STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. 10, June 2022, pp 750–761.
Published online in International Academic Press (www.IAPress.org)

A New Two-parameter Estimator for the Gamma Regression Model

Yasin Asar [1], Zakariya Yahya Algamal [2,*]

[1]Department of Mathematics and Computer Sciences, Necmettin Erbakan University, Turkey
[2]Department of Statistics and Informatics, University of Mosul, Iraq

Abstract In this paper, we propose a new two-parameter biased estimator in gamma regression models when there is collinearity among the regressors. We investigate the mean squared error properties of the newly proposed estimator. Moreover, we provide some theorems to compare the new estimators to the existing ones. We conduct a Monte Carlo simulation study to compare the estimators under different designs of collinearity in the sense of mean squared error. Moreover, we provide a real data application to show the usefulness of the new estimator. The simulations and real data results show that the proposed estimator beats other competitor estimators.

Keywords Gamma regression model, two-parameter estimator, collinearity, Monte Carlo simulation

AMS 2010 subject classifications 62J05, 62J07

DOI: 10.19139/soic-2310-5070-822

1. Introduction

Gamma regression model is one of the widely applied models for studying several real data problems, such as medical science, health-care economics, and automobile insurance claims,(see [12], [13] and [23]). The gamma regression model is used when the values of the response variable under the study is positively skewed following gamma distribution ([1], [35]).

As in linear regression model, in gamma regression model, it is assumed that there is no correlation among the explanatory variables. In practice, however, this assumption often not holds, which leads to the problem of multicollinearity. In the presence of multicollinearity, the maximum likelihood (ML) estimator of the gamma regression coefficients are usually become unstable with high variance, and therefore low statistical significance [13] and [23].

Several remedial methods have been proposed to overcome the problem of multicollinearity. The ridge regression method was proposed by [15] and has been consistently demonstrated to be an attractive and alternative to the ML estimation method. The ridge estimator was considered in the generalized linear models (GLM) by [33]. Moreover, [33] and [32] considered the ridge estimator in logistic regression. [25] and [26] were also adapted the ridge estimator in Poisson regression and negative binomial regression models, respectively. The well-known Liu estimator [20] has been recently generalized to GLM and application of gamma distributed response variable was demonstrated by [19]. We also refer to the following papers for Liu regression: [36], [27], [34]. Another solution to the collinearity problem is the two-parameter estimator [28]. This estimator was also well studied in the literature and generalized to some models which are members of GLMs ([16], [11]). On the other hand, several estimators have been proposed for dealing with the issue of multicollinearity in gamma regression model [24, 29, 8, 21, 9, 6, 7, 3].

*Correspondence to: Zakariya Yahya Algamal (Email:zakariya.algamal@uomosul.edu.iq). Department of Statistics and Informatics, University of Mosul, Mosul, Iraq .

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 2022 International Academic Press

The organization of the paper is as follows: In Section 2, we propose the gamma two-parameter estimator and investigate its mean squared error (MSE) properties. In Section 3, we conduct a Monte Carlo simulation study to evaluate the MSE performances of used estimators in the presence of multicollinearity. Moreover, a real data application is provided to illustrate the benefits of the new estimation technique in Section 4. Finally, a conclusion is given in Section 5.

2. Theory and Method

Let $Y_1, Y_2, ..., Y_n$ be independent random variables and $y_1, y_2, ...y_n$ be the corresponding observations from the gamma distribution having the following probability density function

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

where v is the non-negative shape parameter and τ is the scale parameter such that $E(Y_i) = v\tau = \theta_i$ which is also known as the canonical parameter and $Var(Y_i) = v\tau^2 = 1/(v\theta_i^2)$, $\theta_i = exp\left(x_i^\top \beta\right)$ where $x_i = (x_{i1}, x_{i2}, ..., x_{ip})^\top$, $i = 1, 2, ..., n$ and $j = 1, 2, ..., p$ where n is the sample size and p is the number of explanatory variables $(n > p)$. Generally, the maximum likelihood estimation is used to obtain the parameters. To do so, the following log-likelihood function should be maximized with respect to β

$$l(\beta) = \sum_{i=1}^{n} \left[(v-1)\log(y) - y/\tau - v\log(\tau) - \log(\Gamma(v))\right]. \tag{2}$$

Since the obtained equations are nonlinear in β, we should use some iterative methods to get the solutions. Therefore, by using the Fisher Scoring method, the following iterations can be defined

$$\beta^{t+1} = \beta^t - \left\{E\left[H_l(\beta)\right]_{\beta=\beta^t}\right\}^{-1} \left[\frac{\partial l(\beta)}{\beta}\right]_{\beta=\beta^t} \tag{3}$$

where $H_l(\beta) = -\frac{1}{\phi} X^\top W X$ is the Hessian matrix such that $\phi = 1/v$ is the dispersion parameter and

$$\frac{\partial l(\beta)}{\beta} = \phi \sum_{i=1}^{n} \left[\frac{y_i}{(x_i^\top \beta)^2} - 1\right] x_i. \tag{4}$$

Therefore, Equation (3) can be written as

$$\beta^{t+1} = \beta^t - \left\{\left(X^\top \widehat{W} X\right)^{-1} X^\top \widehat{W} \widehat{z}\right\}_{\beta=\beta^t} \tag{5}$$

where $\widehat{W} = \text{diag}\left(\theta_i^2\right)$ and the ith element of the vector $\widehat{z}$ becomes $\widehat{z}_i = \widehat{\theta}_i + \frac{y_i - \widehat{\theta}_i}{\widehat{\theta}_i^2}$. This iterative process continues until the successive estimates converges to, say, $\widehat{\beta}_{\text{MLE}}$, then we obtain $\widehat{\beta}_{\text{MLE}} = \left(X^\top \widehat{W} X\right)^{-1} X^\top \widehat{W} \widehat{z}$ where $\widehat{W}$ and $\widehat{z}$ are computed at the final iteration.

It is well-known that the covariance matrix of $\widehat{\beta}_{\text{MLE}}$, $\text{Cov}(\widehat{\beta}_{\text{MLE}}) = \phi \left(X^\top \widehat{W} X\right)^{-1}$, may be ill-conditioned so that the variance of the regression coefficients is inflated (see [33], [22]). The mean squared

error (MSE) of MLE is given by

$$\begin{aligned} \mathrm{MSE}\left(\widehat{\beta}_{\mathrm{MLE}}\right) &= \mathrm{E}\left(\widehat{\beta}_{\mathrm{MLE}}-\beta\right)^{\top}\left(\widehat{\beta}_{\mathrm{MLE}}-\beta\right) \\ &= \mathrm{tr}\left[\phi\left(X^{\top}\widehat{W}X\right)^{-1}\right] \\ &= \phi\sum_{j=1}^{p}\frac{1}{\lambda_j} \end{aligned} \tag{6}$$

where λ_j is the jth eigenvalue of the matrix $C = X^{\top}\widehat{W}X$ and tr(.) is the trace of a matrix. Moreover, the eigenvalue decomposition of the matrix C is also considered as follows: $C = Q\Lambda Q^{\top}$ such that Q is the orthogonal matrix consisting of the eigenvectors corresponding to the eigenvalues of C such that $\Lambda = \mathrm{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_p\right)$. It is easy to see that if one or some of the eigenvalues are close to zero, then the MSE of MLE becomes inflated and thus the regression coefficients are affected negatively from this situation.

2.1. Gamma Ridge Estimator

The well-known ridge estimator was proposed by [33] in the generalized linear models. [2] adapted the ridge estimator to the gamma regression models. The author defined the gamma ridge estimator (GRE) as follows

$$\begin{aligned} \widehat{\beta}_{\mathrm{k}} &= \left(C+kI\right)^{-1}X^{\top}\widehat{W}\widehat{z} \\ b_{GRE} &= C_k^{-1}C\widehat{\beta}_{\mathrm{MLE}} \end{aligned} \tag{7}$$

where $k > 0$, $C = X^{\top}\widehat{W}X$ and $C_k = (C + kI)$. The covariance matrix and bias vector of GRE can be obtained respectively by

$$\mathrm{Cov}\left(\widehat{\beta}_{\mathrm{k}}\right) = \phi C_k^{-1}CC_k^{-1} \tag{8}$$

$$b_{GRE} = \mathrm{bias}\left(\widehat{\beta}_{\mathrm{k}}\right) = -kC_k\beta \tag{9}$$

Thus, matrix MSE (MMSE) and MSE of GRE are obtained as

$$\begin{aligned} \mathrm{MMSE}\left(\widehat{\beta}_{\mathrm{k}}\right) &= \mathrm{Cov}\left(\widehat{\beta}_{\mathrm{k}}\right) + b_{GRE}b_{GRE}^{\top} \\ &= \phi C_k^{-1}CC_k^{-1} + k^2C_k^{-1}\beta\beta^{\top}C_k^{-1} \end{aligned} \tag{10}$$

$$\begin{aligned} \mathrm{MSE}\left(\widehat{\beta}_{\mathrm{k}}\right) &= \mathrm{tr}\left[\mathrm{MMSE}\left(\widehat{\beta}_{\mathrm{k}}\right)\right] \\ &= \phi\sum_{j=1}^{p}\frac{\lambda_j}{\left(\lambda_j+k\right)^2} + k^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k\right)^2} \end{aligned} \tag{11}$$

where $\alpha = Q^{\top}\beta$.

2.2. Gamma Liu Estimator

Another popular estimator which is known as Liu estimator has been adopted to the generalized linear models by [19] and the authors considered the gamma dependent variable to study the performance of Liu gamma estimator via Monte Carlo simulation study and real data application. The gamma Liu estimator (GLE) is defined as

$$\widehat{\beta}_{\mathrm{d}} = F_d\widehat{\beta}_{\mathrm{MLE}} \tag{12}$$

where $F_d = (C+I)^{-1}(C+dI)$ and $0 < d < 1$. The covariance matrix and bias vector of GLE can be obtained respectively by

$$\begin{aligned} \mathrm{Cov}\left(\widehat{\beta}_{\mathrm{d}}\right) &= \phi F_d C^{-1} F_d^{\top} && (13)\\ b_{GLE} &= \mathrm{bias}\left(\widehat{\beta}_{\mathrm{d}}\right) = -(1-d)(C+I)^{-1}\beta. && (14) \end{aligned}$$

Using the covariance and bias of GLE, one can obtain the following MMSE and MSE functions respectively

$$\begin{aligned} \mathrm{MMSE}\left(\widehat{\beta}_{\mathrm{d}}\right) &= \mathrm{Cov}\left(\widehat{\beta}_{\mathrm{d}}\right) + b_{GLE} b_{GLE}^{\top} \\ &= \phi F_d C^{-1} F_d^{\top} + (1-d)^2 (C+I)^{-1} \beta\beta^{\top} (C+I)^{-1} && (15)\\ \mathrm{MSE}\left(\widehat{\beta}_{\mathrm{d}}\right) &= \mathrm{tr}\left[\mathrm{MMSE}\left(\widehat{\beta}_{\mathrm{d}}\right)\right] \\ &= \phi \sum_{j=1}^{p} \frac{(\lambda_j + d)^2}{\lambda_j (\lambda_j+1)^2} + (1-d)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\lambda_j+1)^2}. && (16) \end{aligned}$$

2.3. Gamma Two-parameter Estimator

In this paper, we propose to adopt the estimator defined by [28] in the gamma regression model, we call this estimator as gamma two-parameter estimator (GTPE) which is defined as follows

$$\begin{aligned} \widehat{\beta}_{(\mathrm{k,d})} &= {C_k}^{-1} C_{kd} \widehat{\beta}_{\mathrm{MLE}} \\ &= F_{kd} \widehat{\beta}_{\mathrm{MLE}} && (17) \end{aligned}$$

where $0 < d < 1$, $k > 0$, $C_k = C + kI$, $C_{kd} = C + kdI$ and $F_{kd} = {C_k}^{-1} C_{kd}$. We obtain the covariance matrix and bias vector of GTPE as

$$\begin{aligned} \mathrm{Cov}\left(\widehat{\beta}_{(\mathrm{k,d})}\right) &= \phi F_{kd} C^{-1} F_{kd}^{\top} && (18)\\ b_{GLTE} &= \mathrm{bias}\left(\widehat{\beta}_{(\mathrm{k,d})}\right) = k(d-1){C_k}^{-1}\beta. && (19) \end{aligned}$$

Therefore, MMSE and MSE functions of GTPE are respectively computed as

$$\begin{aligned} \mathrm{MMSE}\left(\widehat{\beta}_{(\mathrm{k,d})}\right) &= \mathrm{Cov}\left(\widehat{\beta}_{(\mathrm{k,d})}\right) + b_{GLTE} b_{GLTE}^{\top} \\ &= \phi F_{kd} C^{-1} F_{kd}^{\top} + k^2 (d-1)^2 {C_k}^{-1} \beta\beta^{\top} {C_k}^{-1} && (20)\\ \mathrm{MSE}\left(\widehat{\beta}_{(\mathrm{k,d})}\right) &= \mathrm{tr}\left[\mathrm{MMSE}\left(\widehat{\beta}_{(\mathrm{k,d})}\right)\right] \\ &= \phi \sum_{j=1}^{p} \frac{(\lambda_j + kd)^2}{\lambda_j (\lambda_j+k)^2} + k^2 (d-1)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{(\lambda_j+k)^2}. && (21) \end{aligned}$$

2.4. Theoretical Comparisons Between Estimators

In this subsection, we provide some theorems comparing the MSE and MMSE functions of the listed estimators. To do so, we consider the MSE and MMSE differences and investigate under which conditions they are positive definite. If a matrix A is positive definite, then we write $A > 0$.

Now, we will present three lemmas and use them to prove some of the theorems given in this section.

Lemma 1
[14] Suppose that M be a positive definite matrix, namely $M > 0$, α be some vector, then $M - \alpha\alpha^{\top} \geq 0$ if and only if $\alpha^{\top} M^{-1} \alpha \leq 1$.

Lemma 2
[31] Let $M > 0$, $N > 0$, then $M > N$, if and only if $\eta_{\max}\left(NM^{-1}\right) < 1$, where $\eta_{\max}(A)$ is the maximum eigenvalue of some matrix A.

Lemma 3
[31] Let $\widehat{\beta}_j = A_j y, j = 1, 2$ be two competing estimator of β. Assume that $\Delta = \mathrm{Cov}\left(\widehat{\beta}_1\right) - \mathrm{Cov}\left(\widehat{\beta}_2\right) > 0$, then $MSEM\left(\widehat{\beta}_1\right) - MSEM\left(\widehat{\beta}_2\right) > 0$ if and only if $u_2^{\top}\left(\Delta + u_1 u_1^{\top}\right)^{-1} u_2 \leq 1$, where u_j denotes the bias of $\widehat{\beta}_j$.

In the next theorem, we compare MLE and GTPE using the MMSE functions.

Theorem 1
When $\eta_{max}\left(F_{kd}C^{-1}F_{kd}^{\top}C\right) < 1$ the new estimator GTPE is superior to MLE in the sense of MMSE if and only if $b_{GLTE}^{\top}D_1^{-1}b_{GLTE} < 1$, where $D_1 = \mathrm{Cov}\left(\widehat{\beta}_{\mathrm{MLE}}\right) - \mathrm{Cov}\left(\widehat{\beta}_{(\mathrm{k,d})}\right)$.

Proof
Let us consider the following MMSE difference

$$\begin{aligned}\Delta_1 &= \mathrm{MMSE}\left(\widehat{\beta}_{\mathrm{MLE}}\right) - \mathrm{MMSE}\left(\widehat{\beta}_{(\mathrm{k,d})}\right)\\ &= \phi\left(C^{-1} - F_{kd}C^{-1}F_{kd}\right) - b_{GLTE}b_{GLTE}^{\top}\end{aligned} \tag{22}$$

Since the matrices C^{-1} and $F_{kd}C^{-1}F_{kd}^{\top}$ are positive definite then by Lemma 2, $D_1 = C^{-1} - F_{kd}C^{-1}F_{kd} > 0$ if $\eta_{max}\left(F_{kd}C^{-1}F_{kd}C\right) < 1$. Then, by Lemma 1, $\Delta_1 > 0$ if and only if $b_{GLTE}^{\top}D_1^{-1}b_{GLTE} < 1$. Thus, the proof is finished. □

Theorem 2
The new estimator GLTE is superior to MLE in the sense of MSE if $\min\left\{\frac{2\phi\lambda_{\mathrm{j}}}{\lambda_{\mathrm{j}}\alpha_{\mathrm{j}}^2 - \phi}\right\}_{\mathrm{j=1}}^{\mathrm{p}} > \mathrm{k}(1-\mathrm{d})$ where $k > 0$ and $0 < d < 1$.

Proof
The MSE differences of MLE and GTPE is given as

$$\begin{aligned}\mathrm{MSE}\left(\widehat{\beta}_{\mathrm{MLE}}\right) - \mathrm{MSE}\left(\widehat{\beta}_{(\mathrm{k,d})}\right) &= \phi\sum_{j=1}^{p}\frac{1}{\lambda_j} - \left\{\phi\sum_{j=1}^{p}\frac{(\lambda_j + kd)^2}{\lambda_j(\lambda_j+k)^2} + k^2(d-1)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j+k)^2}\right\}\\ &= \sum_{j=1}^{p}\frac{1}{\lambda_j(\lambda_j+k)^2}\left\{\phi\left[(\lambda_j+k)^2 - (\lambda_j+kd)^2\right] - k^2(d-1)^2\lambda_j\alpha_j^2\right\}\\ &= \sum_{j=1}^{p}\frac{k(1-d)}{\lambda_j(\lambda_j+k)^2}\left\{2\phi\lambda_j + \phi k(1-d) - k(1-d)\lambda_j\alpha_j^2\right\}\end{aligned} \tag{23}$$

The Equation (23) becomes positive if $2\phi\lambda_j + \phi k(1-d) - k(1-d)\lambda_j\alpha_j^2 > 0$ which is satisfied if we have $\min\left\{\frac{2\phi\lambda_{\mathrm{j}}}{\lambda_{\mathrm{j}}\alpha_{\mathrm{j}}^2 - \phi}\right\}_{\mathrm{j=1}}^{\mathrm{p}} > \mathrm{k}(1-\mathrm{d})$. Thus, the proof is finished. □

In the next two theorems, we provide the conditions that GTPE is superior to GRE in both MMSE and MSE sense respectively.

Theorem 3
When $\eta_{max}\left(F_{kd}C^{-1}C_{kd}C^{-1}C_k\right) < 1$, the new estimator GTPE is superior to GRE in the sense of MMSE if and only if $b_{GLTE}^{\top}\left[D_2 + b_{GRE}b_{GRE}^{\top}\right]b_{GLTE} \leq 1$ where $D_2 = C_kC^{-1}C_k - F_{kd}C^{-1}F_{kd}$.

Proof
Now, consider the difference of MMSE functions of GRE and GTPE

$$\begin{aligned}\Delta_2 &= \text{MMSE}\left(\widehat{\beta}_{\text{k}}\right) - \text{MMSE}\left(\widehat{\beta}_{(\text{k,d})}\right)\\ &= \phi\left(C_kCC_k - F_{kd}C^{-1}F_{kd}^{\top}\right) + b_{GRE}b_{GRE}^{\top} - b_{GLTE}b_{GLTE}^{\top}\end{aligned} \tag{24}$$

Since C_kCC_k and $F_{kd}C^{-1}F_{kd}^{\top}$ are positive definite, then by Lemma 2, when

$$\eta_{max}\left(F_{kd}C^{-1}F_{kd}\left[C_kCC_k\right]^{-1}\right) = \eta_{max}\left(F_{kd}C^{-1}C_{kd}C^{-1}C_k\right) < 1,$$

$D_2 = C_kC^{-1}C_k - F_{kd}C^{-1}F_{kd} > 0$. Therefore, by Lemma 3, $b_{GLTE}^{\top}\left[D_2 + b_{GRE}b_{GRE}^{\top}\right]b_{GLTE} \leq 1$ if and only if $\Delta_2 > 0$. The proof is completed. □

Theorem 4
The new estimator GTPE is superior to GRE in the sense of MSE

- if $d > 0$ and $\min\left\{\frac{\phi(2\lambda_j - d)}{\lambda_j\alpha_j^2}\right\}_{j=1}^{p} > d + 2k$ or
- if $d < 0$ and $\max\left\{\frac{\phi(2\lambda_j - d)}{\lambda_j\alpha_j^2}\right\}_{j=1}^{p} < d + 2k$.

where $k > 0$.

Proof
The MSE differences of GRE and GTPE is obtained as

$$\begin{aligned}\text{MSE}\left(\widehat{\beta}_{\text{k}}\right) - \text{MSE}\left(\widehat{\beta}_{(\text{k,d})}\right) &= \sum_{j=1}^{p}\left\{\frac{\phi\lambda_j}{(\lambda_j+k)^2} + \frac{k^2\alpha_j^2}{(\lambda_j+k)^2}\right\} - \sum_{j=1}^{p}\left\{\frac{\phi(\lambda_j+kd)^2}{\lambda_j(\lambda_j+k)^2} + \frac{k^2(1-d)^2\alpha_j^2}{(\lambda_j+k)^2}\right\}\\ &= \sum_{j=1}^{p}\frac{1}{\lambda_j(\lambda_j+k)^2}\left\{\phi\lambda_j^2 - \phi(\lambda_j+kd)^2 + k^2\lambda_j\alpha_j^2 - k^2(1-d)^2\lambda_j\alpha_j^2\right\}\\ &= \sum_{j=1}^{p}\frac{1}{\lambda_j(\lambda_j+k)^2}\left\{k^2d\lambda_j\alpha_j^2(2-d) - \phi kd(2\lambda_j - kd)\right\}.\end{aligned} \tag{25}$$

Thus, Equation (25) becomes positive if $k > \max\left\{\frac{2\phi\lambda_j}{2\lambda_j\alpha_j^2 + d\left(\phi - \lambda_j\alpha_j^2\right)}\right\}_{j=1}^{p}$. □

Finally, in the next theorem, we obtain the conditions that GTPE is better than GLE in the sense of MMSE.

Theorem 5
When $\mu_{max}\left(F_kdC^{-1}F_kd^{\top}\left[F_dC^{-1}F_d^{\top}\right]^{-1}\right) < 1$, the new estimator GTPE is better than GLE in MMSE sense if $b_{GLTE}^{\top}\left[D_3 + b_{GLE}b_{GLE}^{\top}\right]b_{GLTE} \leq 1$ where $D_3 = F_dC^{-1}F_d^{\top} - F_kdC^{-1}F_kd^{\top}$.

Proof
Consider the following MMSE difference

$$\begin{aligned}\Delta_3 &= \text{MMSE}\left(\widehat{\beta}_{\text{d}}\right) - \text{MMSE}\left(\widehat{\beta}_{(\text{k,d})}\right)\\ &= \phi\left(F_dC^{-1}F_d^{\top} - F_kdC^{-1}F_kd^{\top}\right) + b_{GLE}b_{GLE}^{\top} - b_{GLTE}b_{GLTE}^{\top}\end{aligned} \tag{26}$$

Since $F_dC^{-1}F_d^\top$ and $F_kdC^{-1}F_kd^\top$ are positive definite, then by Lemma 2, when

$$\mu_{max}\left(F_kdC^{-1}F_kd^\top\left[F_dC^{-1}F_d^\top\right]^{-1}\right)<1$$

$D_3=F_dC^{-1}F_d^\top-F_kdC^{-1}F_kd^\top>0$. Therefore, by Lemma 3, $b_{GLTE}^\top\left[D_3+b_{GLE}b_{GLE}^\top\right]b_{GLTE}\leq 1$ if and only if $\Delta_3>0$. The proof is completed. □

2.5. Selection of the biasing parameters k and d

In order to obtain less MSE values and efficient regression coefficients, one needs to estimate the parameters of GTPE accordingly. Following [15] and [10], the values of k and d may be chosen iteratively as follows: Firstly, taking derivative of Equation (21) with respect to the parameter k and equating the resulting quantity to zero then one obtains the following

$$\frac{\partial \text{MSE}\left(\widehat{\beta}_{(\text{k,d})}\right)}{\partial k}=2\sum_{j=1}^{p}\frac{\phi\left(\lambda_j+kd\right)(d-1)+k(d-1)^2\lambda_j\alpha_j^2}{\left(\lambda_j+k\right)^3}=0. \tag{27}$$

As in [15] and [17], it is possible to equate the numerator of Eq. (27) to zero and solve for each individual parameter as

$$k_j=\frac{\phi\lambda_j}{(1-d)\lambda_j\alpha_j^2-\phi d} \tag{28}$$

where $j=1,2,...p$. Since each k_j should be positive, we can obtain a condition such that $0<d<1$ by

$$d<\min\left\{\frac{\lambda_j\alpha_j^2}{\phi+\lambda_j\alpha_j^2}\right\}. \tag{29}$$

Now, after obtaining an estimate of d using Eq. (29), it is easy to compute the individual parameters k_j's. However, we only need an estimate of the parameter k. Therefore, following [10], we propose the following estimators of k:

- $k_1=\frac{1}{p}\sum_{j=1}^{p}\left\{\frac{\phi\lambda_j}{(1-d)\lambda_j\alpha_j^2-\phi d}\right\}$ which is the arithmetic mean of k_j's.
- $k_2=\text{median}\left\{\frac{\phi\lambda_j}{(1-d)\lambda_j\alpha_j^2-\phi d}\right\}$
- $k_3=\min\left\{\frac{\phi\lambda_j}{(1-d)\lambda_j\alpha_j^2-\phi d}\right\}$
- $k_4=\max\left\{\frac{\phi\lambda_j}{(1-d)\lambda_j\alpha_j^2-\phi d}\right\}$
- $k_5=\left\{\prod_{j=1}^{p}\left\{\frac{\phi\lambda_j}{(1-d)\lambda_j\alpha_j^2-\phi d}\right\}\right\}^{\frac{1}{p}}$ which is the geometric mean of k_j's.

3. Monte Carlo Simulation

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

3.1. Simulation design

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

$$y_i\sim Gamma(\theta_i,1.5), \tag{30}$$

where $\theta_i = \exp(x_i^\top \beta)$, $\beta = (\beta_1, ..., \beta_p)$ with $\sum_{j=1}^{p} \beta_j^2 = 1$ and $\beta_1 = \beta_2 = ... = \beta_p$ [17]. The explanatory variables $x_i^\top = (x_{i1}, x_{i2}, ..., x_{in})$ have been generated from the following formula

$$x_{ij} = (1 - \rho^2)^{1l2} w_{ij} + \rho w_{ip}, \quad i = 1, 2, ..., n, \quad j = 1, 2, ..., p, \tag{31}$$

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: 50, 100 and 150. 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, since we are interested in the effect of multicollinearity, in which the degrees of correlation considered to be more important, three values of the pairwise correlation are taken into account such that $\rho = \{0.90, 0.95, 0.99\}$. For a combination of these different values of n, p and ρ the generated data is repeated 1000 times. For each replication, the MSE is calculated as

$$\text{MSE}_\text{i}\left(\widehat{\beta}\right) = \left(\widehat{\beta} - \beta\right)^\top \left(\widehat{\beta} - \beta\right), \quad i = 1, 2, ..., 1000, \tag{32}$$

where $\hat{\beta}$ is defined as

- $\widehat{\beta}_{\text{MLE}}$
- $\widehat{\beta}_{\text{k}}$ with $\hat{k} = \frac{\hat{\phi}}{\hat{\alpha}^T \hat{\alpha}}$, $\widehat{\alpha}_j = Q^\top \widehat{\beta}_{\text{MLE}}$ and $\hat{\phi}$ is the estimated dispersion parameter which is computed as $\hat{\phi} = (n-p)^{-1} \sum_{i=1}^{n} (y_i - \hat{\mu}_i / \hat{\mu}_i)^2$, (Pearson residual [18]).
- $\widehat{\beta}_{\text{d}}$ with an optimal d value as in [27], $d = \max(0, max(\lambda_j(\hat{\alpha}_j^2 - \hat{\phi})/(\hat{\phi} + \lambda_j \hat{\alpha}_j^2)))$.
- $\widehat{\beta}_{(\text{k,d})}$ with the optimal d (Eq. (29)) after substituting $k = \hat{k}$, $\lambda_j = \hat{\lambda}_j$, $\alpha_j = \hat{\alpha}_j$, and $\phi = \hat{\phi}$.

The averaged performance from 1000 simulations are summarized in terms of the MSE. All the computations are done using R programming language [30].

3.2. Simulation results

The average MSE values of the MLE, GRE, GLE, and GTPE are presented in Tables 1-3 for $n = 50$, $n = 100$, and $n = 150$, respectively. We may observe that the MSE of GTPE in all situations of selecting k is less than the other used estimators, which clearly shows that the GTPE outperforms the MLE, GRE, and GLE in all of the cases. It is clearly seen that the optimal selection k_3 in GTPE works fine comparing with k_1, k_2, k_4, and k_4 with respect to MSE.

Further, the MSE values clearly increase when p increases. However, there is a clear advantage of the GTPE over the other competitor estimators. In addition, it is clear that increasing the n values leading to decreasing in MSE. For the large n, however, there is a clear advantage of using the GTPE.

Regarding the degree of correlation, the MSE increases monotonously when the correlation between the explanatory variables increasing. The GTPE again outperforms the MLE, GRE, and GLE.

4. A Real Data Application

To further demonstrate the usefulness of the GLTE in real application, we present here a chemistry dataset with $(n, p) = (65, 15)$, where n is representing the number of imidazo[4,5-b]pyridine derivatives, which are used as anticancer compounds. While, p is denoting the number of molecular descriptors, which are treated as explanatory variables [4]. The response of interest is the biological activities (IC_{50}). Quantitative structure-activity relationship (QSAR) study has become a great deal of importance in chemometrics. The principle of QSAR is to model several biological activities over a collection of chemical

Table 1. Averaged MSE values when $n = 50$

p	ρ	MLE	GRE	GLE	GTPE				
					k_1	k_2	k_3	k_4	k_5
4	0.90	4.311	4.091	1.834	1.138	1.379	1.126	1.148	1.412
	0.95	4.567	4.394	1.954	1.159	1.411	1.142	1.163	1.444
	0.99	6.993	4.737	2.283	1.181	1.439	1.146	1.168	1.472
8	0.90	4.348	4.746	1.874	1.245	1.488	1.243	1.265	1.521
	0.95	4.891	4.843	1.987	1.253	1.569	1.255	1.276	1.602
	0.99	10.417	4.877	2.775	1.255	1.605	1.259	1.28	1.638

Table 2. Averaged MSE values when $n = 100$

p	ρ	MLE	GRE	GLE	GTPE				
					k_1	k_2	k_3	k_4	k_5
4	0.90	4.281	4.071	1.814	1.118	1.359	1.106	1.128	1.392
	0.95	4.547	4.374	1.934	1.14	1.391	1.122	1.143	1.425
	0.99	6.973	4.717	2.264	1.161	1.419	1.127	1.148	1.452
8	0.90	4.328	4.726	1.854	1.225	1.468	1.224	1.245	1.501
	0.95	4.871	4.823	1.967	1.233	1.549	1.235	1.257	1.583
	0.99	10.398	4.858	2.755	1.235	1.585	1.239	1.261	1.618

Table 3. Averaged MSE values when $n = 150$

p	ρ	MLE	GRE	GLE	GTPE				
					k_1	k_2	k_3	k_4	k_5
4	0.90	4.228	4.018	1.761	1.066	1.306	1.054	1.075	1.34
	0.95	4.495	4.322	1.881	1.087	1.339	1.069	1.091	1.372
	0.99	6.921	4.664	2.211	1.108	1.366	1.074	1.095	1.404
8	0.90	4.275	4.673	1.802	1.172	1.415	1.171	1.192	1.448
	0.95	4.818	4.771	1.914	1.182	1.497	1.183	1.204	1.532
	0.99	10.345	4.805	2.702	1.182	1.532	1.186	1.208	1.565

compounds in terms of their structural properties [5]. Consequently, using of multiple regression model is one of the most important tools for constructing the QSAR model. A description of the used explanatory variables is provided in Table 4. All the variables are numerical.

First, to check whether the response variable belongs to the gamma distribution, Chi-square test is used. The result of the test equals to 9.3657 with p-value equals to 0.4534. It is indicated form this result that the gamma distribution fits very well to this response variable.

Second, to check whether there is a relationship among the explanatory variables or not, Figure 1 displays the correlation matrix among the 15 explanatory variables. It is obviously seen that there are correlations greater than 0.90 among MW, SpMaxA_D, and ATS8v ($r = 0.96$), between SpMax3_Bh(s) and ATS8v ($r = 0.92$), and between Mor21v and Mor21e ($r = 0.93$).

Third, to test the existence of multicollinearity, the eigenvalues of the matrix $X^\top \widehat{W} X$ are obtained as 2.14×10^9,3.85×10^6, 2.42×10^5, 1.26×10^4, 1.29×10^3, 2.14×10^9,9.01×10^2, 4.71×10^2, 1.71×10^2, 5.93×10^1, 3.24×10^1, 2.77×10^1, 1.78×10^1, 9.56, and 1.23. The determined condition number $\text{CN} = \sqrt{\lambda_{\max}/\lambda_{\min}}$ of the data is 41652.77 indicating that the severe multicollinearity issue exists.

The estimated gamma regression coefficients and MSE values for the MLE, GRE, GLE, and GTPE using k_3 estimators are listed in Table 5. According to Table 5, it is clearly seen that the GTPE shrinkages the value of the estimated coefficients efficiently. Additionally, in terms of the MSE, there is an important

Table 4. Description of the used explanatory variables

Variable name's	description
MW	molecular weight
IC3	Information Content index (neighborhood symmetry of 3-order)
SpMaxA_D	normalized leading eigenvalue from topological distance matrix
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
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
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
TDB08m	3D Topological distance based descriptors - lag 8 weighted by mass
RDF100m	Radial Distribution Function - 100 / weighted by mass
Mor21v	signal 21 / weighted by van der Waals volume
Mor21e	signal 21 / weighted by Sanderson electronegativity
HATS6v	leverage-weighted autocorrelation of lag 6 / weighted by van der Waals volume

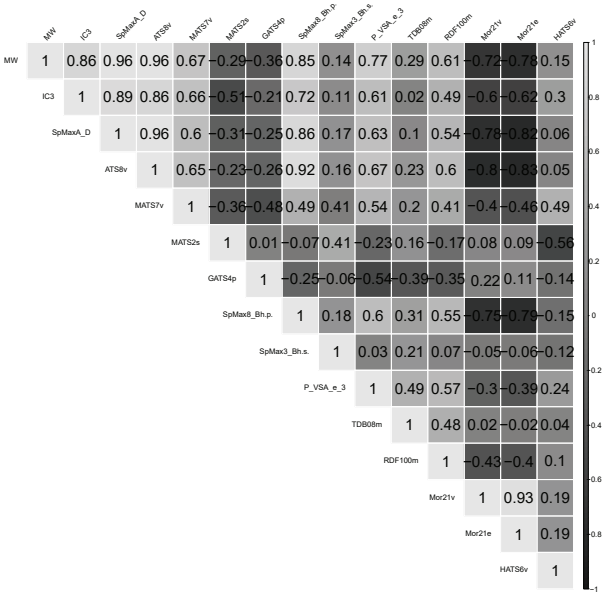

Figure 1. Correlation matrix of the real data.

reduction in favor of the GTPE. Specifically, it can be seen that the MSE of the GTPE was about 46.11%, 31.09%, and 26.40% lower than that of MLE, GRE, and GLE, respectively.

5. Conclusion

In this paper, gamma two-parameter estimator is proposed to overcome the multicollinearity problem in the gamma regression model. According to Monte Carlo simulation studies, the GTPE has better performance than MLE, GRE, and GLE estimators, in terms of MSE. Additionally, a real data application is also considered to illustrate benefits of using the GTPE in the context of gamma regression model.

Table 5. The estimated coefficients and MSE values of the listed estimators

	MLE	GRE	GLE	GTPE (k_3)
$\widehat{\beta}_{\text{MW}}$	1.0383	0.7773	0.8713	0.7703
$\widehat{\beta}_{\text{IC3}}$	1.2733	1.0133	1.1063	1.0053
$\widehat{\beta}_{\text{SpMaxA\_D}}$	-1.0657	-1.3267	-1.2327	-0.8657
$\widehat{\beta}_{\text{ATS8v}}$	-1.3427	-1.6037	-1.5097	-1.1427
$\widehat{\beta}_{\text{MATS7v}}$	-1.1827	-1.4437	-1.3497	-0.9827
$\widehat{\beta}_{\text{MATS2s}}$	-1.1787	-1.4397	-1.3457	-0.9787
$\widehat{\beta}_{\text{GATS4p}}$	-1.2007	-1.4617	-1.3687	-1.0007
$\widehat{\beta}_{\text{SpMax8\_Bh(p)}}$	2.5423	2.2813	2.3753	2.7433
$\widehat{\beta}_{\text{SpMax3\_Bh(s)}}$	2.1053	1.8443	1.9383	2.3053
$\widehat{\beta}_{\text{P\_VSA\_e\_3}}$	2.0373	1.7753	1.8693	2.2363
$\widehat{\beta}_{\text{TDB08m}}$	-2.0667	-2.3287	-2.2337	-1.8667
$\widehat{\beta}_{\text{RDF100m}}$	1.6073	1.3453	1.4393	1.8063
$\widehat{\beta}_{\text{Mor21v}}$	-2.3977	-2.6587	-2.5647	-2.1987
$\widehat{\beta}_{\text{Mor21e}}$	-2.3157	-2.5767	-2.4827	-2.1157
$\widehat{\beta}_{\text{HATS6v}}$	2.2473	1.9863	2.0803	2.4473
MSE	4.075	3.187	2.984	2.196

The superiority of the GTPE based on the resulting MSE was observed and it was shown that the results are consistent with Monte Carlo simulation results. In conclusion, the use of the GTPE is recommended when multicollinearity is present in the gamma regression model.

REFERENCES

1. A. M.Al-Abood, and D. H. Young, Improved deviance goodness of fit statistics for a gamma regression model, Communications in Statistics-Theory and Methods, vol. 15, no. 6, pp. 1865-1874, 1986.
2. Z. Y. Algamal, Developing a ridge estimator for the gamma regression model, Journal of Chemometrics, vol. 32, e3054, 2018
3. Z. Y. Algamal, Shrinkage estimators for gamma regression model, Electronic Journal of Applied Statistical Analysis, vol. 11, no. 1, pp. 253-268, 2018.
4. Z. Y. Algamal, M. H. Lee, A. M. Al-Fakih, and M. Aziz, High-dimensional QSAR prediction of anticancer potency of imidazo [4, 5-b] pyridine derivatives using adjusted adaptive LASSO, Journal of Chemometrics, vol. 29, pp. 547-556, 2015.
5. Z. Y. Algamal, and M. H. Lee, A novel molecular descriptor selection method in QSAR classification model based on weighted penalized logistic regression, Journal of Chemometrics, vol. 31, pp. e2915, 2017
6. Z. Y. Algamal, and Y. Asar, Liu-type estimator for the gamma regression model, Communications in Statistics-Simulation and Computation, vol. 49, no. 8, pp. 2035-2048, 2020.
7. N. A. Al-Thanoon, O. S. Qasim, and Z. Y. Algamal, Variable selection in Gamma regression model using binary gray Wolf optimization algorithm, Journal of Physics: Conference Series, vol. 1591, no. 1, pp. 12-36, 2020.
8. M. Amin, M. Qasim, A. Yasin, and M. Amanullah, Almost unbiased ridge estimator in the gamma regression model, Communications in Statistics-Simulation and Computation, vol. 1, no. 1, pp. 1-20, 2020.
9. M. Amin, M. Qasim, M. Amanullah, and S. Afzal, Performance of some ridge estimators for the gamma regression model, Statistical papers, vol. 61, no. 3, pp. 997-1026, 2020.
10. Y. Asar Y, and A. Genç, New shrinkage parameters for the Liu-type logistic estimators, Communications in Statistics-Simulation and Computation, vol. 45, no. 3, pp. 1094-1103, 2016.
11. Y. Asar, and A. Genç, A new two-parameter estimator for the Poisson regression model, Iranian Journal of Science and Technology, Transactions A: Science, vol. 42, no. 2, pp. 793-803, 2018.
12. P. De Jong, P., and G. Z. Heller, Generalized linear models for insurance data, Cambridge: Cambridge University Press, 2008.
13. E. Dunder, S. Gumustekin, and M. A. Cengiz, Variable selection in gamma regression models via artificial bee colony algorithm, Journal of Applied Statistics, vol. 45, no. 1, pp. 8-16, 2018.
14. R. W. Farebrother, Further Results on the Mean Square Error of Ridge Regression, Journal of the Royal Statistical Society B, vol. 38, pp. 248-250, 1976.

15. A. E. Hoerl, and R. W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems Technometrics, vol. 12, no. 1, pp. 55-67, 1970.
16. J. Huang, and H. Yang, A two-parameter estimator in the negative binomial regression model, Journal of Statistical Computation and Simulation, vol. 84, no. 1, pp. 124-134, 2012.
17. B. G. Kibria, Performance of some new ridge regression estimators, Communications in Statistics-Simulation and Computation, vol. 32, vol. 2, pp. 419-435, 2003.
18. A. I. Khuri, Linear model methodology, FL Chapman & Hall/CRC Press, 2010.
19. F. Kurtoglu, and M. R. Özkale, Liu estimation in generalized linear models: application on gamma distributed response variable, Statistical Papers, vol. 57, no. 4, pp. 911-928, 2016.
20. K. Liu, A new class of biased estimate in linear regression, Communications in Statistics-Theory and Methods, vol. 22, no. 2, pp. 393-402, 1993.
21. A. F. Lukman, K. Ayinde, B. M. Kibria, and E. T. Adewuyi, Modified ridge-type estimator for the gamma regression model, Communications in Statistics-Simulation and Computation, vol. 1, no. 2, pp. 1-15, 2020.
22. M. J. Mackinnon, and M. L. Puterman, Collinearity in generalized linear models, Communications in Statistics-Theory and Methods, vol. 18, no. 9, pp. 3463-3472, 1989.
23. A. S. Malehi, F. Pourmotahari, and K. A. Angali, Statistical models for the analysis of skewed healthcare cost data: A simulation study, Health Economics Review, vol. 5, no. 11, 2013.
24. S. Mandal, A. Belaghi, A. Mahmoudi, and M. Aminnejad, Stein-type shrinkage estimators in gamma regression model with application to prostate cancer data, Statistics in Medicine, vol. 38, no. 22, pp. 4310-4322, 2019.
25. K. Månsson, and G. Shukur, A Poisson ridge regression estimator, Economic Modelling, vol. 28, no. 4, pp. 1475-1481, 2011.
26. K. Månsson, On ridge estimators for the negative binomial regression model, Economic Modelling, vol. 29, no. 2 , pp. 178-184, 2012.
27. K. Månsson, B. G. Kibria, G. Shukur, On Liu estimators for the logit regression model, Economic Modelling, vol. 29, no. 4, pp. 1483-1488, 2012.
28. M. R. Özkale, and S. Kaçıranlar, The restricted and unrestricted two-parameter estimators, Communications in Statistics-Theory and Methods, vol. 36, no. 15, pp. 2707-2725, 2007.
29. M. Qasim, M. Amin, and M. Amanullah, On the performance of some new Liu parameters for the gamma regression model, Journal of Statistical Computation and Simulation, vol. 88, no. 16, pp. 3065-3080, 2018.
30. R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2018.
31. R. Rao, C. Helge, S. Toutenburg, and C. Heumann, Linear models and generalizations, least squares and alternatives, Springer Berlin Heidelberg New York, 2008.
32. R. L. Schaefer, L. D. Roi, and R. A. Wolfe, A ridge logistic estimator, Communications in Statistics–Theory and Methods, vol. 13, no. 1, pp. 99-113, 1984.
33. B. Segerstedt, On ordinary ridge regression in generalized linear models, Communications in Statistics-Theory and Methods, vol. 21, no. 8, pp. 2227-2246, 1992.
34. N. N. Urgan, and M. Tez, Liu estimator in logistic regression when the data are collinear, In 20th EURO Mini Conference (pp. 323-327), 2008.
35. H. W. Wasef, A derivation of prediction intervals for gamma regression, Journal of Statistical Computation and Simulation, vol. 86, no. 17, pp. 3512-3526, 2016.
36. J. Wu, and Y. Asar, More on the restricted Liu estimator in the logistic regression model, Communications in Statistics-Simulation and Computation, vol. 46, no. 5, pp. 3680-3689, 2017.