Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v11n1p253

Shrinkage estimators for gamma regression model

By Algamal

Published: 26 April 2018

Electronic Journal of Applied Statistical Analysis
Vol. 11, Issue 01, April 2018, 253-268
DOI: 10.1285/i20705948v11n1p253

Shrinkage estimators for gamma regression model

Zakariya Yahya Algamal*

Department of Statistics and Informatics, University of Mosul

Published: 26 April 2018

To reduce the effects of multicollinearity, the ridge regression model has been efficiently demonstrated to be an attractive shrinkage method. The gamma regression model is a widely used model in application when the response variable is a positively skewed. However, it is known that the variance of maximum likelihood estimator of the gamma regression coefficients can negatively affects when the multicollinearity exists. To deal with this problem, a gamma ridge regression model (GRRM) has been proposed. The performance of GRRM is fully depending on the shrinkage parameter. In this paper, numerous selection methods of the shrinkage parameter are explored and investigated. In addition, their predictive performances are considered. Our Monte Carlo simulation results suggest that some estimators can bring significant improvement relative to others, in terms of mean squared error and prediction mean squared error.

keywords: Multicollinearity; ridge estimator; gamma regression model; shrinkage; Monte Carlo simulation.

1 Introduction

In studying several real data problems, such as health-care economics, automobile insurance claims, and medical science, gamma regression model (GRM) is a widely applied model (De Jong and Heller, 2008; Dunder et al., 2018; Malehi et al., 2015). In specific, GRM is used when the response variable under the study is not following the normal distribution or the response variable is positively skewed. Consequently, the GRM assumes that the response variable has a gamma distribution (Al-Abood and Young, 1986; Wasef Hattab, 2016).

*Corresponding author: zakariya.algamal@uomosul.edu.iq

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

In dealing with the GRM, it is assumed that there is no correlation among the regressors. In practice, however, this assumption often does not holds, which leads to the problem of multicollinearity. In the existence of multicollinearity, when estimating the regression coefficients for GRM using the maximum likelihood (ML) method, the estimated coefficients are usually become unstable with a high variance, and therefore low statistical significance (Asar and Genç, 2015; Kurtoğlu and Özkale, 2016). 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.

In classical linear regression models the following relationship is usually adopted

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon, \tag{1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations of the response variable, $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_p)$ 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, and ε is an $n \times 1$ vector of random errors with mean 0 and variance σ^2.

Ridge regression is a shrinkage method that shrinks all regression coefficients toward zero to reduce the large variance (Asar and Genç, 2015; Batah et al., 2008). This is 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 guarantees a smaller mean squared error than the ML estimator. In linear regression, the ridge estimator is defined as

$$\hat{\beta}_{Ridge} = (\mathbf{X}^T\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}, \tag{2}$$

where $\mathbf{I}$ 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 β toward zero. For larger value of k, the $\hat{\beta}_{Ridge}$ estimator yields greater shrinkage approaching zero (Hoerl and Kennard, 1970).

2 Statistical methodology

2.1 Gamma ridge regression model

Positively skewed data often arise in epidemiology, social, and economic studies. This type of data consists of nonnegative values. Gamma distribution is a well-known distribution that fits to such type of data. Gamma regression model (GRM) is used to model the relationship between the positively skewed response variable and potentially regressors (Uusipaikka, 2009; Algamal and Lee, 2017a,b; Algamal and Ali, 2017b,a; Kahya et al., 2017a; Algamal, 2008; Algamal and Lee, 2015a,b,c; Algamal et al., 2017b; Kahya et al., 2017b; Al-Fakih et al., 2015; Algamal et al., 2017a, 2016b, 2015, 2016a; Algamal, 2011; Algamal and Allyas, 2017; Al-Fakih et al., 2016; Algamal, 2017).

Let y_i be the response variable and follows a gamma distribution with nonnegative shape parameter θ_2 and nonnegative scale parameter θ_1, i.e. $y_i \sim Gamma(\theta_2, \theta_1)$, then the probability density function is defined as

$$f\left(y_i\right) = \frac{\theta_1}{\Gamma(\theta_2)}(\theta_1 y_i)^{\theta_2 - 1} e^{-\theta_1 y_i}, \quad y_i \geq 0, \tag{3}$$

with $E(y) = \theta_2/\theta_1 = \mu$ and $\text{var}(y) = \theta_2/\theta_1^2 = \mu^2/\theta_2$. Given that $\theta_1 = \theta_2/\mu$, Eq. (3) can re-parameterized as a function of the mean (μ) and the shape (θ_2) parameters and written depending on the exponential function as

$$f\left(y_i\right) = \text{EXP}\left\{\frac{y_i(-1/\mu) - log(-1/\mu)}{1/\theta_2} + c(y_i, \theta_2)\right\}, \tag{4}$$

where the canonical link function is $-1/\mu$, the dispersion parameter is $\phi = 1/\theta_2$ and $c(y_i, \theta_2) = \theta_2 \log(\theta_2) + \theta_2 \log(y_i) - \log(y_i) - \log(\Gamma(\theta_2))$. Gamma regression model is usually modeled using the canonical link function (reciprocal), $\mu_i = -1/\mathbf{x}_i^T \beta$ which is expressed as a linear combination of regressors $\mathbf{x}_i = (x_{i1}, ..., x_{ip})^T$. The log link function, $\mu_i = \exp(\mathbf{x}_i^T \beta)$, is alternatively used rather than the reciprocal link function because it ensures that $\mu_i > 0$.

The most common method of estimating the coefficients of GRM is to use the maximum likelihood method of Eq. (4). Given the assumption that the observations are independent and $\mu_i = -1/\mathbf{x}_i^T \beta$, the log-likelihood function is given by

$$\ell(\beta) = \sum_{i=1}^{n} \left\{\frac{y_i \mathbf{x}_i^T \beta - log(\mathbf{x}_i^T \beta)}{1/\theta_2} + c(y_i, \theta_2)\right\}, \tag{5}$$

the ML estimator is then obtained by computing the first derivative of the Eq. (5) and setting it equal to zero, as

$$\frac{\partial \ell(\beta)}{\partial \beta} = \frac{1}{\theta_2} \sum_{i=1}^{n} \left[y_i - \frac{1}{\mathbf{x}_i^T \beta}\right] \mathbf{x}_i = 0. \tag{6}$$

Unfortunately, the first derivative cannot be solved analytically because Eq. (6) is nonlinear in β (Algamal and Lee, 2017a, 2015b; Algamal, 2012). The iteratively weighted least squares (IWLS) algorithm or Fisher-scoring algorithm can be used to obtain the ML estimators of the gamma regression parameters. In each iteration, the parameters are updated by

$$\beta^{(r+1)} = \beta^{(r)} + I^{-1}(\beta^{(r)}) S(\beta^{(r)}), \tag{7}$$

where

$$S(\beta) = \partial \ell(\beta)/\partial \beta$$

and

$$I^{-1}(\beta) = \left(-E\left(\partial^2 \ell(\beta)/\partial \beta \partial \beta^T\right)\right)^{-1}$$

. The final step of the estimated coefficients is defined as

$$\hat{\beta}_{GRM} = (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{W}} \hat{\mathbf{u}}, \tag{8}$$

where $\hat{\mathbf{W}} = \text{diag}(\hat{\mu}_i^2)$ and $\hat{\mathbf{u}}$ is a vector where ith element equals to $\hat{u}_i = \hat{\mu}_i + ((y_i - \hat{\mu}_i)/\hat{\mu}_i^2)$.

The ML estimator is asymptotically normally distributed with a covariance matrix that corresponds to the inverse of the Hessian matrix

$$\text{cov}(\hat{\beta}_{GRM}) = \left[-E\left(\frac{\partial^2 \ell(\beta)}{\partial \beta_i \, \partial \beta_k} \right) \right]^{-1} = \phi\, (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X})^{-1}, \tag{9}$$

where the dispersion parameter $\phi = 1/\theta_2$ is estimated by (Uusipaikka, 2009)

$$\hat{\phi} = \frac{1}{(n-p)} \sum_{i=1}^{n} \left(\frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} \right)^2. \tag{10}$$

The mean squared error (MSE) of Eq. (8) can be obtained as

$$\begin{aligned} \text{MSE}\,(\hat{\beta}_{GRM}) &= E(\hat{\beta}_{GRM} - \hat{\beta})^T (\hat{\beta}_{GRM} - \hat{\beta}) \\ &= \phi\, tr[(\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X})^{-1}] \\ &= \phi \sum_{j=1}^{p} \tfrac{1}{\lambda_j}, \end{aligned} \tag{11}$$

where λ_j is the eigenvalue of the $\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X}$ matrix. In the presence of multicollinearity, the matrix $\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X}$ becomes ill-conditioned leading to high variance and instability of the ML estimator of the gamma regression parameters. As a remedy, the gamma ridge regression model (GRRM) can be defined as

$$\begin{aligned} \hat{\beta}_{GRRM} &= (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^T \hat{\mathbf{W}} \mathbf{X} \hat{\beta}_{GRM} \\ &= (\mathbf{X}^T \hat{\mathbf{W}} \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^T \hat{\mathbf{W}} \hat{\mathbf{u}}, \end{aligned} \tag{12}$$

where $k \geq 0$. The ML estimator can be considered as a special estimator from Eq. (11) with $k = 0$.

2.2 Estimating the shrinkage parameter k

The efficiency of ridge estimator is fully depending on k which controls the amount of the shrinkage. For $k = 0$, the ML estimates are obtained. On the contrary, when k takes large values, the influence of the shrinkage amount increases on the coefficient estimates. In practice, it is better to estimate the value of k. Numerous methods are available for estimating a ridge parameter, especially in linear regression. In this paper, several methods are considered and extended to estimate the value of k in gamma ridge regression model. The idea behind these used methods is obtained from the work by Hoerl and Kennard (1970), Dorugade and Kashid (2010), Asar et al. (2014), and Bhat (2016).

1. Hoerl et al. (1975); Hoerl and Kennard (1970) (HK1 and HK2), which are, respectively, defined as

$$HK1 = \frac{p\hat{\sigma}^2}{\hat{\alpha}^T \hat{\alpha}}, \;\; j = 1, 2, ..., p, \tag{13}$$

$$HK2 = \frac{\hat{\sigma}^2}{\hat{\alpha}^2_{\max}}, \tag{14}$$

Where $\hat{\alpha}$ is defined as the jth element of $\gamma\,\hat{\beta}_{GRM}$and γ is the eigenvector of the $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$ matrix, $\hat{\alpha}_{\max}$is the maximum value of $\hat{\alpha}$, and $\hat{\sigma}^2 = (y_i - \hat{\mu}_i)/\hat{\mu}_i\sqrt{\hat{\phi}}$ (scaled Pearson residual (Uusipaikka, 2009)).

2. Kibria et al. (2015) used several methods which were proposed by Kibria et al. (2012) and Muniz and Kibria (2009)(K1-K12). They are, respectively, defined as

$$\mathrm{K1} = \max\left\{\frac{1}{m_j}\right\}, \tag{15}$$

$$\mathrm{K2} = \max\left\{m_j\right\}, \tag{16}$$

$$\mathrm{K3} = \prod_{j=1}^{p}\left\{\frac{1}{m_j}\right\}^{\frac{1}{p}}, \tag{17}$$

$$\mathrm{K4} = \prod_{j=1}^{p}\left\{m_j\right\}^{\frac{1}{p}}, \tag{18}$$

$$\mathrm{K5} = \mathrm{median}\left\{\frac{1}{m_j}\right\}, \tag{19}$$

$$\mathrm{K6} = \mathrm{median}\left\{m_j\right\}, \tag{20}$$

$$\mathrm{K7} = \max\left\{\frac{1}{q_j}\right\}, \tag{21}$$

$$\mathrm{K8} = \max\left\{q_j\right\}, \tag{22}$$

$$\mathrm{K9} = \prod_{j=1}^{p}\left\{\frac{1}{q_j}\right\}^{\frac{1}{p}}, \tag{23}$$

$$\mathrm{K10} = \prod_{j=1}^{p}\left\{q_j\right\}^{\frac{1}{p}}, \tag{24}$$

$$\mathrm{K11} = \mathrm{median}\left\{\frac{1}{q_j}\right\}, \tag{25}$$

$$\mathrm{K12} = \mathrm{median}\left\{q_j\right\}, \tag{26}$$

where $m_j = \sqrt{\hat{\sigma}^2/\hat{\alpha}_j^2}$ and $q_j = \lambda_{\max}/(n-p)\hat{\sigma}^2 + \lambda_{\max}\hat{\alpha}_j^2$.

3. Dorugade and Kashid (2010) proposed to use variance inflation factor (VIF) by adding it to the HK1. This method defined as

$$\mathrm{DK} = \frac{p\hat{\sigma}^2}{\hat{\alpha}^T\hat{\alpha}} - \frac{1}{n\left(\mathrm{VIF}_j\right)_{\max}}, \tag{27}$$

where VIF_j= $1/(1-\mathrm{R}_j^2)$.

4. Asar et al. (2014) proposed five modifications of ridge parameter. They are defined as, respectively

$$\mathrm{A1} = \frac{p^2}{\lambda_{\max}^2} \frac{\hat{\sigma}^2}{\sum_{j=1}^{p} \hat{\alpha}_j^2}, \tag{28}$$

$$\mathrm{A2} = \frac{p^3}{\lambda_{\max}^3} \frac{\hat{\sigma}^2}{\sum_{j=1}^{p} \hat{\alpha}_j^2}, \tag{29}$$

$$\mathrm{A3} = \frac{p}{(\lambda_{\max})^{1/3}} \frac{\hat{\sigma}^2}{\sum_{j=1}^{p} \hat{\alpha}_j^2}, \tag{30}$$

$$\mathrm{A4} = \frac{p}{(\sum_{j=1}^{p} \sqrt{\lambda_i})^{1/3}} \frac{\hat{\sigma}^2}{\sum_{j=1}^{p} \hat{\alpha}_j^2}, \tag{31}$$

$$\mathrm{A5} = \frac{2p}{\sqrt{\lambda_{\max}}} \frac{\hat{\sigma}^2}{\sum_{j=1}^{p} \hat{\alpha}_j^2}, \tag{32}$$

5. Bhat (2016) proposed two modifications of HK1. They are defined as, respectively

$$\mathrm{B1} = \frac{p\hat{\sigma}^2}{\hat{\alpha}^T\hat{\alpha}} + \frac{1}{\lambda_{\max}\hat{\alpha}^T\hat{\alpha}}, \tag{33}$$

$$\mathrm{B2} = \frac{p\hat{\sigma}^2}{\hat{\alpha}^T\hat{\alpha}} + \frac{1}{2\left(\sqrt{\lambda_{\max}/\lambda_{\max}}\right)^2}, \tag{34}$$

3 Simulation study

In this section, a Monte Carlo simulation experiment is used to examine the performance of these methods in GRRM with different degrees of multicollinearity.

3.1 Simulation design

The response variable of n observations from gamma regression model is generated by Kurtoğlu and Özkale (2016)

$$y_i \sim Gamma(\theta^2/var, \mathrm{var}/\theta) \tag{35}$$

where $\theta = \exp(\mathbf{X}^T\beta)$, var denotes θ^2, $\beta = (\beta_1, ..., \beta_p)$ with $\sum_{j=1}^{p} \beta_j^2 = 1$ and $\beta_1 = \beta_2 = ... = \beta_p$ (Kibria, 2003). The explanatory variables $\mathbf{x}_i^T = (x_{i1}, x_{i2}, ..., x_{in})$ have been generated from the following formula

$$x_{ij} = (1 - \rho^2)^{1l2} w_{ij} + \rho w_{ip}, \quad i = 1, 2, ..., n, \quad j = 1, 2, ..., p, \tag{36}$$

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, and ρ the generated data is repeated 1000 times and the averaged mean squared errors (MSE) is calculated as

$$\mathrm{MSE}(\hat{\beta}) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\beta}_{\mathrm{GRRM}} - \beta)^T (\hat{\beta}_{\mathrm{GRRM}} - \beta). \tag{37}$$

Additionally, the predictive performance according to the prediction mean squared error (PMSE) of the gamma ridge regression models depending on the type of selecting k is calculated. The PMSE is defined by

$$\mathrm{PMSE}(\hat{y}_{\mathrm{GRRM}}) = \frac{1}{1000} \sum_{i=1}^{1000} (\exp(\mathbf{X}\hat{\beta}_{\mathrm{GRRM}}) - \exp(\mathbf{X}\beta))(exp(\mathbf{X}\hat{\beta}_{\mathrm{GRRM}}) - \exp(\mathbf{X}\beta))^T. \tag{38}$$

3.2 Simulation results

The estimated MSE of Eq. (35) and PMSE of Eq. (36) for all the different selection methods of k and the combination of n, p, and ρ, are respectively summarized in Tables 1-6. Several observations can be obtained as follows:

1. In terms of ρ values, there is increasing in the MSE and PMSE values when the correlation degree increases regardless the value of n and p.
2. Regarding the number of regressors, it is easily seen that there is a negative impact on both MSE and PMSE, where there are increasing in their values when the p increasing from four regressors to eight regressors.
3. With respect to the value of n, The MSE and PMSE values decrease when n increases, regardless the value of $\rho$and p.
4. All the selection methods of k are superior to the ML estimator in terms of both MSE and PMSE.

5. Clearly, in terms of MSE, K_2 and K_8 improved the performance of the gamma ridge regression compared to ML estimator in all the cases without any domination. In contrast, A_2 estimator attained poor results comparing with the other used estimators in all cases.

6. For comparisons between the modification estimators of HK_1, i.e. DK, B_1 and B_2, it is seen that B_2 achieves the lowest MSE and PMSE compared to DK and B_1 whilst DK obtains the highest MSE and PMS among them.

7. In terms of PMSE, K_8 noticeably shows large reduction amongst others. On the other hand, A_2 appears in the second position for all cases.

Table 1: MSE values when $n = 30$

Method	$p = 4$			$p = 8$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	1.376	2.737	3.386	9.479	12.306	13.323
HK1	1.244	2.804	2.977	4.361	4.761	5.791
HK2	1.295	2.897	3.044	4.157	9.687	10.704
K1	1.310	1.325	2.847	5.189	3.911	4.928
K2	1.022	2.222	2.901	1.111	1.149	2.166
K3	1.354	2.554	2.614	8.685	9.378	10.395
K4	1.196	2.396	2.640	2.843	2.883	3.907
K5	1.331	2.531	2.701	8.533	8.667	9.684
K6	1.238	2.438	2.876	3.114	3.595	4.612
K7	1.367	2.567	2.802	2.272	3.395	4.412
K8	1.014	2.214	2.981	1.642	1.718	2.735
K9	1.373	2.573	2.772	8.691	9.347	10.364
K10	1.047	2.247	2.664	2.834	2.856	3.873
K12	1.371	2.571	2.618	8.914	9.331	10.348
DK	1.241	2.444	2.480	4.358	4.751	5.768
A1	1.377	2.577	3.386	9.479	12.305	13.322
A2	1.379	2.579	3.388	9.482	12.306	13.323
A3	1.345	2.545	3.193	8.211	10.206	11.223
A4	1.317	2.517	2.972	7.031	8.244	9.261
A5	1.351	2.551	3.223	8.552	10.761	11.778
B1	1.243	2.443	2.469	4.360	4.749	5.766
B2	1.238	2.438	2.466	4.302	4.669	5.686

Table 2: MSE values when $n = 50$

Method	$p = 4$			$p = 8$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	1.387	2.582	3.175	9.428	11.776	13.283
HK1	1.185	2.737	2.724	4.309	4.497	5.722
HK2	1.236	2.732	2.781	4.295	9.117	10.654
K1	1.251	1.158	2.584	5.127	5.341	4.878
K2	0.963	2.055	2.638	1.049	1.579	2.116
K3	1.295	2.387	2.351	8.623	8.808	10.345
K4	1.137	2.229	2.377	2.781	2.913	3.857
K5	1.272	2.364	2.438	8.471	8.597	9.634
K6	1.179	2.271	2.613	3.052	3.125	4.562
K7	1.308	2.404	2.539	2.212	2.825	4.362
K8	0.955	2.047	2.718	1.582	1.648	2.685
K9	1.314	2.406	2.509	8.629	8.777	10.314
K10	0.988	2.082	2.401	2.772	2.886	3.823
K11	1.312	2.404	2.355	8.852	8.861	10.298
DK	1.182	2.277	2.217	4.299	4.381	5.718
A1	1.318	2.411	3.123	9.417	11.735	13.272
A2	1.322	2.416	3.128	9.418	11.736	13.273
A3	1.286	2.378	2.931	8.149	9.636	11.173
A4	1.258	2.352	2.709	6.969	7.674	9.211
A5	1.292	2.384	2.961	8.491	10.191	11.728
B1	1.184	2.276	2.283	4.299	4.379	5.716
B2	1.179	2.271	2.236	4.243	4.369	5.636

Table 3: MSE values when $n = 100$

Method	$p = 4$			$p = 8$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	3.358	4.445	5.962	9.395	11.266	13.241
HK1	1.136	2.441	2.651	4.277	4.678	5.768
HK2	1.177	1.565	2.518	4.233	8.547	10.604
K1	1.192	0.991	2.321	5.065	4.771	4.828
K2	0.904	1.888	2.375	0.987	1.009	2.066
K3	1.236	2.22	2.088	8.561	8.238	10.295
K4	1.078	2.062	2.114	2.719	2.343	3.807
K5	1.213	2.197	2.175	8.409	8.527	9.584
K6	1.12	2.104	2.35	2.99	2.555	4.512
K7	1.249	2.237	2.276	2.151	2.255	4.312
K8	0.896	1.881	2.455	1.521	1.678	2.635
K9	1.255	2.239	2.246	8.567	8.207	10.264
K10	0.929	1.913	2.138	2.711	2.816	3.773
K11	1.253	2.237	2.092	8.79	8.291	10.248
DK	1.126	2.111	1.954	4.237	4.811	5.668
A1	1.259	2.244	2.861	9.352	11.165	13.222
A2	1.261	2.247	2.866	9.357	11.169	13.231
A3	1.227	2.211	2.668	8.087	9.066	11.123
A4	1.199	2.185	2.446	6.907	7.104	9.161
A5	1.233	2.217	2.698	8.429	9.621	11.678
B1	1.125	2.109	2.481	4.237	4.829	5.686
B2	1.121	2.104	2.473	4.181	4.809	5.566

Table 4: PMSE values when $n = 30$

Method	$p = 4$			$p = 8$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	8.663	8.875	9.026	8.871	8.986	9.255
HK1	4.112	4.238	4.489	4.329	4.457	4.798
HK2	4.323	4.432	4.686	4.532	4.646	4.915
K1	4.385	4.494	4.748	4.592	4.708	4.977
K2	3.096	3.205	3.459	3.303	3.419	3.688
K3	4.572	4.681	4.935	4.779	4.895	5.164
K4	3.894	4.003	4.257	4.101	4.217	4.486
K5	4.474	4.583	4.837	4.681	4.797	5.066
K6	4.079	4.188	4.442	4.286	4.402	4.671
K7	4.624	4.733	4.987	4.831	4.947	5.216
K8	3.059	3.168	3.422	3.266	3.382	3.651
K9	4.647	4.756	5.011	4.854	4.971	5.239
K10	3.209	3.318	3.572	3.416	3.532	3.801
K11	4.639	4.748	5.002	4.846	4.962	5.231
DK	4.106	4.215	4.469	4.313	4.429	4.698
A1	4.663	4.772	5.026	4.871	4.986	5.255
A2	4.691	4.801	5.102	4.886	4.989	5.257
A3	4.534	4.643	4.897	4.741	4.857	5.126
A4	4.417	4.526	4.781	4.624	4.741	5.009
A5	4.554	4.663	4.917	4.761	4.877	5.146
B1	4.081	4.212	4.464	4.308	4.424	4.693
B2	4.101	4.192	4.444	4.288	4.404	4.673

Table 5: PMSE values when $n = 50$

Method	$p = 4$			$p = 8$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	8.367	8.415	8.808	8.771	9.086	9.138
HK1	4.111	4.211	4.486	4.322	4.711	4.817
HK2	4.226	4.335	4.589	4.433	4.549	4.818
K1	4.288	4.397	4.651	4.495	4.611	4.881
K2	2.999	3.108	3.362	3.206	3.322	3.591
K3	4.475	4.584	4.838	4.682	4.798	5.067
K4	3.797	3.906	4.161	4.004	4.122	4.389
K5	4.377	4.486	4.741	4.584	4.701	4.969
K6	3.982	4.091	4.345	4.189	4.305	4.574
K7	4.527	4.636	4.892	4.734	4.852	5.119
K8	2.962	3.071	3.325	3.169	3.285	3.554
K9	4.552	4.659	4.913	4.757	4.873	5.142
K10	3.112	3.221	3.475	3.319	3.435	3.704
K11	4.542	4.651	4.905	4.749	4.865	5.134
DK	4.009	4.118	4.372	4.216	4.332	4.601
A1	4.566	4.675	4.929	4.773	4.888	5.154
A2	4.582	4.679	4.931	4.779	4.889	5.158
A3	4.437	4.546	4.802	4.644	4.762	5.029
A4	4.322	4.429	4.683	4.527	4.643	4.912
A5	4.457	4.566	4.822	4.664	4.782	5.049
B1	4.004	4.113	4.367	4.211	4.327	4.596
B2	3.984	4.093	4.347	4.191	4.307	4.576

Table 6: PMSE values when $n = 100$

Method	$p = 4$			$p = 8$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
ML	8.173	8.509	8.773	8.608	8.882	9.107
HK1	4.023	4.132	4.386	4.231	4.346	4.615
HK2	4.241	4.349	4.603	4.447	4.563	4.832
K1	4.302	4.411	4.665	4.509	4.625	4.894
K2	3.013	3.122	3.376	3.221	3.336	3.605
K3	4.489	4.598	4.852	4.696	4.812	5.081
K4	3.811	3.922	4.174	4.018	4.134	4.403
K5	4.391	4.501	4.754	4.598	4.714	4.983
K6	3.996	4.105	4.359	4.203	4.319	4.588
K7	4.541	4.651	4.904	4.748	4.864	5.133
K8	2.976	3.085	3.339	3.183	3.299	3.568
K9	4.564	4.673	4.927	4.771	4.887	5.156
K10	3.126	3.235	3.489	3.333	3.449	3.718
K11	4.556	4.665	4.919	4.763	4.879	5.148
DK	4.013	4.103	4.354	4.212	4.329	4.601
A1	4.582	4.685	4.939	4.783	4.903	5.168
A2	4.586	4.6899	4.943	4.787	4.904	5.172
A3	4.451	4.562	4.814	4.658	4.774	5.043
A4	4.334	4.443	4.697	4.541	4.657	4.926
A5	4.471	4.581	4.834	4.678	4.794	5.063
B1	4.018	4.127	4.381	4.225	4.341	4.611
B2	3.998	4.107	4.361	4.205	4.321	4.591

4 Conclusion

In this paper, numerous selection methods of the shrinkage parameter are explored and investigated of gamma ridge regression model. In addition, their predictive performances are considered. According to Monte Carlo simulation studies, it has been seen that some estimator can bring significant improvement relative to others, in terms of MSE and PSEM. The K_2 and K_8 improved the performance of the gamma ridge regression compared to ML estimator in all the cases without any domination but with superiority of K_8 in terms of PMSE. In contrast, A_2 estimator showed poor results comparing with others in all cases. Besides, B_2 achieves the lowest MSE and PMSE compared to DK and B_1 whilst DK obtains the highest MSE and PMS among them. In conclusion, the use of these estimators is recommended when multicollinearity is present in the gamma ridge regression model.

References

Al-Abood, A. M. and Young, D. H. (1986). Improved deviance goodness of fit statistics for a gamma regression model. *Communications in Statistics - Theory and Methods*, 15(6):1865–1874.

Al-Fakih, A. M., Algamal, Z. Y., Lee, M. H., Abdallah, H. H., Maarof, H., and Aziz, M. (2016). Quantitative structure-activity relationship model for prediction study of corrosion inhibition efficiency using two-stage sparse multiple linear regression. *Journal of Chemometrics*, 30(7):361–368.

Al-Fakih, A. M., Aziz, M., Abdallah, H. H., Algamal, Z. Y., Lee, M. H., and Maarof, H. (2015). High dimensional qsar study of mild steel corrosion inhibition in acidic medium by furan derivatives. *International Journal of Electrochemical Science*, 10:3568–3583.

Algamal, Z. and Allyas, I. I. (2017). Prediction of blood lead level in maternal and fetal us-ing generalized linear model. *International Journal of Advanced Statistics and Probability*, 5(2):65.

Algamal, Z. and Lee, M. (2017a). A new adaptive l1-norm for optimal descriptor selection of high-dimensional qsar classification model for anti-hepatitis c virus activity of thiourea derivatives. *SAR and QSAR in Environmental Research*, 28(1):75–90.

Algamal, Z. Y. (2008). Exponentiated exponential distribution as a failure time distribution. *IRAQI Journal of Statistical science*, 14:63–75.

Algamal, Z. Y. (2011). Paired bootstrapping procedure in gamma regression model using r. *Journal of Basrah Researches*, 37(4):201–211.

Algamal, Z. Y. (2012). Diagnostic in poisson regression models. *Electronic Journal of Applied Statistical Analysis*, 5(2):178–186.

Algamal, Z. Y. (2017). Using maximum likelihood ratio test to discriminate between the inverse gaussian and gamma distributions. *International Journal of Statistical Distributions*, 1(1):27–32.

Algamal, Z. Y. and Ali, H. T. M. (2017a). Bootstrapping pseudo - r2 measures for binary response variable model. *Biomedical Statistics and Informatics*, 2(3):107–110.

Algamal, Z. Y. and Ali, H. T. M. (2017b). An efficient gene selection method for high-dimensional microarray data based on sparse logistic regression. *Electronic Journal of Applied Statistical Analysis*, 10(1):242–256.

Algamal, Z. Y. and Lee, M. H. (2015a). Adjusted adaptive lasso in high-dimensional poisson regression model. *Modern Applied Science*, 9(4):170–179.

Algamal, Z. Y. and Lee, M. H. (2015b). Applying penalized binary logistic regression with correlation based elastic net for variables selection. *Journal of Modern Applied Statistical Methods*, 14(1):168–179.

Algamal, Z. Y. and Lee, M. H. (2015c). High dimensional logistic regression model using adjusted elastic net penalty. *Pakistan Journal of Statistics and Operation Research*, 11(4):1–10.

Algamal, Z. Y. and Lee, M. H. (2017b). A novel molecular descriptor selection method in qsar classification model based on weighted penalized logistic regression. *Journal of Chemometrics*, page e2915.

Algamal, Z. Y., Lee, M. H., and Al-Fakih, A. M. (2016a). High-dimensional quantitative structure-activity relationship modeling of influenza neuraminidase a/pr/8/34 (h1n1) inhibitors based on a two-stage adaptive penalized rank regression. *Journal of Chemometrics*, 30(2):50–57.

Algamal, Z. Y., Lee, M. H., Al-Fakih, A. M., and Aziz, M. (2015). High-dimensional qsar prediction of anticancer potency of imidazo[4,5-b]pyridine derivatives using adjusted adaptive lasso. *Journal of Chemometrics*, 29(10):547–556.

Algamal, Z. Y., Lee, M. H., Al-Fakih, A. M., and Aziz, M. (2016b). High-dimensional qsar modelling using penalized linear regression model with l1/2-norm. *SAR and QSAR in Environmental Research*, 27(9):703–19.

Algamal, Z. Y., Lee, M. H., Al-Fakih, A. M., and Aziz, M. (2017a). High-dimensional qsar classification model for anti-hepatitis c virus activity of thiourea derivatives based on the sparse logistic regression model with a bridge penalty. *Journal of Chemometrics*, 31:1–8.

Algamal, Z. Y., Qasim, M. K., and Ali, H. T. M. (2017b). A qsar classification model for neuraminidase inhibitors of influenza a viruses (h1n1) based on weighted penalized support vector machine. *SAR and QSAR in Environmental Research*, pages 1–12.

Asar, Y. and Genç, A. (2015). New shrinkage parameters for the liu-type logistic estimators. *Communications in Statistics - Simulation and Computation*, 45(3):1094–1103.

Asar, Y., Karaibrahimoğlu, A., and Genç, A. (2014). Modified ridge regression parameters: A comparative monte carlo study. *Hacettepe Journal of Mathematics and Statistics*, 43(5):827–841.

Batah, F. S. M., Ramanathan, T. V., and Gore, S. D. (2008). The efficiency of modefied jackknife and ridge type regression estimators - a comparison. *Surveys in Mathematics and its Applications*, 3:111 122.

Bhat, S. S. (2016). A comparative study on the performance of new ridge estimators. *Pakistan Journal of Statistics and Operation Research*, 12(2):317–325.

De Jong, P. and Heller, G. Z. (2008). *Generalized linear models for insurance data*, volume 10. Cambridge University Press Cambridge.

Dorugade, A. and Kashid, D. (2010). Alternative method for choosing ridge parameter for regression. *Applied Mathematical Sciences*, 4(9):447–456.

Dunder, E., Gumustekin, S., and Cengiz, M. A. (2018). Variable selection in gamma regression models via artificial bee colony algorithm. *Journal of Applied Statistics*, 45(1):8–16.

Hoerl, A. E., Kannard, R. W., and Baldwin, K. F. (1975). Ridge regression:some simulations. *Communications in Statistics*, 4(2):105–123.

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

Kahya, M. A., Al-Hayani, W., and Algamal, Z. Y. (2017a). Classification of breast cancer histopathology images based on adaptive sparse support vector machine. *Journal of Applied Mathematics & Bioinformatics*, 7(1):49–69.

Kahya, M. A., Al-Hayani, W., and Algamal, Z. Y. (2017b). Gene selection inside pathways using weighted l1-norm support vector machine. *American Journal of Computational and Applied Mathematics*, 7(4):87–94.

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

Kibria, B. M. G., Månsson, K., and Shukur, G. (2012). Performance of some logistic ridge regression estimators. *Computational Economics*, 40(4):401–414.

Kibria, B. M. G., Månsson, K., and Shukur, G. (2015). A simulation study of some biasing parameters for the ridge type estimation of poisson regression. *Communications in Statistics - Simulation and Computation*, 44(4):943–957.

Kurtoğlu, F. and Özkale, M. R. (2016). Liu estimation in generalized linear models: application on gamma distributed response variable. *Statistical Papers*, 57(4):911–928.

Malehi, A. S., Pourmotahari, F., and Angali, K. A. (2015). Statistical models for the analysis of skewed healthcare cost data: A simulation study. *Health Economics Review*, 5:1–11.

Muniz, G. and Kibria, B. G. (2009). On some ridge regression estimators: An empirical comparisons. *Communications in Statistics-Simulation and Computation*, 38(3):621–630.

Uusipaikka, E. (2009). *Confidence intervals in generalized regression models*. Chapman & Hall/CRC Press, NW.

Wasef Hattab, M. (2016). A derivation of prediction intervals for gamma regression. *Journal of Statistical Computation and Simulation*, 86(17):3512–3526.