IRAQI STATISTICIANS JOURNAL

https://isj.edu.iq/index.php/isj

ISSN: 3007-1658 (Online)

Generalized Ridge Estimator for Conway-Maxwell Poisson Regression Model

Sura Mohammed Jamal Alden Hussein, Zakariya Yahya Algamal

Department of statistics and information, College of Computer Sciences and Mathematics, University of Mosul, Iraq

ARTICLE INFO

Article history:

Received	11/11/2024
Revised	11/11/2024
Accepted	13/01/2025
Available online	15/05/2025

Keywords:

Ridge Estimator
Regression Model
Conway-Maxwell Poisson
Multicollinearity

ABSTRACT

The ridge estimator has been shown as a preferable shrinkage strategy to counter the impacts of multicollinearity. Interestingly, the Conway-Maxwell Poisson regression model is one of the most frequently used models in application where the response variable is positively skewed. It is a flexible extension of Poisson regression designed to handle count data with overdispersion or underdispersion. It generalizes the Poisson distribution by introducing a dispersion parameter that controls the tail behavior of the distribution. However, it is a well-established fact that the variance of maximum likelihood estimator (MLE) of the Conway-Maxwell Poisson regression coefficients can get dragged down due to multicollinearity. Thus, in this paper, a new approach named the generalized ridge estimator is developed to fix the flaw of the ridge estimator. Many methods for estimating the shrinkage matrix have been borrowed. These findings, based on our Monte Carlo simulation and the using of real data application results. No matter what kind of estimating method of shrinkage matrix, the proposed estimator is better than MLE estimator and ridge estimator, in terms of MSE. In addition, some estimating method of shrinkage matrix can make the improvement relatively large compared to others.

1. Introduction

Conway-Maxwell Poisson regression model is used in numerous real data problems related to automobile insurance claims, healthcare economy, and medical science [1, 2, 3]. Specifically, Conway-Maxwell Poisson regression model is used when the response variable under the study is not distributed as normal distribution or the response variable is positively skewed. Consequently, the Conway-Maxwell Poisson regression assumes that the response variable has a beta distribution [4, 5].

Looking at the estimation of the Conway-Maxwell Poisson regression model a few assumptions made include: There is no correlation between the regressing factors. In practice, nonetheless, this assumption Tart often not holds; if the independent variables in a regression model are highly interrelated, there exists multicollinearity. Also, in any regression analysis conducted while the multicollinearity problem is present, the estimated regression coefficients of Conway-Maxwell Poisson regression model while employing the maximum likelihood (ML) method are normally highly variable and statistically insignificant [6, 7]. Numerous remedial methods have been proposed to overcome the problem of multicollinearity. The ridge regression method [8] 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

* Corresponding author.E-**mail** address: sura.alalwlia@uomosul.edu.iq
ttps://doi.org/10.62933/1nzxxd87

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations of the response variable, $\mathbf{X} = (\mathbf{x_1}, \ldots, \mathbf{x_p})$ is an $n \times p$ known design matrix of explanatory variables, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ is a $p \times 1$ vector of unknown regression coefficients, and $\boldsymbol{\varepsilon}$ 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 [6, 9]. 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 guaranties a smaller mean squared error than the ML estimator.

In linear regression, the ridge estimator is defined as

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

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 $\boldsymbol{\beta}$ toward zero. For larger value of k, the $\hat{\boldsymbol{\beta}}_{Ridge}$ estimator yields greater shrinkage approaching zero [8]".

2. Statistical Methodology

2.1 Conway-Maxwell Poisson ridge regression model

In real application, count data have sometimes demonstrated overdispersion that is to say, variance > mean and at other times have indicated underdispersion, that is variance < mean. There is a form pursuant to the Conway – Maxwell – Poisson distribution (CMPD) that provides a straightforward method for addressing the extra and defective distribution [10, 11]. "The CMPD is an extension of the Poisson distribution with two parameters λ (centering parameter related to the observations mean) and θ (the shape parameter) [12]. Suppose $y \in \{0,1,2,.....\}$ is a random variable that follows a CMPD, then the probability mass function is defined as

$$P(Y = y; \lambda, \theta) = \frac{\lambda^y}{(y!)^\theta Z(\lambda, \theta)}, \quad \lambda > 1,\ \theta \geq 0, \quad (3)$$

where $Z(\lambda,\theta) = \sum_{s=0}^{\infty}\left(\lambda^s / (s!)^\theta\right)$ is a normalizing constant. The CMPD can model both underdispersed ($\theta > 1$) and overdispersed ($\theta < 1$) data.

According to Eq. (1), there is no closed form representation available for the mean. This is because the normalizing constant, $Z(\lambda,\theta)$, is an infinite series with no closed form representation [13]. Shmueli, Minka [14] used the asymptotic expression for $Z(\lambda,\theta)$ in Eq. (1) to express the mean and variance of the CMPD as

$$E(Y) \approx \lambda^{\frac{1}{\theta}} - \frac{\theta - 1}{2\theta},$$

$$Var(Y) \approx \frac{1}{\theta}\lambda^{\frac{1}{\theta}} \quad (4)$$

For regression modeling in which the count responses may change depending on a set of explanatory variables, it is more convenient and interpretable to model the mean of the CMPD directly. By setting $\mu = \lambda^{\frac{1}{\theta}}$ [15], a re-parameterization of Eq. (1) to provide a clear centering parameter is can be defined as

$$P(Y = y; \mu, \theta) = \left(\frac{\mu^y}{y!}\right)^\theta \frac{1}{S(\mu,\theta)}, \quad (5)$$

where $S(\mu,\theta) = \sum_{n=0}^{\infty}\left(\mu^n / n!\right)^\theta$.

Depending on Eq. (3) and in terms of generalized linear model framework, the Conway–Maxwell–Poisson regression model (CMPR) can be formulated as

$$\ln(\mu) = \beta_0 + \sum_{j=1}^{p}\beta_j \mathbf{x}_j, \quad (6)$$

$$\ln(\theta) = \gamma_0 + \sum_{k=1}^{q}\gamma_k \mathbf{m}_j. \quad (7)$$

In Eqs. (4) and (5), $\mathbf{x}_j$ and $\mathbf{m}_j$ are explanatory variables, and there are assumed to be p covariates used in the centering link function and q covariates used in the shape link function. Assuming θ as a dispersion parameter and using single link function, Eq. (4), with $\eta = \ln(\mu) = \boldsymbol{\beta}\mathbf{x}$ as a linear predictor with log link, where $\boldsymbol{\beta}$ is the vector of regression coefficients including intercept, the log likelihood function can be written a [16]

$$\ell(\beta) = \theta\sum_{i=1}^{n} y_i(\boldsymbol{\beta}\mathbf{x}_i) - \theta\sum_{i=1}^{n}\ln(y_i!) - \sum_{i=1}^{n}\ln\left[S(\boldsymbol{\beta}\mathbf{x}_i, \theta)\right]. \quad (8)$$

Solving Eq. (6), the estimation of the regression parameters, $\boldsymbol{\beta}$, and the estimation of the dispersion parameter, θ, can be obtained as, respectively,

$$\frac{\partial\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} = \sum_{i=1}^{n}(y_i\theta - \frac{\partial}{\partial\eta_i}\ln[S(\eta_i, \theta)])x_{ij} \quad (9)$$

$$\frac{\partial\ell(\boldsymbol{\beta})}{\partial\theta} = \sum_{i=1}^{n}(-\ln(y_i!) - \frac{\partial}{\partial\theta}\log[S(\eta_i, \theta)]) \quad (10)$$

Iterative reweighted least square (IRLS) is used to solve both Eq. (7) and Eq. (8). By fixing θ, the maximum likelihood estimator (MLE) of $\boldsymbol{\beta}$ is defined as

$$\hat{\boldsymbol{\beta}}_{\text{MLE}} = (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\hat{\mathbf{u}}, \quad (11)$$

where $\hat{\mathbf{u}} = \ln(\hat{\mu}) + \frac{(y-\hat{\mu})}{\hat{\mu}^2}$ is a vector of the adjusted response variable, and $\hat{\mathbf{W}}$ is a matrix of weights [13].

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 MLE of the Conway–Maxwell–Poisson regression model. As a remedy, a ridge estimator of Hoerl and Kennard [8] for Conway–Maxwell–Poisson regression model (CMPRE) can be defined as

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{\text{CMPRE}} &= (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}\hat{\boldsymbol{\beta}}_{\text{MLE}} \\ &= (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\hat{\mathbf{u}},\end{aligned} \quad (12)$$

where $k > 0$.

The mean squared error (MSE) of Eq. (9) can be obtained as

$$\begin{aligned}\text{MSE}(\hat{\boldsymbol{\beta}}_{\text{MLE}}) &= E(\hat{\boldsymbol{\beta}}_{\text{MLE}} - \hat{\boldsymbol{\beta}})^T(\hat{\boldsymbol{\beta}}_{\text{MLE}} - \hat{\boldsymbol{\beta}}) \\ &= \hat{\theta}\, tr[(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X})^{-1}] \\ &= \hat{\theta}\sum_{j=1}^{p}\frac{1}{\lambda_j},\end{aligned} \quad (13)$$

where λ_j is the eigenvalue of the $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$ matrix and $\hat{\theta}$ is the estimated dispersion parameter. On the other hand, the mean squared error (MSE) of Eq. (10) can be obtained as

$$\text{MSE}(\hat{\boldsymbol{\beta}}_{\text{CMPRE}}) = \hat{\theta}\sum_{j=1}^{p}\frac{\lambda_j}{(\lambda_j + k)^2} + k^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j + k)^2}, \quad (14)$$

where α_j is defined as the jth element of $\gamma\hat{\boldsymbol{\beta}}_{\text{MLE}}$ and γ is the eigenvector of the $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$.

Generalized ridge estimator

The generalized ridge estimator (GRE) is suggested to generalize the ridge estimator, the difference between ridge estimator and GRR is there are p values of k, such that [8]

$$\hat{\boldsymbol{\beta}}_{GRE} = (\mathbf{X}^T\mathbf{X} + \mathbf{K})^{-1}\mathbf{X}^T\mathbf{y}, \quad (15)$$

where $\mathbf{K} = \text{diag}(k_1, k_2, ..., k_p)$. The good thing where using GRE is to find the best values of k_i so as to obtain the MSE which is less than when we using the ridge estimator and OLS.

The generalized ridge estimator for Conway–Maxwell–Poisson regression model (GCRM) is defined as

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{GCRM} &= (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + \mathbf{K})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}\hat{\boldsymbol{\beta}}_{MLE} \\ &= (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + \mathbf{K})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\hat{\mathbf{u}}.\end{aligned} \quad (16)$$

The selection of the matrix $\mathbf{K}$ is essential. In this paper, several methods are adapted to estimate $\mathbf{K}$, such as [17], [18], [19], [20], [21],

[22], [23], [24], [25], [26], [27]and [28]. These methods are given below, respectively

$$\hat{k}_{i\,(HK)} = \frac{\hat{\nu}}{\alpha_i^2}, \quad [8] \qquad (17)$$

where α_j is defined as the jth element of $\gamma\hat{\boldsymbol{\beta}}_{GRM}$ and γ is the eigenvector of the $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$ and the dispersion parameter, ν, is" estimated by $\hat{\nu} = (1/(n-p))\sum_{i=1}^{n}\left((y_i - \hat{\theta}_i)^2 / \hat{\theta}_i^2\right)$ [29].

$$\hat{k}_{i\,(N)} = \frac{\hat{\nu}}{\hat{\alpha}_i^2}\left\{1+\left[1+\lambda_i(\hat{\alpha}_i^2/\hat{\nu})^{1/2}\right]\right\} \quad [18] \qquad (18)$$

$$\hat{k}_{i\,(TC)} = \frac{\lambda_i\hat{\nu}}{\lambda_i\hat{\alpha}_i^2+\hat{\nu}} \quad [19] \qquad (19)$$

$$\hat{k}_{i\,(F)} = \frac{\lambda_i\hat{\nu}}{\lambda_i\hat{\alpha}_i^2+(n-p)\hat{\nu}} \quad [20] \qquad (20)$$

$$\hat{k}_{i\,(HSL)} = \hat{\nu}\frac{\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2)^2}{(\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2))^2} \quad [17] \qquad (21)$$

$$\hat{k}_{i\,(AH)} = \hat{\nu}\frac{\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2)^2}{(\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i^2))^2}+\frac{1}{\lambda_{\max}} \quad [23] \qquad (22)$$

$$\hat{k}_{i\,(D)} = \frac{\hat{\nu}}{\lambda_{\max}\hat{\alpha}_i^2} \quad [26] \qquad (23)$$

$$\hat{k}_{i\,(SB)} = \frac{\lambda_i\hat{\nu}}{\lambda_i\hat{\alpha}_i^2+\hat{\nu}}+\frac{1}{\lambda_{\max}} \quad [28] \qquad (24)$$

$$\hat{k}_{i\,(SV1)} = \frac{p\hat{\nu}}{\hat{\alpha}_i^2}+\frac{1}{\lambda_{Max}\hat{\alpha}_i^2} \quad [28] \qquad (25)$$

$$\hat{k}_{i\,(SV2)} = \frac{p\hat{\nu}}{\hat{\alpha}_i^2}+\frac{1}{2\left(\sqrt{\lambda_{Max}/\lambda_{Min}}\right)^2} \quad [28] \qquad (26)$$

$$\hat{k}_{i\,(M)} = \frac{1}{\frac{\lambda_{Max}\hat{\alpha}_i^2}{(n-p)\hat{\nu}+\lambda_{Max}\hat{\alpha}_i^2}} \quad [28] \qquad (27)$$

$$\hat{k}_{i\,(AS)} = \frac{\hat{\nu}}{\hat{\alpha}_i^2}+\frac{1}{\lambda_i} \quad [27] \qquad (28)$$

.

3. Results and discussion

3.1. Simulation study

For this section, a Monte Carlo experiment to compare the performance of these methods in GCRM with varying levels of multicollinearity is performed.

3.1 Simulation design

The response variable of n observations from gamma regression model is generated by [7, 30, 31, 32, 33, 34, 35, 36]

$$y_i \sim CMP(\mu_i, \theta) \qquad (29)$$

where $\mu_i = \exp(\mathbf{x}_i^T\boldsymbol{\beta})$, $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)$ with $\sum_{j=1}^{p}\beta_j^2 = 1$ and $\beta_1 = \beta_2 = ... = \beta_p$ [37], and $\nu \in \{0.50, 2\}$ [38]. The explanatory variables $\mathbf{x}_i^T = (x_{i1}, x_{i2}, ..., x_{in})$ have been generated from the following formula

$$x_{ij} = (1-\rho^2)^{1/2}w_{ij} + \rho w_{ip}, \; i = 1,2,...,n, \quad j = 1,2,...,p, \qquad (30)$$

where ρ represents the correlation between the explanatory variables and w_{ij}'s are independent standard normal pseudo-random numbers. n= 50, 100 and 200. In addition, $p = 4$ and $p = 8$ because increasing the number of explanatory variables can lead to increase the MSE. $\rho = \{0.90, 0.95, 0.99\}$. The averaged mean squared errors (MSE) is calculated as

$$\text{MSE}(\hat{\boldsymbol{\beta}}) = \frac{1}{1000}\sum_{i=1}^{1000}(\hat{\boldsymbol{\beta}}_{\text{GCRM}} - \boldsymbol{\beta})^T(\hat{\boldsymbol{\beta}}_{\text{GCRM}} - \boldsymbol{\beta}). \quad (31)$$

3.2 Simulation results

The averaged MSE all the combinations of n, ν, p, and ρ, are respectively summarized in Tables 1 – 6. "The best value of the averaged MSE is highlighted in bold. Several observations can be obtained as follows:

1- In general analysis, the MSE of GRRM is less than that of MLE.

2- As expected GCRM achieved a lower MSE compared to GRRM regardless the type of estimating method of K matrix.

3- Clearly, in terms of MSE, F method (Eq. (14)), which was proposed by [20], Indeed, in proportional to the MSE, F method (14) by [20] indicated that the gamma generalized ridge estimator outperformed the other approaches in all examined cases. On the other hand, HK & SB methods gave very low results compared with the other used methods in all the cases.The results regarding the number of explanatory variables are quite simple; increasing their values have a negative impact on the MSE as the r increasing.In terms of ρ values, there is increasing in the MSE values when the correlation degree increases regardless the value of n, ν and p with superiority of F method.

4- Regarding the number of explanatory variables, it is easily seen that there is a negative impact on MSE, where there are increasing in their values when the p increasing.

5- Concerning the value of n, the MSE values decrease when n increases, regardless the value of ρ, ν and p.

6- For fixed n, p and degree of multicollinearity ρ, as the ν increases the MSE of all methods decreases".

Table 1: MSE on average when $n = 50$ and $p = 4$

Methods	$\nu = 0.5$			$\nu = 2$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.2535	5.4166	5.6312	5.1599	5.3215	5.5917
GRRM	3.7795	3.8008	3.7785	3.6636	3.7105	3.7562
HK	3.3342	3.3885	3.4031	3.3271	3.341	3.3532
N	2.9976	3.0058	3.0082	2.9904	2.9886	2.9838
TC	3.0762	3.1305	3.1451	3.0691	3.083	3.0952
F	2.5508	2.6051	2.6197	2.5437	2.5576	2.5698
HSL	2.9632	3.0175	3.0321	2.9561	2.97	2.9822
AH	2.8884	2.9427	2.9573	2.8813	2.8952	2.9074
D	2.7837	2.7945	2.8001	2.7753	2.7835	2.796
SB	3.0809	3.1352	3.1498	3.0738	3.0877	3.0999
SV1	2.9218	2.9761	2.9907	2.9147	2.9286	2.9408
SV2	2.8671	2.9214	2.936	2.86	2.8739	2.8861
M	2.8859	2.9397	2.9543	2.8783	2.8922	2.9045
AS	2.9468	3.0011	3.0157	2.9397	2.9536	2.9658

Table 2: MSE on average when $n = 50$ and $p = 8$

Methods	$\nu = 0.5$			$\nu = 2$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.5782	5.7413	5.9559	5.4846	5.6462	5.9164
GRRM	4.1042	4.1255	4.1032	3.9883	4.0352	4.0809
HK	3.6589	3.7132	3.7278	3.6518	3.6657	3.6779
N	3.3223	3.3305	3.3329	3.3151	3.3133	3.3085
TC	3.4009	3.4552	3.4698	3.3938	3.4077	3.4199
F	2.8755	2.9298	2.9444	2.8684	2.8823	2.8945
HSL	3.2879	3.3422	3.3568	3.2808	3.2947	3.3069
AH	3.2131	3.2674	3.282	3.206	3.2199	3.2321
D	3.1084	3.1192	3.1248	3.1	3.1082	3.1207
SB	3.4056	3.4599	3.4745	3.3985	3.4124	3.4246
SV1	3.2465	3.3008	3.3154	3.2394	3.2533	3.2655
SV2	3.1918	3.2461	3.2607	3.1847	3.1986	3.2108
M	3.2106	3.2644	3.279	3.203	3.2169	3.2292
AS	3.2715	3.3258	3.3404	3.2644	3.2783	3.2905

Table 3: MSE on average when $n = 100$ and $p = 4$

Methods	$\nu = 0.5$			$\nu = 2$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.1479	5.311	5.5256	5.0543	5.2159	5.4861
GRRM	3.6739	3.6952	3.6729	3.558	3.6049	3.6506
HK	3.2286	3.2829	3.2975	3.2215	3.2354	3.2476
N	2.892	2.9002	2.9026	2.8848	2.883	2.8782
TC	2.9706	3.0249	3.0395	2.9635	2.9774	2.9896
F	2.4452	2.4995	2.5141	2.4381	2.452	2.4642
HSL	2.8576	2.9119	2.9265	2.8505	2.8644	2.8766
AH	2.7828	2.8371	2.8517	2.7757	2.7896	2.8018
D	2.6781	2.6889	2.6945	2.6697	2.6779	2.6904
SB	2.9753	3.0296	3.0442	2.9682	2.9821	2.9943
SV1	2.8162	2.8705	2.8851	2.8091	2.823	2.8352
SV2	2.7615	2.8158	2.8304	2.7544	2.7683	2.7805
M	2.7803	2.8341	2.8487	2.7727	2.7866	2.7989
AS	2.8412	2.8955	2.9101	2.8341	2.848	2.8602

Table 4: MSE on average when $n = 100$ and $p = 8$

Methods	$\nu = 0.5$			$\nu = 2$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.2688	5.4319	5.6465	5.1752	5.3368	5.607
GRRM	3.7948	3.8161	3.7938	3.6789	3.7258	3.7715
HK	3.3495	3.4038	3.4184	3.3424	3.3563	3.3685
N	3.0129	3.0211	3.0235	3.0057	3.0039	2.9991

TC	3.0915	3.1458	3.1604	3.0844	3.0983	3.1105
F	2.5661	2.6204	2.635	2.559	2.5729	2.5851
HSL	2.9785	3.0328	3.0474	2.9714	2.9853	2.9975
AH	2.9037	2.958	2.9726	2.8966	2.9105	2.9227
D	2.799	2.8098	2.8154	2.7906	2.7988	2.8113
SB	3.0962	3.1505	3.1651	3.0891	3.103	3.1152
SV1	2.9371	2.9914	3.006	2.93	2.9439	2.9561
SV2	2.8824	2.9367	2.9513	2.8753	2.8892	2.9014
M	2.9012	2.955	2.9696	2.8936	2.9075	2.9198
AS	2.9621	3.0164	3.031	2.955	2.9689	2.9811

Table 5: MSE on average when $n = 200$ and $p = 4$

Methods	$\nu = 0.5$			$\nu = 2$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.01616	5.17926	5.39386	4.92256	5.08416	5.35436
GRRM	3.54216	3.56346	3.54116	3.42626	3.47316	3.51886
HK	3.09686	3.15116	3.16576	3.08976	3.10366	3.11586
N	2.76026	2.76846	2.77086	2.75306	2.75126	2.74646
TC	2.83886	2.89316	2.90776	2.83176	2.84566	2.85786
F	2.31346	2.36776	2.38236	2.30636	2.32026	2.33246
HSL	2.72586	2.78016	2.79476	2.71876	2.73266	2.74486
AH	2.65106	2.70536	2.71996	2.64396	2.65786	2.67006
D	2.54636	2.55716	2.56276	2.53796	2.54616	2.55866
SB	2.84356	2.89786	2.91246	2.83646	2.85036	2.86256
SV1	2.68446	2.73876	2.75336	2.67736	2.69126	2.70346
SV2	2.62976	2.68406	2.69866	2.62266	2.63656	2.64876
M	2.64856	2.70236	2.71696	2.64096	2.65486	2.66716
AS	2.70946	2.76376	2.77836	2.70236	2.71626	2.72846

Table 6: MSE on average when $n = 200$ and $p = 8$

Methods	$\nu = 0.5$			$\nu = 2$		
	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
MLE	5.11346	5.27656	5.49116	5.01986	5.18146	5.45166
GRRM	3.63946	3.66076	3.63846	3.52356	3.57046	3.61616
HK	3.19416	3.24846	3.26306	3.18706	3.20096	3.21316
N	2.85756	2.86576	2.86816	2.85036	2.84856	2.84376
TC	2.93616	2.99046	3.00506	2.92906	2.94296	2.95516
F	2.41076	2.46506	2.47966	2.40366	2.41756	2.42976
HSL	2.82316	2.87746	2.89206	2.81606	2.82996	2.84216
AH	2.74836	2.80266	2.81726	2.74126	2.75516	2.76736
D	2.64366	2.65446	2.66006	2.63526	2.64346	2.65596
SB	2.94086	2.99516	3.00976	2.93376	2.94766	2.95986
SV1	2.78176	2.83606	2.85066	2.77466	2.78856	2.80076
SV2	2.72706	2.78136	2.79596	2.71996	2.73386	2.74606
M	2.74586	2.79966	2.81426	2.73826	2.75216	2.76446
AS	2.80676	2.86106	2.87566	2.79966	2.81356	2.82576

4. Conclusions

In this paper, I modified the procedure to use the generalized ridge estimator which can solve the multicollinearity problem in the choice of beta regression model. Several techniques for estimating the K matrix have been adopted by various sources. The MJS prediction studies indicate that GCRM irrespective of the type estimating method of K matrix has always been more accurate than MLE and GRRM in terms of MSE has been adapted. According to Monte Carlo simulation studies, the GCRM estimator, regardless the type of estimating method of K matrix, has better performance than MLE and GRRM, in terms of MSE. Moreover, a simple real data example is also examined to also investigate the advantages of applying the GCRM estimator in the case of beta regression model. This was accompanied by an observation of the superiority of the GCRM estimator relative to the raw CM estimator in terms of the resulting MSE and a confirmation that the findings tally with Monte Carlo simulation outcomes.

References

[1] De Jong P, Heller GZ. Generalized linear models for insurance data. Vol. 10. Cambridge University Press Cambridge; 2008.

[2] Dunder E, Gumustekin S, Cengiz MA. Variable selection in gamma regression models via artificial bee colony algorithm. Journal of Applied Statistics. 2016:1-9. doi: 10.1080/02664763.2016.1254730.

[3] Malehi AS, Pourmotahari F, Angali KA. Statistical models for the analysis of skewed healthcare cost data: A simulation study. Health Economics Review. 2015;5:1-11. doi: 10.1186/s13561-015-0045-7. PubMed PMID: 26029491; PubMed Central PMCID: PMCPMC4442782.

[4] Al-Abood AM, Young DH. Improved deviance goodness of fit statistics for a gamma regression model. Communications in Statistics - Theory and Methods. 1986;15(6):1865-1874. doi: 10.1080/03610928608829223.

[5] Wasef Hattab M. A derivation of prediction intervals for gamma regression. Journal of Statistical Computation and Simulation. 2016;86(17):3512-3526. doi: 10.1080/00949655.2016.1169421.

[6] Asar Y, Genç A. New shrinkage parameters for the Liu-type logistic estimators. Communications in Statistics - Simulation and Computation. 2015;45(3):1094-1103. doi: 10.1080/03610918.2014.995815.

[7] Kurtoğlu F, Özkale MR. Liu estimation in generalized linear models: application on gamma distributed response variable. Statistical Papers. 2016;57(4):911-928. doi: 10.1007/s00362-016-0814-3.

[8] Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics. 1970;12(1):55-67.

[9] Batah FSM, Ramanathan TV, Gore SD. The efficiency of modefied jackknife and ridge type regression estimators - A comparison. Surveys in Mathematics and its Applications. 2008;3:111 – 122.

[10] Lord D, Geedipally SR, Guikema SD. Extension of the application of conway-maxwell-poisson models: analyzing traffic crash data exhibiting underdispersion. Risk Anal. 2010 Aug;30(8):1268-76. doi: 10.1111/j.1539-6924.2010.01417.x. PubMed PMID: 20412518.

[11] Lord D, Guikema SD, Geedipally SR. Application of the Conway-Maxwell-Poisson generalized linear model for analyzing motor vehicle crashes. Accid Anal Prev. 2008 May;40(3):1123-34. doi: 10.1016/j.aap.2007.12.003. PubMed PMID: 18460381.

[12] Santarelli MF, Della Latta D, Scipioni M, et al. A Conway-Maxwell-Poisson (CMP) model to address data dispersion on positron emission tomography. Comput Biol Med. 2016 Oct 1;77:90-101. doi: 10.1016/j.compbiomed.2016.08.006. PubMed PMID: 27522237.

[13] Chatla SB, Shmueli G. Efficient estimation of COM–Poisson regression and a generalized additive model. Computational Statistics & Data Analysis. 2018;121:71-88. doi: 10.1016/j.csda.2017.11.011.

[14] Shmueli G, Minka TP, Kadane JB, et al. A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. 2005;54(1):127-142.

[15] Guikema SD, Coffelt JP. A flexible count data regression model for risk analysis.

Risk Anal. 2008 Feb;28(1):213-23. doi: 10.1111/j.1539-6924.2008.01014.x. PubMed PMID: 18304118.

[16] Francis RA, Geedipally SR, Guikema SD, et al. Characterizing the performance of the conway-maxwell poisson generalized linear model. 2012;32(1):167-183.

[17] Hocking RR, Speed F, Lynn M. A class of biased estimators in linear regression. Technometrics. 1976;18(4):425-437.

[18] Nomura M. On the almost unbiased ridge regression estimator. Communications in Statistics-Simulation and Computation. 1988;17(3):729-743.

[19] Troskie C, Chalton D, editors. Detection of outliers in the presence of multicollinearity. Multidimensional statistical analysis and theory of random matrices, Proceedings of the Sixth Lukacs Symposium, eds. Gupta, AK and VL Girko; 1996.

[20] Firinguetti L. A generalized ridge regression estimator and its finite sample properties: A generalized ridge regression estimator. Communications in Statistics-Theory and Methods. 1999;28(5):1217-1229.

[21] Alkhamisi MA, Shukur G. A Monte Carlo study of recent ridge parameters. Communications in Statistics—Simulation and Computation®. 2007;36(3):535-547.

[22] Batah FSM, Ramanathan TV, Gore SD. The efficiency of modified jackknife and ridge type regression estimators: a comparison. Surveys in Mathematics & its Applications. 2008;3.

[23] Al-Hassan YM. Performance of a new ridge regression estimator. Journal of the Association of Arab Universities for Basic and Applied Sciences. 2010;9(1):23-26.

[24] Dorugade A, Kashid D. Alternative method for choosing ridge parameter for regression. Applied Mathematical Sciences. 2010;4(9):447-456.

[25] Månsson K, Shukur G, Golam Kibria B. A simulation study of some ridge regression estimators under different distributional assumptions. Communications in Statistics-Simulation and Computation. 2010;39(8):1639-1670.

[26] Dorugade A. New ridge parameters for ridge regression. Journal of the Association of Arab Universities for Basic and Applied Sciences. 2014;15(1):94-99.

[27] Asar Y, Karaibrahimoğlu A, Genç A. Modified ridge regression parameters: A comparative Monte Carlo study. Hacettepe Journal of Mathematics and Statistics. 2014;43(5):827-841.

[28] Bhat S, Raju V. A class of generalized ridge estimators. Communications in Statistics-Simulation and Computation. 2017;46(7):5105-5112.

[29] Uusipaikka E. Confidence intervals in generalized regression models. NW: Chapman & Hall/CRC Press; 2009.

[30] Amin M, Amanullah M, Aslam M, et al. Influence diagnostics in gamma ridge regression model. Journal of Statistical Computation and Simulation. 2018;89(3):536-556. doi: 10.1080/00949655.2018.1558226.

[31] Amin M, Amanullah M, Cordeiro GM. Influence diagnostics in the Gamma regression model with adjusted deviance residuals. Communications in Statistics - Simulation and Computation. 2016;46(9):6959-6973. doi: 10.1080/03610918.2016.1222420.

[32] Amin M, Qasim M, Amanullah M, et al. Performance of some ridge estimators for the gamma regression model. Statistical Papers. 2017;61(3):997-1026. doi: 10.1007/s00362-017-0971-z.

[33] Amin M, Qasim M, Yasin A, et al. Almost unbiased ridge estimator in the gamma regression model. Communications in Statistics - Simulation and Computation. 2020:1-21. doi: 10.1080/03610918.2020.1722837.

[34] Lukman AF, Ayinde K, Kibria BMG, et al. Modified ridge-type estimator for the gamma regression model. Communications in Statistics - Simulation and Computation. 2020:1-15. doi: 10.1080/03610918.2020.1752720.

[35] Mandal S, Arabi Belaghi R, Mahmoudi A, et al. Stein-type shrinkage estimators in gamma regression model with application to prostate cancer data. Stat Med. 2019 Sep 30;38(22):4310-4322. doi: 10.1002/sim.8297. PubMed PMID: 31317564.

[36] Qasim M, Amin M, Amanullah M. On the performance of some new Liu parameters for the gamma regression model. Journal of Statistical Computation and Simulation. 2018;88(16):3065-3080. doi: 10.1080/00949655.2018.1498502.

[37] Kibria BMG. Performance of some new ridge regression estimators.

Communications in Statistics - Simulation and Computation. 2003 2003/01/06;32(2):419-435. doi: 10.1081/SAC-120017499.

[38] Algamal ZY, Abonazel MR, Awwad FA, Eldin ET. Modified jackknife ridge estimator for the Conway-Maxwell-Poisson model. Scientific African. 2023; 19, doi:10.1016/j.sciaf.2023.e01543.