



# Improving Generalized Ridge Estimator for the Gamma Regression Model

A.Al-Saffar<sup>1</sup> and Z.Y. Algamal<sup>2</sup>

<sup>1</sup>Department of Statistics, College of Administration and Economics, University of Duhok, Duhok, Iraq, <sup>2</sup>Department of Statistics and Informatics, College Computer Science and Mathematics, University of Mosul, Mosul, Iraq.

## Article information

### Article history:

Received November 08, 2023

Revised :April 9,2024

Accepted: April 29, 2024

Available online June 21, 2024

### Keywords:

Multicollinearity

ridge estimator

gamma regression model

generalized ridge estimator

Monte Carlo simulation

### Correspondence:

A. Al-Saffar

[Avan.elias@uod.ac](mailto:Avan.elias@uod.ac)

## Abstract

It has been consistently proven that the ridge estimator is an effective shrinking strategy for reducing the effects of multicollinearity. An effective model to use when the response variable is positively skewed is the Gamma Regression Model (GRM). However, it is well known that the existence of multicollinearity can have a detrimental impact on the variance of the maximum likelihood estimator (MLE) of the gamma regression coefficients. The generalized ridge estimator is suggested in this study as a solution to the ridge estimator's limitation. The shrinkage matrix has been estimated using a number of different techniques. Our Monte Carlo simulation and actual data application findings indicate that the suggested estimator, regardless of the kind of estimating method of shrinkage matrix, is superior to the MLE and ridge estimator in terms of Mean Square Error (MSE). Additionally, compared to other methods, some shrinkage matrix estimation techniques can significantly enhance results.

DOI [10.3389/IQJOSS.2024.183251](https://doi.org/10.3389/IQJOSS.2024.183251), ©Authors, 2024, College of Computer Science and Mathematics University of Mosul.

This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The appearance of puberty in the female Many real data problems, such as automobile insurance claims, healthcare economics, and medical science, can be studied using the Gamma Regression Model (GRM) (1, 2, 3). A GRM is used particularly when a study's response variable is positively skewed or not normally distributed. As a result, gamma regression requires gamma distributions for the response variables (4, 5, 6).

The GRM presumes that there is no correlation between the regressors. However, in reality, this presumption frequently fails, which creates the multicollinearity issue. In the presence of multicollinearity, gamma regression coefficients are typically unstable with a large variance and poor statistical significance when estimated using the maximum likelihood (ML) approach (7, 8). To solve the multicollinearity issue, many solutions have been presented out. It has been frequently shown that the ridge regression approach (9) is a desirable replacement for the ML estimation method.

The following relationship is typically used in classical linear regression models:

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

where  $y$  is an  $n \times 1$  vector of response variable observations,  $X = (1, X_1, X_2, \dots, X_p)$  is a known design matrix of explanatory variables  $n \times (p + 1)$ ,  $\beta = (\beta_0, \beta_1, \dots, \beta_p)$  is a  $(p + 1) \times 1$  vector of unknown regression coefficients, and  $\varepsilon$  is an  $n \times 1$  vector of random errors with mean 0 and variance  $\sigma^2$ .

In order to decrease the high variance, the ridge regression shrinkage approach compresses all regression coefficients in the direction of zero (7, 10). The diagonal of  $X^T X$  is raised in a positive direction to achieve this. The ridge estimator has a lower mean squared error than the ML estimator due to its bias.

The ridge estimator in linear regression is defined as:

$$\hat{\beta}_{Ridge} = (X^T X + kI)^{-1} X^T y \quad (2)$$

With  $I$  as the identity matrix of size  $p \times p$  and  $k \geq 0$  as the ridge parameter (shrinkage parameter) which controls the shrinkage of  $\beta$  toward zero. A larger value of  $k$  yields greater shrinkage for the  $\hat{\beta}_{Ridge}$  estimator (9).

## 2. Statistical Methodology

### Gamma Ridge Regression Model (GRRM)

There are often positive skewed data used in studies in sociology, economics, and epidemiology, these kinds of data do not have any negative numbers, making the gamma distribution an ideal choice for these kinds of data (5). If  $y_i$  be the response variable and has a gamma distribution with nonnegative shape parameter  $v$  and nonnegative scale parameter  $\tau$ , i.e.  $y_i \sim \text{Gamma}(v, \tau)$ , then the probability density function is defined as (6, 11):

$$f(y_i) = \frac{\tau^v}{\Gamma(v)} (\tau y_i)^{v-1} e^{-\tau y_i}; \quad y_i \geq 0 \quad (3)$$

with  $E(y) = v/\tau = \theta$  and  $\text{var}(y) = v/\tau^2 = \theta^2/v$ . When the parameter is known, it is shown that the response variable's variance is proportional to the square of its mean.

In a GRM,  $\theta_i = \exp(X_i^T \beta)$  is expressed as a linear combination of regressors  $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ . The  $\theta_i$  is called the log link function is what gives the relationship between the predictors and the response variable its linear shape. This log like function is alternatively used rather than the canonical link function (reciprocal link function,  $\theta_i = -1/X_i^T \beta$ ) because it ensures that  $\theta_i > 0$ .

Using the Maximum likelihood technique is the most typical way to estimate the GRM coefficients. Considering that the observations are presumed to be independent and  $\tau = v/\theta$ , the log-likelihood function is given by:

$$\ell(\beta) = \sum_{i=1}^n \left\{ v \left[ -\frac{y_i}{X_i^T \beta} - X_i^T \beta \right] - \ln \Gamma v + v \ln(v y_i) - \ln(y_i) \right\} \quad (4)$$

The first derivative of Eq. (4) is then calculated and set to zero to get the ML estimator, as:

$$\frac{\partial \ell(\beta)}{\partial \beta} = \sum_{i=1}^n v \left[ \frac{y_i}{(X_i^T \beta)^2} - 1 \right] X_i = 0 \quad (5)$$

Unfortunately, the first derivative cannot be analytically calculated since Eq. (5) is nonlinear. The ML estimators of the gamma regression parameters may be obtained using either the iteratively weighted least squares (IWLS) technique or Newton-Raphson approach. In each iteration, the parameters are updated by:

$$\beta^{(r+1)} = \beta^{(r)} + I^{-1}(\beta^{(r)}) S(\beta^{(r)}) \quad (6)$$

Where  $S(\beta) = \partial \ell(\beta) / \partial \beta$  and  $I^{-1}(\beta) = (-E(\partial^2 \ell(\beta) / \partial \beta \partial \beta^T))^{-1}$ . The estimated coefficients final step is defined as

$$\hat{\beta}_{GRM} = (X^T \hat{W} X)^{-1} X^T \hat{W} \hat{u} \quad (7)$$

Where  $\hat{W} = \text{diag}(\hat{\theta}_i^2)$  and  $\hat{u}$  is a vector where  $i^{th}$  element equals to  $\hat{u}_i = \hat{\theta}_i + ((y_i - \hat{\theta}_i) / \hat{\theta}_i^2)$ . ML estimators are normally distributed with covariance matrices that are inverses of Hessian matrices.

$$\text{cov}(\hat{\beta}_{GRM}) = [-E(\frac{\partial^2 \ell(\beta)}{\partial \beta_i \partial \beta_k})]^{-1} = (X^T \hat{W} X)^{-1} \quad (8)$$

Eq. (7)'s mean squared error (MSE) can be calculated as follows:

$$\begin{aligned} MSE(\hat{\beta}_{GRM}) &= E(\hat{\beta}_{GRM} - \beta)^T (\hat{\beta}_{GRM} - \beta) \\ &= \text{tr } v^{-1} [(X^T \hat{W} X)^{-1}] \\ &= v^{-1} \sum_{j=1}^p \frac{1}{\lambda_j} \end{aligned} \quad (9)$$

where  $\lambda_j$  is the eigenvalue of the  $X^T \widehat{W} X$  matrix. The matrix  $X^T \widehat{W} X$  becomes ill-conditioned in the presence of multicollinearity, the ML estimator of the gamma regression parameters becomes unstable and has an excessive amount of variation. As a remedy, the gamma ridge regression model (GRRM) can be described as:

$$\begin{aligned}\hat{\beta}_{GRRM} &= (X^T \widehat{W} X + kI)^{-1} X^T \widehat{W} X \hat{\beta}_{GRM} \\ &= (X^T \widehat{W} X + kI)^{-1} X^T \widehat{W} \hat{u}\end{aligned}\quad (10)$$

where  $k \geq 0$ . A specific estimator from Eq. (10) with  $k = 0$  might be thought of as the ML estimator.

### Generalized Ridge Estimator

The generalized ridge estimator (GRE), differs from the generalized ridge regression model (GRR) in that it takes  $p$  values of  $k$  into account (9).

$$\hat{\beta}_{GRE} = (X^T X + K)^{-1} X^T y, \quad (11)$$

where  $K = \text{diag}(k_1, k_2, \dots, k_p)$ . Finding the optimal values of  $k_i$  while using GRE is advantageous because the MSE is smaller than when the ridge estimator and OLS are used.

The definition of the GRE for the gamma regression model (GRM) is:

$$\begin{aligned}\hat{\beta}_{GGRRM} &= (X^T \widehat{W} X + K)^{-1} X^T \widehat{W} X \hat{\beta}_{GRM} \\ &= (X^T \widehat{W} X + K)^{-1} X^T \widehat{W} \hat{u}.\end{aligned}\quad (12)$$

$K$  matrix selection must be carefully considered. Several approaches are modified to estimate  $K$  in this study, including (9, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22). These approaches are listed below, in order.

$$\hat{k}_{i(HK)} = \frac{\hat{v}}{\hat{\alpha}_i^2}, \quad (13)$$

where  $\hat{\alpha}_i$  is defined as the  $i^{th}$  element of  $\gamma \hat{\beta}_{GRM}$  and  $\gamma$  is the eigenvector of the  $X^T \widehat{W} X$  and the dispersion parameter,  $v$ , is estimated by  $\hat{v} = (1/(n-p)) \sum_{i=1}^n ((y_i - \hat{\theta}_i)^2 / \hat{\theta}_i^2)$  (5).

$$\hat{k}_{i(N)} = \frac{\hat{v}}{\hat{\alpha}_i^2} \{1 + [1 + \lambda_i (\hat{\alpha}_i^2 / \hat{v})^{1/2}]\} \quad (14)$$

$$\hat{k}_{i(TC)} = \frac{\lambda_i \hat{v}}{\lambda_i \hat{\alpha}_i^2 + \hat{v}} \quad (15)$$

$$\hat{k}_{i(F)} = \frac{\lambda_i \hat{v}}{\lambda_i \hat{\alpha}_i^2 + (n-p) \hat{v}} \quad (16)$$

$$\hat{k}_{i(HSL)} = \hat{v} \frac{\sum_{i=1}^p (\lambda_i \hat{\alpha}_i^2)^2}{(\sum_{i=1}^p (\lambda_i \hat{\alpha}_i^2))^2} \quad (17)$$

$$\hat{k}_{i(AH)} = \hat{v} \frac{\sum_{i=1}^p (\lambda_i \hat{\alpha}_i^2)^2}{(\sum_{i=1}^p (\lambda_i \hat{\alpha}_i^2))^2} + \frac{1}{\lambda_{max}} \quad (18)$$

$$\hat{k}_{i(D)} = \frac{\hat{v}}{\lambda_{max} \hat{\alpha}_i^2} \quad (19)$$

$$\hat{k}_{i(SB)} = \frac{\lambda_i \hat{v}}{\lambda_i \hat{\alpha}_i^2 + \hat{v}} + \frac{1}{\lambda_{max}} \quad (20)$$

$$\hat{k}_{i(SV1)} = \frac{p \hat{v}}{\hat{\alpha}_i^2} + \frac{1}{\lambda_{max} \hat{\alpha}_i^2} \quad (21)$$

$$\hat{k}_{i(SV2)} = \frac{p\hat{v}}{\hat{\alpha}_i^2} + \frac{1}{2(\sqrt{\lambda_{\max}/\lambda_{\min}})^2} \quad (22)$$

$$\hat{k}_{i(M)} = \frac{(n-p)\hat{v} + \lambda_{\max}\hat{\alpha}_i^2}{\lambda_{\max}\hat{\alpha}_i^2} \quad (23)$$

$$\hat{k}_{i(AS)} = \frac{\hat{v}}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i} \quad (24)$$

### 3. Modeling and Simulation

With the help of Monte Carlo simulations, the effectiveness of these approaches is examined using the GGRRM and different levels of multicollinearity.

#### The Design of Simulations

The GRM's response variable  $y_i \sim \text{Gamma}(\theta_i, \nu)$  for  $n$  observations is produced by (8, 11, 23, 24, 25, 26, 27, 28). where  $\theta = \exp(X^T \beta)$ ,  $\beta = (\beta_0, \beta_1, \dots, \beta_p)$  with  $\sum_{j=1}^p \beta_j^2 = 1$  and  $\beta_1 = \beta_2 = \dots = \beta_p$  [29], and  $\nu \in \{0.50, 2\}$ .

Explanatory variables  $X_i^T = (x_{i1}, x_{i2}, \dots, x_{in})$  have been generated from the following formula:

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

where  $\rho$  represents the correlation between explanatory variables and  $w_{ij}$ 's are independent standard normal pseudo-random numbers. Three exemplary sample size values 50, 100, and 200 are taken into consideration since the sample size directly influences prediction accuracy. Additionally, the number of explanatory factors is taken into account as  $p = 4$  and  $p = 8$  because doing so might result in an increase in the MSE. Further, three values of the pairwise correlation are taken into consideration with  $\rho = \{0.90, 0.95, 0.99\}$  since we are interested in the influence of multicollinearity, in which the degrees of correlation are deemed more essential. The produced data is repeated 1,000 times for a combination of these various values of  $n, p$ , and  $\rho$ , the averaged mean squared errors (MSE) is determined as follows:

$$MSE(\hat{\beta}) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\beta}_{GGRRM} - \beta)^T (\hat{\beta}_{GGRRM} - \beta). \quad (26)$$

#### 4. The Results of Simulations

There are six tables showing the averaged MSE for the combinations of  $n, \nu, p$ , and  $\rho$ . Throughout the table, the best MSE value is highlighted to emphasize its importance. The following are some possible observations:

- 1- GRRM frequently has a lower MSE than MLE.
- 2- GGRRM achieved less MSE than GRRM, regardless of the estimating method of the matrix K.
- 3- A comparison of the F method with other approaches revealed that the gamma generalized ridge estimator was significantly enhanced by Firinguetti (15) in Eq. (16)). HK and SB procedures consistently produced inadequate results, when compared with other approaches tested.
- 4- MSE values increase as the degree of correlation increases with respect to  $p$ , regardless of the values of  $n, \nu$  and  $p$ .
- 5- In terms of the number of explanatory variables, it is easy to see that there is a negative impact on MSE, where their values rise as  $p$  increases.
- 6- The MSE values decrease with increasing  $n$ , regardless of the values of  $\rho, \nu$ , or  $p$ .
- 7- As  $\nu$  increases, the MSE of all methods decreases for fixed  $n, p$ , and degree of multicollinearity.

Table 1: Average MSE values 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	3.2411	3.4042	3.6188	3.1475	3.3091	3.5793
GRRM	1.7671	1.7884	1.7661	1.6512	1.6981	1.7438
HK	1.3218	1.3761	1.3907	1.3147	1.3286	1.3408
N	0.9852	0.9934	0.9958	0.9780	0.9762	0.9714
TC	1.0638	1.1181	1.1327	1.0567	1.0706	1.0828
<b>F</b>	<b>0.5384</b>	<b>0.5927</b>	<b>0.6073</b>	<b>0.5313</b>	<b>0.5452</b>	<b>0.5574</b>
HSL	0.9508	1.0051	1.0197	0.9437	0.9576	0.9698
AH	0.876	0.9303	0.9449	0.8689	0.8828	0.895
D	0.7713	0.7821	0.7877	0.7629	0.7711	0.7836
SB	1.0685	1.1228	1.1374	1.0614	1.0753	1.0875
SV1	0.9094	0.9637	0.9783	0.9023	0.9162	0.9284
SV2	0.8547	0.909	0.9236	0.8476	0.8615	0.8737
M	0.8735	0.9273	0.9419	0.8659	0.8798	0.8921
AS	0.9344	0.9887	1.0033	0.9273	0.9412	0.9534

Table 2: Average MSE values 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	3.5658	3.7289	3.9435	3.4722	3.6338	3.904
GRRM	2.0918	2.1131	2.0908	1.9759	2.0228	2.0685
HK	1.6465	1.7008	1.7154	1.6394	1.6533	1.6655
N	1.3099	1.3181	1.3205	1.3027	1.3009	1.2961
TC	1.3885	1.4428	1.4574	1.3814	1.3953	1.4075
<b>F</b>	<b>0.8631</b>	<b>0.9174</b>	<b>0.932</b>	<b>0.856</b>	<b>0.8699</b>	<b>0.8821</b>
HSL	1.2755	1.3298	1.3444	1.2684	1.2823	1.2945
AH	1.2007	1.255	1.2696	1.1936	1.2075	1.2197
D	1.096	1.1068	1.1124	1.0876	1.0958	1.1083
SB	1.3932	1.4475	1.4621	1.3861	1.4	1.4122
SV1	1.2341	1.2884	1.303	1.227	1.2409	1.2531
SV2	1.1794	1.2337	1.2483	1.1723	1.1862	1.1984
M	1.1982	1.252	1.2666	1.1906	1.2045	1.2168
AS	1.2591	1.3134	1.328	1.252	1.2659	1.2781

Table 3: Average MSE values 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	3.1355	3.2986	3.5132	3.0419	3.2035	3.4737
GRRM	1.6615	1.6828	1.6605	1.5456	1.5925	1.6382
HK	1.2162	1.2705	1.2851	1.2091	1.223	1.2352
N	0.8796	0.8878	0.8902	0.8724	0.8706	0.8658
TC	0.9582	1.0125	1.0271	0.9511	0.965	0.9772
<b>F</b>	<b>0.4328</b>	<b>0.4871</b>	<b>0.5017</b>	<b>0.4257</b>	<b>0.4396</b>	<b>0.4518</b>
HSL	0.8452	0.8995	0.9141	0.8381	0.852	0.8642

AH	0.7704	0.8247	0.8393	0.7633	0.7772	0.7894
D	0.6657	0.6765	0.6821	0.6573	0.6655	0.678
SB	0.9629	1.0172	1.0318	0.9558	0.9697	0.9819
SV1	0.8038	0.8581	0.8727	0.7967	0.8106	0.8228
SV2	0.7491	0.8034	0.818	0.742	0.7559	0.7681
M	0.7679	0.8217	0.8363	0.7603	0.7742	0.7865
AS	0.8288	0.8831	0.8977	0.8217	0.8356	0.8478

Table 4: Average MSE values 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	3.2564	3.4195	3.6341	3.1628	3.3244	3.5946
GRRM	1.7824	1.8037	1.7814	1.6665	1.7134	1.7591
HK	1.3371	1.3914	1.406	1.33	1.3439	1.3561
N	1.0005	1.0087	1.0111	0.9933	0.9915	0.9867
TC	1.0791	1.1334	1.148	1.072	1.0859	1.0981
<b>F</b>	<b>0.5537</b>	<b>0.608</b>	<b>0.6226</b>	<b>0.5466</b>	<b>0.5605</b>	<b>0.5727</b>
HSL	0.9661	1.0204	1.035	0.959	0.9729	0.9851
AH	0.8913	0.9456	0.9602	0.8842	0.8981	0.9103
D	0.7866	0.7974	0.803	0.7782	0.7864	0.7989
SB	1.0838	1.1381	1.1527	1.0767	1.0906	1.1028
SV1	0.9247	0.979	0.9936	0.9176	0.9315	0.9437
SV2	0.87	0.9243	0.9389	0.8629	0.8768	0.889
M	0.8888	0.9426	0.9572	0.8812	0.8951	0.9074
AS	0.9497	1.004	1.0186	0.9426	0.9565	0.9687

Table 5: Average MSE values 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	3.00376	3.16686	3.38146	2.91016	3.07176	3.34196
GRRM	1.52976	1.55106	1.52876	1.41386	1.46076	1.50646
HK	1.08446	1.13876	1.15336	1.07736	1.09126	1.10346
N	0.74786	0.75606	0.75846	0.74066	0.73886	0.73406
TC	0.82646	0.88076	0.89536	0.81936	0.83326	0.84546
<b>F</b>	<b>0.30106</b>	<b>0.35536</b>	<b>0.36996</b>	<b>0.29396</b>	<b>0.30786</b>	<b>0.32006</b>
HSL	0.71346	0.76776	0.78236	0.70636	0.72026	0.73246
AH	0.63866	0.69296	0.70756	0.63156	0.64546	0.65766
D	0.53396	0.54476	0.55036	0.52556	0.53376	0.54626
SB	0.83116	0.88546	0.90006	0.82406	0.83796	0.85016
SV1	0.67206	0.72636	0.74096	0.66496	0.67886	0.69106
SV2	0.61736	0.67166	0.68626	0.61026	0.62416	0.63636
M	0.63616	0.68996	0.70456	0.62856	0.64246	0.65476
AS	0.69706	0.75136	0.76596	0.68996	0.70386	0.71606

Table 6: Average MSE values 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	3.10106	3.26416	3.47876	3.00746	3.16906	3.43926
GRRM	1.62706	1.64836	1.62606	1.51116	1.55806	1.60376
HK	1.18176	1.23606	1.25066	1.17466	1.18856	1.20076
N	0.84516	0.85336	0.85576	0.83796	0.83616	0.83136
TC	0.92376	0.97806	0.99266	0.91666	0.93056	0.94276
<b>F</b>	<b>0.39836</b>	<b>0.45266</b>	<b>0.46726</b>	<b>0.39126</b>	<b>0.40516</b>	<b>0.41736</b>
HSL	0.81076	0.86506	0.87966	0.80366	0.81756	0.82976
AH	0.73596	0.79026	0.80486	0.72886	0.74276	0.75496
D	0.63126	0.64206	0.64766	0.62286	0.63106	0.64356
SB	0.92846	0.98276	0.99736	0.92136	0.93526	0.94746
SV1	0.76936	0.82366	0.83826	0.76226	0.77616	0.78836
SV2	0.71466	0.76896	0.78356	0.70756	0.72146	0.73366
M	0.73346	0.78726	0.80186	0.72586	0.73976	0.75206
AS	0.79436	0.84866	0.86326	0.78726	0.80116	0.81336

#### 4. Application Of Real Data

Here, we offer a chemical dataset with  $(n, p) = (212, 10)$ , where  $n$  denotes the quantity of antifungal drugs, to illustrate the applicability of the GRRM estimator in practical applications. pMIC (the logarithm of reciprocal of MIC, where MIC is the lowest inhibitory concentration against *C. albicans* in mM/L) was used to quantify the antibacterial activity. As explanatory variables, molecular descriptors are represented by the integer  $p$  (29, 30). In chemometrics, the quantitative structure-activity relationship (QSAR) investigation has gained significant attention. The fundamental idea behind QSAR is to simulate various biological functions across a group of chemical substances in terms of their structural characteristics. Regression modeling is therefore one of the most crucial techniques for building the QSAR model. Table 7 lists the explanatory variables that were employed. Every variable is a number.

The Chi-square test is performed first to determine if the answer variable is part of the gamma distribution. The test yielded a result of 10.0286 and a p-value of 0.9117. The gamma distribution closely matches this response variable, with an estimated dispersion parameter of 0.0153. Using the predicted dispersion parameter of 0.0153 and log link function to construct the gamma regression model, the test for multicollinearity, the eigenvalues of the matrix  $X^T \hat{W} X$  are obtained as:  $1.97 \times 10^9, 3.74 \times 10^6, 1.21 \times 10^4, 1.34 \times 10^3, 1.22 \times 10^3, 1.07 \times 10^3, 4.63 \times 10^2, 2.08 \times 10^1, 10.68$ , and 1.57. The determined condition number  $CN = \sqrt{\lambda_{\max}/\lambda_{\min}}$  of the data is 35422.83 demonstrating the existence of the serious multicollinearity problem.

Table 8 lists the estimated MSE values for the MLE, GRRM, and GRRM estimators using various estimating matrices. Table 8 makes it abundantly evident that the F approach effectively reduces the value of the calculated coefficients. The MSE has also been considerably decreased in favor of the F method. It is clear that the MSE of the F technique was around 64.97%, 60.63%, 59.33%, 47.98%, 44.34%, 46.79%, 45.37%, 42.43%, 48.26%, 65.96%, 44.34%, 44.47%, and 45.49% lower than that of MLE, GRRM, HK, N, TC, HSL, AH, D, SB, SV1, SV2, M, and AS estimators, respectively.

Table 7: Description of the used explanatory variables

Variable name's	Description
SpMax3_Bh(s)	largest eigenvalue n. 3 of Burden matrix weighted by I-state
P_VSA_e_3	P_VSA-like on Sanderson electronegativity, bin 3
IC3	Information Content index (neighborhood symmetry of 3-order)
Mor21e	signal 21 / weighted by Sanderson electronegativity
MATS2s	Moran autocorrelation of lag 2 weighted by I-state
GATS4p	Geary autocorrelation of lag 4 weighted by polarizability
SpMax8_Bh(p)	largest eigenvalue n. 8 of Burden matrix weighted by polarizability
ATS8v	Broto-Moreau autocorrelation of lag 8 (log function) weighted by van der Waals

	volume
MATS7v	Moran autocorrelation of lag 7 weighted by van der Waals volume
TDB08m	3D Topological distance based descriptors - lag 8 weighted by mass

Table 8: The estimated MSE values for the real data application

Methods	MSE
MLE	4.3291
GRRM	3.8507
HK	3.3008
N	2.9147
TC	2.9348
F	<b>1.5161</b>
HSL	2.8492
AH	2.7751
D	2.6335
SB	2.9301
SV1	2.7571
SV2	2.7238
M	2.7304
AS	2.7816

## 5. Conclusion

In this study, a generalized ridge estimator was suggested as a solution to the gamma regression model's multicollinearity issue. The  $\mathbf{K}$  matrix has been estimated using a variety of techniques. According to Monte Carlo simulation tests, the GGRRM estimator performs better than MLE and GRRM in terms of MSE regardless of the kind of estimating method used for the  $\mathbf{K}$  matrix. To further demonstrate the advantages of utilizing the GGRRM estimator in the context of gamma regression models, a real data application is also taken into consideration. It was determined that the GGRRM estimator is superior based on the resultant MSE, and it was further demonstrated that the outcomes are compatible with those of Monte Carlo simulations.

## Conflict of Interest

The author has no conflict of interest.

## 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](https://doi.org/10.1080/02664763.2016.1254730).
3. Malehi AS, Pourmohammadi 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](https://doi.org/10.1186/s13561-015-0045-7). PubMed PMID: 26029491; PubMed Central PMCID: PMC4442782.
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](https://doi.org/10.1080/03610928608829223).
5. Uusipaikka E. Confidence intervals in generalized regression models. NW: Chapman & Hall/CRC Press; 2009.
6. 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](https://doi.org/10.1080/00949655.2016.1169421).
7. 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](https://doi.org/10.1080/03610918.2014.995815).
8. 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](https://doi.org/10.1007/s00362-016-0814-3).



9. Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*. 1970;12(1):55-67.
10. Batah FSM, Ramanathan TV, Gore SD. The efficiency of modified jackknife and ridge type regression estimators - A comparison. *Surveys in Mathematics and its Applications*. 2008;3:111 – 122.
11. 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](https://doi.org/10.1080/03610918.2020.1752720).
12. Hocking RR, Speed F, Lynn M. A class of biased estimators in linear regression. *Technometrics*. 1976;18(4):425-437.
13. Nomura M. On the almost unbiased ridge regression estimator. *Communications in Statistics-Simulation and Computation*. 1988;17(3):729-743.
14. 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.
15. 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.
16. Alkhamisi MA, Shukur G. A Monte Carlo study of recent ridge parameters. *Communications in Statistics—Simulation and Computation®*. 2007;36(3):535-547.
17. 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.
18. Dorugade A, Kashid D. Alternative method for choosing ridge parameter for regression. *Applied Mathematical Sciences*. 2010;4(9):447-456.
19. 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.
20. 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.
21. Asar Y, Karabrahimoğlu A, Genç A. Modified ridge regression parameters: A comparative Monte Carlo study. *Hacettepe Journal of Mathematics and Statistics*. 2014;43(5):827-841.
22. Bhat S, Raju V. A class of generalized ridge estimators. *Communications in Statistics-Simulation and Computation*. 2017;46(7):5105-5112.
23. 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](https://doi.org/10.1080/03610918.2016.1222420).
24. 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](https://doi.org/10.1007/s00362-017-0971-z).
25. 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](https://doi.org/10.1080/00949655.2018.1558226).
26. 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](https://doi.org/10.1080/03610918.2020.1722837).
27. 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](https://doi.org/10.1002/sim.8297). PubMed PMID: 31317564.
28. 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](https://doi.org/10.1080/00949655.2018.1498502).
29. Al-Fakih AM, Algarni ZY, Lee MH, et al. QSAR classification model for diverse series of antifungal agents based on improved binary differential search algorithm. *SAR QSAR Environ Res*. 2019 Feb;30(2):131-143. DOI: [10.1080/1062936X.2019.1568298](https://doi.org/10.1080/1062936X.2019.1568298). PubMed PMID: 30734580.
30. Alharthi AM, Lee MH, Algarni ZY, et al. Quantitative structure-activity relationship model for classifying the diverse series of antifungal agents using ratio weighted penalized logistic regression. *SAR QSAR Environ Res*. 2020 Jul 6;1-13. DOI: [10.1080/1062936X.2020.1782467](https://doi.org/10.1080/1062936X.2020.1782467). PubMed PMID: 32628042.

## تحسين مقدر ريدج المعمم لنموذج انحدار كاما

آفان الصفار<sup>1</sup> و زكريا يحيى الجمال<sup>2</sup>

<sup>1</sup>قسم الإحصاء، كلية الإدارة والاقتصاد، جامعة دهوك، دهوك، إقليم كردستان، العراق. <sup>2</sup>قسم الإحصاء والمعلوماتية، كلية علوم الحاسوب والرياضيات، جامعة الموصل، الموصل، العراق.

**الخلاصة:** لقد ثبت باستمرار أن مقدر ريدج هو استراتيجية تقليص فعالة لتقليل تأثيرات مشكلة التعدد الخطي. فعندما يكون متغير الاستجابة منحرفاً بشكل إيجابي، عندئذ يتم استخدام النموذج المتداول والمعروف والذي هو نموذج انحدار كاما. مع العلم بأن ظهور مشكلة التعدد الخطي يمكن أن يكون له تأثير غير محبب على تباين مقدر الإمكان الأعظم لمعاملات انحدار كاما. ولهذا تم اقتراح مقدر ريدج المعمم في هذه الورقة البحثية كحل لمحدودية مقدر ريدج. تم تقدير مصفوفة الانكماش باستخدام عدد من التقنيات المختلفة. تشير نتائج محاكاة Monti Carlo وتطبيق البيانات الفعلية إلى أن المقدر المقترح، بغض النظر عن نوع طريقة التقدير المستخدمة لمصفوفة الانكماش، يتفوق على مقدر الإمكان الأعظم ومقدر ريدج من حيث متوسط مربعات الخطأ. بالإضافة إلى ذلك، فإن بعض تقنيات تقدير مصفوفة الانكماش يمكن أن تعزز النتائج بشكل كبير مقارنة بطرائق التقدير الأخرى. الكلمات المفتاحية: تعدد العلاقة الخطية، مقدر الحرف، نموذج انحدار كاما، مقدر الحرف المعمم، محاكاة مونت كارلو.