.121	2.173	1.077	0.066	0.093	4.166	1.642
MSE	4.148	2.102	1.024	0.987	0.659	1.311	0.975

Acknowledgements

The authors are very grateful to the University of Mosul/College of Computers Sciences and Mathematics for their provided facilities, which helped to improve the quality of this work.

References

Algamal ZY. Diagnostic in Poisson regression models. Electr J Appl Stat Anal. 2012; 5: 178-186.

Cameron AC, Trivedi PK. Regression analysis of count data. Cambridge: Cambridge University Press; 2013.

De Jong P, Heller GZ. Generalized linear models for insurance data. Cambridge: Cambridge University Press; 2008.

Kibria BMG, Månsson K, Shukur G. A simulation study of some biasing parameters for the ridge type estimation of Poisson regression. Commun Stat - Simul Comput. 2015; 44: 943-957.

Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics. 1970; 12: 55-67.

Asar Y and Genç A. New shrinkage parameters for the Liu-type logistic estimators. Commun Stat - Simul Comput. 2015; 45: 1094-1103.

Algamal ZY, Lee MH. Penalized Poisson regression model using adaptive modified elastic net penalty. Electr J Appl Stat Anal. 2015; 8: 236-245.

Rashad NK, Algamal ZY. A new ridge estimator for the Poisson regression model. Iran J Sci Technol Trans A Sci. 43(6): 2921-2928.

Månsson K. On ridge estimators for the negative binomial regression model. Econ Model. 2012; 29: 178-184.

Nyquist H. Applications of the jackknife procedure in ridge regression. Comput Stat Data Anal. 1988; 6: 177-183.

Khurana M, Chaubey YP, Chandra S. Jackknifing the ridge regression estimator: a revisit. Commun Stat - Theory Methods. 2014; 43: 5249-5262.

Singh B, Chaubey Y, Dwivedi T. An almost unbiased ridge estimator. Sankhyā Ser B. 1986; 13: 342-346.

Batah FSM, Ramanathan TV, Gore SD. The efficiency of modified jackknife and ridge type regression estimators: a comparison. Surveys Math Appl. 2008; 3: 111-122.

Türkan S, Özel G. A jackknifed estimators for the negative binomial regression model. Commun Stat Simulat Comput. 2017: 1845-1865.

Huang J, Yang H. A two-parameter estimator in the negative binomial regression model. J Stat Comput Simul. 2012; 84: 124-134.

KaÇiranlar S, Dawoud I. On the performance of the Poisson and the negative binomial ridge predictors. Commun Stat - Simul Comput. 2018; 47: 1751-1770.

Kandemir Çetinkaya M, Kaçıranlar S. Improved two-parameter estimators for the negative binomial and Poisson regression models. J Stat Comput Simul. 2019: 1-16.

Månsson K. Developing a Liu estimator for the negative binomial regression model: method and application. J Stat Comput Simul. 2013; 83: 1773-1780.

Kibria BMG. Performance of some new ridge regression estimators. Commun Stat - Simul Comput. 2003; 32: 419-435.

Månsson K, Shukur G. A Poisson ridge regression estimator. Econ Model. 2011; 28: 1475-1481.

Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. New York: John Wiley & Sons; 2015.