AIMS Mathematics

AIMS Mathematics, 11(6): 19321–19337.
DOI: 10.3934/math.2026785
Received: 10 March 2026
Revised: 30 May 2026
Accepted: 15 June 2026
Published: 30 June 2026

https://www.aimspress.com/journal/Math

Research article

Mitigating multicollinearity issues to improve the two-parameter estimators for a Beta regression model

Muteb Faraj Alharthi[1] and Nadeem Akhtar[2,*]

[1] Department of Mathematics and Statistics, College of Science, Taif University, Taif 21944, Saudi Arabia; Email: muteb.faraj@tu.edu.sa
[2] Government Degree College Achini Payan, Higher Education, Archives and Libraries Department Peshawar 25000, Khyber Pakhtunkhwa, Pakistan; Email: nadeemscholar@icp.edu.pk

* **Correspondence:** Email: nadeemscholar@icp.edu.pk; Tel: +923459805047.

Abstract: The Maximum Likelihood Estimation (MLE) method is commonly used for parameter estimation in statistical models. However, it can be significantly affected by multicollinearity issues, which undermine the reliability of parameter estimates. In Beta regression models (BRMs) for continuous proportions in (0,1), multicollinearity complicates estimation and produces unstable results when using MLE. To address this challenge, ridge-based and alternative regression estimators have been introduced for BRMs with logit links. Selecting a two-parameter estimator remains a challenge for obtaining reliable parameter estimates for the BRM. In this study, we proposed new two-parameter ridge estimators for the BRM that effectively mitigate multicollinearity. We evaluated the performance of our newly proposed estimators and existing methods based on the mean squared error (MSE) criterion as well as bias, variance, and predictive accuracy. Simulation results showed that the proposed estimators generally performed better than the existing methods across most scenarios. Furthermore, a real-world data application confirmed the simulation results, demonstrating that our proposed estimators provide reliable and robust parameter estimates compared to MLE and the other existing methods.

Keywords: maximum likelihood estimation; Beta regression model; Ridge parameter; Monte Carlo simulation; mean squared error
Mathematics Subject Classification: 62J07, 62F10, 62G05

1. Introduction

When the response variable is a continuous proportion rate or percentage bounded between 0 and 1, traditional linear regression models are not suitable. In such situations, the Beta regression model (BRM) is used for parameter estimation. The BRM is widely applied in fields such as ecology, economics, digital currency trading, and epidemiology, where the response variables are continuous proportions, rates, or percentages [1,2].

Multicollinearity can distort parameter estimates in Beta regression when standard Maximum Likelihood Estimation (MLE) is used. To reduce this problem, shrinkage (regularization) methods are often applied, and an appropriate bias (tuning) parameter is selected to obtain more stable and reliable regression estimates [3,4]. In practice, parameter estimation for Beta regression is mostly carried out using either MLE or a Bayesian approach. MLE method estimates the model parameters by maximizing the likelihood of the observed data and generally performs well when multicollinearity is not severe. By contrast, the Bayesian framework incorporates prior information; it is typically used when credible prior knowledge is available (e.g., from subject-matter experts) and can be especially helpful for improving estimation stability in challenging data settings [5,6]. The MLE method is generally preferred for large samples because, when multicollinearity is not a concern, it tends to provide efficient and reliable parameter estimates. Conversely, Bayesian methods are often favored in small-sample settings, where incorporating prior information can help stabilize estimation. The Beta model has useful properties and is widely applied in real-world settings, where predictors may be related. In these situations, standard estimation approaches such as MLE or Bayesian estimation can still produce reasonable results. However, an important limitation of these models is their sensitivity to multicollinearity, which can inflate standard errors, destabilize coefficient estimates, and weaken overall model fit [7]. This issue is somewhat analogous to challenges faced in interpolation methods, where techniques like those developed by [8] for unstructured data and [9] for kernel interpolation aim to adapt to complex data structures and improve accuracy.

In practice, predictors are frequently dependent on one another, which undermines the independence assumptions commonly relied on in regression modeling to ensure stable estimations. When predictors are highly correlated, the resulting estimates can become unstable and less trustworthy [10,11]. To address these issues, regularization methods introduce bias in a controlled way to reduce variance, making Beta regression estimates more stable and often more reliable than those obtained using the MLE alone under severe multicollinearity.

In ridge regression applied to the Beta regression framework, the shrinkage parameter (k) controls the bias in parameter estimation, which helps reduce overfitting and produces reliable estimates for the BRM when multicollinearity is present, as compared to the MLE method. The researchers in [12] introduced a Beta Liu-type of estimator to mitigate severe multicollinearity issues and to estimate the parameters of the Beta model.

To mitigate the issues of multicollinearity, many researchers have introduced two-parameter ridge estimators (k, d) to address multicollinearity issues, as compared to traditional or basic ridge estimators based on the estimated mean squared error (MSE) criterion [13,14]. The basic ridge estimators and two-parameter ridge estimators are biased; therefore, the efficiency of these estimators is measured by their estimated MSE. An estimator is considered more efficient than others if its MSE is lower than that of the compared estimators. Moreover, the researchers in [15,16] introduced two-parameter ridge estimators for improved parameter estimates in the BRM under multicollinearity. Evidence from other studies suggests that no ridge-type estimator is uniformly best across all multicollinearity settings. For this reason, selecting an appropriate shrinkage parameter remains a difficult and still-developing topic, particularly when the goal is to obtain stable and accurate inference in Beta regression with strongly correlated predictors.

Motivated by this limitation, we develop a new modified two-parameter ridge-type estimator (MTRTE) for the BRM and examine its performance relative to the MLE and several competitors. The proposed approach introduces two tuning parameters, k and d, to regulate the degree of shrinkage and improve numerical stability when multicollinearity is severe. By combining these two controls, the MTRTE is intended to yield coefficient estimates that are less variable and more reliable than those obtained from the MLE and other common ridge or Liu-type alternatives under strong predictor dependence.

The paper is organized as follows: In Section 2, we present the proposed MTRTE and review the competing estimators with derivations. In Section 3, we describe the Monte Carlo simulation design and evaluation criteria. In Section 4, we provide an empirical application to real data. In Section 5, we conclude the paper and summarize the key findings.

2. Materials and methods

In this section, we introduce the BRM and review the major shrinkage-based estimators considered in this study, including the MLE, the single parameter ridge-type estimators, the Liu estimators, the two-parameter ridge estimators, and the newly proposed modified two-parameter ridge-type estimators.

2.1. Beta regression model

The MLE method is used to estimate the parameters of the BRM when the sample size is large, as it has good asymptotic properties. However, the MLE method is highly sensitive to multicollinearity issues. When the explanatory variables are strongly correlated with each other, the MLE method estimates the parameters of the regression model in an unstable manner, leading to large standard deviations, even when the sample size is large. Additionally, the Beta model is nonlinear, and the bounded nature of the response variable (0, 1) further exacerbates the challenges of using this method in the presence of multicollinearity. Within the regression modeling framework, the expected value of the response can be linked to a set of explanatory variables by choosing a suitable link function [17]. The model includes a precision parameter, which is commonly interpreted as being inversely related to the variance (i.e., larger precision implies less variability).

Assume that Y follows a Beta distribution with parameters α and β. Then, the probability density function of Y is given by:

$$f(Y;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}, \quad 0 < y < 1. \tag{2.1}$$

$\Gamma(.)$ is the Gamma function, and the mean and variance of Y is given by

$$\mathrm{E(Y)} = \frac{\alpha}{\alpha+\beta} \text{ and } \mathrm{Var(Y)} = \frac{\alpha\ \beta}{(\alpha+\beta)^2((\alpha+\beta+1)}.$$

The $B(\alpha,\beta)$ is the Beta function, which can be written in terms of the Gamma function as follows:

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

For regression, we use the mean-precision parameterization:

$$\mu = \frac{\alpha}{\alpha+\beta} \text{ and } \alpha + \beta = \varphi.$$

By re-parameterization, the Beta model parameters improve the efficiency of the estimations, where $\mu\varphi = \alpha$ and $\beta = (1-\mu)\varphi,$ and the probability distribution can be expressed in the following form:

$$f(y;\mu,z) = \frac{\boldsymbol{\Gamma}(\varphi)}{\boldsymbol{\Gamma}(\mu\varphi)\boldsymbol{\Gamma}((1-\mu)\varphi)} y^{\mu\varphi-1}(1-y)^{(1-\mu)\varphi-1},\ 0 < y < 1. \quad (2.2)$$

In the BRM, the response variable is bounded between 0 and 1. The model is typically characterized by two major components: The mean (or location) parameter (often denoted by (μ), which lies in (0,1)), and a precision parameter (φ), which must be strictly positive. The precision parameter controls the spread of the response around its mean, where larger values of the precision parameter indicate lower variability, and smaller values imply greater dispersion. In the Beta distribution, φ also appears through normalizing constants involving the Gamma function. Precision φ is assumed constant across observations (for simplicity; it could also be modeled but it is a nuisance parameter). Thus, we adjust the BRM mean and variance follows as:

$$Mean(y) = \mu \text{ and } \quad Var(y) = \frac{\mu(1-\mu)}{\varphi+1}.$$

For the BRM, the mean (or location) parameter is modeled as a function of the covariates through a link function. In most applications, the logit link is used, so the relationship is written as:

$$g(\mu_i) = \boldsymbol{\eta_i} = \boldsymbol{X_i^T}\beta, i = 1,2,\ldots,n, \quad (2.3)$$

where $g(.)$ denotes a monotonic and differentiable link function that connects the systematic component to the random component, $\boldsymbol{\eta_i}$ represents the linear predictor, $\beta = \left(\beta_{\boldsymbol{1}},\beta_{\boldsymbol{2}},\ldots,\beta_{\boldsymbol{p}}\right)^{\boldsymbol{T}} \in \boldsymbol{R^{p\times 1}}$ is the vector of unknown parameters of the model, and $\boldsymbol{X_i} = (\boldsymbol{x_{i1}},\boldsymbol{x_{i2}},\ldots,\boldsymbol{x_{ip}}) \ \in \boldsymbol{R^{n\times p}}$ is the vector of the predictors. The logit link $g(\mu) = log(\frac{\mu}{1-\mu})$ is most common, so we have

$$\mu_i = \frac{\exp(\boldsymbol{X_i^T\beta})}{1+\exp(\boldsymbol{X_i^T\beta})}.$$

2.2. *Maximum likelihood estimator method*

Estimation parameters of the BRM can be obtained using the MLE method [18]. The likelihood function of the model can be written as:

$$\ell(y_i;\beta,\varphi) = \sum_{i=i}^{n}\{log\Gamma(\varphi) - log\Gamma(\mu_i\varphi) - \log(1-\mu_i)\varphi) + (\mu_i\varphi - 1)\log(y_i) + ((1-\mu_i)\varphi - 1)\log(1-y_i)\}. \quad (2.4)$$

The parameters are estimated by maximizing ℓ. Because the score equations are non-linear, we use Fisher scoring, which requires the 1st and 2nd derivatives.
Define the di-gamma function $\psi(z) = \frac{d}{dz}log\Gamma(z)$. Differentiating Eq (2.4) with respect to μ_i:
Where $\eta_i = \boldsymbol{X_i^T\beta}$ is the linear predictor. The individual derivatives are: $\frac{\partial\mu_i}{\partial\eta_i} = \mu_i(1-\mu_i)$,

$$\frac{\partial\ell}{\partial\mu_i} = \varphi\left[\psi(1-\mu_i)\varphi) - \psi\ (\mu_i\varphi) + log\ \frac{y_i}{1-y_i}\right]. \quad (2.5)$$

Let $\boldsymbol{y_i^*} = \boldsymbol{log}\left(\frac{\boldsymbol{y_i}}{\boldsymbol{1-y_i}}\right)$, and $\mu_{\boldsymbol{i}}^* = \boldsymbol{\psi}(\mu_{\boldsymbol{i}}\varphi) - \boldsymbol{\psi}((\boldsymbol{1}-\mu_{\boldsymbol{i}})\varphi)$. Then Eq (2.5) simplifies to:

$$\frac{\partial \ell}{\partial \mu_i} = \varphi \ \ (\mu_i{}^* - \boldsymbol{y}^*).$$

Now apply the chain rule for each coefficient β_j:

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n} \frac{\partial \ell}{\partial \mu_i} . \frac{\partial \mu_i}{\partial \boldsymbol{\eta}_\mathbf{i}} . \frac{\partial \eta_\mathbf{i}}{\partial \beta_j}.$$

We have:

$$\frac{\partial \eta_\mathbf{i}}{\partial \beta_j} = x_{ij}, \text{and the logit link,} \frac{\partial \mu_i}{\partial \boldsymbol{\eta}_\mathbf{i}} = \mu_i(1-\mu_i).$$

Thus,

$$S(\beta) = \frac{\partial \ell}{\partial \beta} = \varphi \quad \sum_{i=1}^{n} (\mu_i{}^* - \boldsymbol{y}^*)\, \mu_i(1-\mu_i) \ \ X_i. \tag{2.6}$$

In matrix form, we have

$$\boldsymbol{T} = (\mu_1(1-\mu_1), \ldots, \mu_n(1-\mu_n)\ \ ), \qquad \boldsymbol{y}^* = (\boldsymbol{y}_\mathbf{1}^*, \boldsymbol{y}_\mathbf{2}^*, \ldots, \boldsymbol{y}_\boldsymbol{n}^*)^{\boldsymbol{T}},$$
$$\mu^* = (\mu_\mathbf{1}^*, \mu_\mathbf{2}^*, \ldots, \mu_\boldsymbol{n}^*)^{\boldsymbol{T}}, \qquad S(\beta) = \varphi \ \ \boldsymbol{X}^{\boldsymbol{T}}\mathbf{T} \ \ (\mathbf{y}^* - \mu^*). \tag{2.7}$$

The observed information matrix is $\frac{\partial^2 \ell}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}$. To obtain the expected Fisher information, we use $\boldsymbol{E}(\boldsymbol{y}_\boldsymbol{i}^*) = \mu_\boldsymbol{i}^*$ and the fact that:

$$Var(y_\boldsymbol{i}^*) = \varphi^{-1}[\psi^T(\mu_i\varphi) + \psi^T(1-\mu_i\varphi)], \tag{2.8}$$

where $\psi^T(\mathrm{z}) = \frac{d}{dz} \ \ \psi(\mathrm{z})$ is the tri-gamma function. The derivative of $\mu_\boldsymbol{i}^*$ with respect to μ_i is:

$$\frac{\partial \mu_\boldsymbol{i}^*}{\partial \mu_i} = \varphi[\psi'(\mu_i\varphi) - \psi' \ \ (1-\mu_i)\varphi)].$$

Using the chain rule and the fact that $-E\left[\frac{\partial^2 \ell}{\partial \boldsymbol{\beta}_j \partial \boldsymbol{\beta}_k}\right] = Cov(\frac{\partial \ell}{\partial \boldsymbol{\beta}_j}, \frac{\partial \ell}{\partial \boldsymbol{\beta}_k})$, after simplification, we obtain the Fisher information matrix for β: $\boldsymbol{I}(\beta) = \varphi \ \ \boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{X}$, where $W = dia(w_1, w_2, \ldots, w_n)$ with $\mathrm{w_i} = [\psi^T(\mu_i) + \psi^T(1-\mu_i)] \ \ \mu_\boldsymbol{i}^\mathbf{2} \ \ (1\text{-}\mu_i)^2$.

Derivation outline: From Eq (2.6), the score for β_j is $\varphi \ \ \sum_{i=1}^{n} (\mu_i{}^* - \boldsymbol{y}^*)\, \mu_i(1-\mu_i) \ \ x_{ij}$. The covariance between scores for β_j and β_k is:

$$\varphi^2 \quad \sum_{i=1}^{n} \mu_i^2 (1-\mu_i)^2 x_{ij} \ \ x_{ik} \ \ Var(\boldsymbol{y}_\boldsymbol{i}^*),$$

$Var(\boldsymbol{y}_\boldsymbol{i}^*)$ is defined in Eq (2.8).

Here, $\boldsymbol{y}_\boldsymbol{i}^*$ are independent. The MLE of $\widehat{\boldsymbol{\beta}}$ solves $S(\boldsymbol{\beta}) = 0$. Starting from an initial estimate $\beta^{(0)}$ (e.g., from least squares on $log \ \ \left(\frac{y_i}{1-y_i}\right)$), we iterate:

$$\boldsymbol{\beta}^{(r+1)} = \boldsymbol{\beta}^r + \left(\boldsymbol{I}(\boldsymbol{\beta}^{(r)})\right)^{-\mathbf{1}} \boldsymbol{S}_\beta^r\left(\boldsymbol{\beta}^{(r)}\right) \quad r = 0, 1, 2, \ldots$$

Here, S_β^r denotes the iteration (r) function, and $I_{\beta\beta}$ is information matrix for β (see [19] for more details). The least square method is usually used for the value of β, while the precision parameters are initialized as follows:

$$\varphi_i^* = \frac{\hat{\mu}_i(1-\hat{\mu}_i)}{\delta^2}, \tag{2.9}$$

where $\hat{\beta}_i$ is the estimated parameters of the BRM. The iterative process continues until the difference between successive estimates is below a predefined threshold, signifying convergence. The final step involves calculating the MLE of the model, with $\boldsymbol{X} \in R^{n\times p}$.

$$\hat{\beta}_{\boldsymbol{MLE}} = (\boldsymbol{X^T W X})^{-\boldsymbol{1}}\boldsymbol{X^T W z}, \tag{2.10}$$

where $\widehat{\boldsymbol{W}} = \boldsymbol{diag}\ (\widehat{\boldsymbol{w}}_1, \widehat{\boldsymbol{w}}_2, \dots, \widehat{\boldsymbol{w}}_n)$ and $\hat{\boldsymbol{z}} = \hat{\boldsymbol{\eta}} + \widehat{\boldsymbol{w}}^{-\boldsymbol{1}}\widehat{\boldsymbol{T}}\ (\mathbf{y}^* - \ \boldsymbol{\mu}^*)$.

$$\widehat{\boldsymbol{W}}_{\boldsymbol{i}} = \frac{\boldsymbol{1}-\hat{\delta}^2}{\hat{\delta}^2}\left[\psi^{\mathrm{T}}\left(\frac{(\omega_i^T(1-\hat{\delta}^2)}{\hat{\delta}^2}\right) + \psi^{\mathrm{T}}\left(\frac{(1-\omega_i^T)(1-\hat{\delta}^2)}{\hat{\delta}^2}\right)\right]\frac{\boldsymbol{1}}{\left(g^T(\widehat{\omega}_i)\right)^2}.$$

The estimated matrices $\boldsymbol{W}$ and T are denoted by $\widehat{\boldsymbol{W}}$ and $\widehat{\boldsymbol{T}}$, respectively, and evaluated at an ML estimator of the model. The MLE parameter estimates $\widehat{\omega}$ of the BRM approach converges to the true parameter as the sample size increases. Under regularity, the estimated parameter follows the condition:

$$E(\hat{\beta}_{MLE}) \to \beta \text{ as n} \to \infty.$$

The asymptotic covariance and the estimated MSE of the MLE estimates of the model are expressed as follows:

$$\sqrt{n}(\hat{\beta} - \beta) \overset{d}{\to} N(0, \frac{1}{\varphi}(\boldsymbol{X^T W X})^{-\boldsymbol{1}}).$$

Thus, the asymptotic covariance matrix is

$$Cov(\hat{\beta}_{MLE}) = \frac{1}{\varphi}(\boldsymbol{X^T W X})^{-\boldsymbol{1}}. \tag{2.11}$$

The total mean squared error (TMSE) is

$$\boldsymbol{TMSE}(\hat{\beta}_{MLE}) = Cov(\hat{\beta}_{MLE}) = tr\left\{\frac{1}{\varphi}(\boldsymbol{X^T W X})^{-\boldsymbol{1}}\right\} = \frac{\boldsymbol{1}}{\varphi}\sum_{\boldsymbol{j=1}}^{\boldsymbol{p}}\frac{\boldsymbol{1}}{\lambda_j}, \tag{2.12}$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p > 0$ are the eigenvalues of the $\boldsymbol{X^T W X}$.

2.3. Ridge-type biased estimators for Beta regression

To reduce the effects of multicollinearity, Altukhaes et al. [20] developed a shrinkage parameter to obtain reliable parameter estimates for the regression model. The researchers in [21] modified the ridge-biased parameter to better address the multicollinearity issue compared to other methods, thereby improving the efficiency of parameter estimates for the model.
Define $\boldsymbol{A} = \boldsymbol{X^T W X}$. Let $\mathbf{Q}$ be the orthogonal matrix with $\boldsymbol{A}$ such that $\boldsymbol{Q^T A\ Q} = \boldsymbol{\Lambda} = diag(\lambda_1, \dots, \lambda_p)$. Define $\gamma = \boldsymbol{Q^T \beta}$. Then $E(\hat{\gamma}_{MLE}) = \gamma$ and $Cov(\hat{\gamma}_{MLE}) = \frac{\boldsymbol{1}}{\varphi}\boldsymbol{\Lambda}^{-\boldsymbol{1}}$. For any estimator $\widehat{\boldsymbol{\beta}}$, TMSE is defined as $TMSE(\widehat{\boldsymbol{\beta}}) = \boldsymbol{tr}\left(\boldsymbol{Cov}(\widehat{\boldsymbol{\beta}})\right) + \left\|\boldsymbol{Bias}(\widehat{\boldsymbol{\beta}})\right\|^2$.

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}} = \left(\boldsymbol{A} + k\boldsymbol{I_p}\right)^{-\boldsymbol{1}}\boldsymbol{A}\ \widehat{\boldsymbol{\beta}}_{\boldsymbol{MLE}}, \quad k > 0. \tag{2.13}$$

In canonical form: $\hat{\gamma}_{RE} = (\boldsymbol{\Lambda} + \boldsymbol{kI})^{-1}\boldsymbol{\Lambda}\ \hat{\gamma}_{MLE}$. Then

$$\boldsymbol{E}(\hat{\gamma}_{\boldsymbol{RE}}) = (\boldsymbol{\Lambda} + \boldsymbol{kI})^{-1}\boldsymbol{\Lambda}\ \gamma.$$

Hence,

$$\boldsymbol{Bias}(\hat{\gamma}_{\boldsymbol{RE}}) = [(\boldsymbol{\Lambda} + k\boldsymbol{I})^{-1}\boldsymbol{\Lambda} - \boldsymbol{I}]\ \ \gamma.$$

$$Bias(\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}}) = -k\boldsymbol{Q}(\boldsymbol{\Lambda} + k\boldsymbol{I})^{-1}\gamma,\ k > 0. \tag{2.14}$$

Now, suppose that $\mathbf{Q}$ matrix is constructed such that its columns having of the eigenvectors of the $\boldsymbol{A}$ matrix and $\boldsymbol{A}$ matrix have eigenvalues $\lambda_1 \geq \lambda_2 \geq . \ \ . \ \ . \geq \lambda_p > 0$, then

$$\boldsymbol{\Lambda} = \boldsymbol{Q}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{Q} = diag(\lambda_1, \lambda_2, \ \ . \ \ \ ., \lambda_p),$$

$$\boldsymbol{Cov}(\hat{\gamma}_{\boldsymbol{RE}}) = (\boldsymbol{\Lambda} + \boldsymbol{kI})^{-1}\boldsymbol{\Lambda}\ .\frac{1}{\varphi}\ \ \boldsymbol{\Lambda}^{-1}.\boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \boldsymbol{kI})^{-1} = \frac{1}{\varphi}\ \ \ \boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \boldsymbol{kI})^{-2}.$$

Thus,

$$\boldsymbol{Cov}\left(\widehat{\widehat{\boldsymbol{\beta}}}_{\boldsymbol{RE}}\right) = \frac{1}{\varphi}\ \ \boldsymbol{Q}\ \ \boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \boldsymbol{kI})^{-2}\boldsymbol{Q}^{\boldsymbol{T}}.$$

Thus, the matrix mean squared error (MMSE) is written as follows:

$$MMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}}) = \boldsymbol{E}\left[(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})^{\boldsymbol{T}}\right] = \boldsymbol{Cov}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}}) + \boldsymbol{Bias}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}})\boldsymbol{Bias}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}})^{\boldsymbol{T}}$$

$$= \frac{1}{\varphi}\boldsymbol{Q}\ \ \boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \mathrm{k}\mathbf{I})^{-\mathbf{2}}\boldsymbol{Q}^{\boldsymbol{T}} + k^2\boldsymbol{Q}(\boldsymbol{\Lambda} + \mathrm{k}\mathbf{I})^{-\mathbf{1}}\ \ \gamma\ \ \gamma^T(\boldsymbol{\Lambda} + \mathrm{k}\mathbf{I})^{-\mathbf{1}}\boldsymbol{Q}^{\boldsymbol{T}}. \tag{2.15}$$

Taking the trace and using $tr(\boldsymbol{Q}^{\boldsymbol{T}}\boldsymbol{M}\boldsymbol{Q}) = \boldsymbol{tr}(\boldsymbol{M})$ gives the TMSE:

$$TMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{RE}}) == \frac{\mathbf{1}}{\varphi}\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}}\left[\frac{\lambda_j}{(\lambda_j + k)^2} + \frac{\gamma_j^2 k^2}{(\lambda_j + k)^2}\right]. \tag{2.16}$$

2.4. Liu-type estimators

Mansson et al. [22] modified the ridge bias estimators to obtain stable estimated parameters for the BRM in the presence of collinearity issues, which is referred to as the Liu ridge estimator. It is written as follows:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}} = (\boldsymbol{A} + \boldsymbol{I})^{-\mathbf{1}}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{z} + \boldsymbol{d}\widehat{\boldsymbol{\beta}}_{\boldsymbol{MLE}});\ \ 0 < d < 1, \tag{2.17}$$

$$\boldsymbol{Bias}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}}) = (\boldsymbol{A})^{-\mathbf{1}}(d - 1)^{-\mathbf{1}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{MLE}}\ \ \ \ , 0 < d < 1. \tag{2.18}$$

The MMSE can be expressed as:

$$MMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}}) = \boldsymbol{Cov}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}}) + Bias(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}})Bias(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}})^{\boldsymbol{T}}\ \ \widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}}$$

$$= \frac{1}{\varphi}\left[\boldsymbol{Q}\ \ \boldsymbol{\Lambda}_{\mathbf{L}}^{-\mathbf{1}}\boldsymbol{\Lambda}_{\mathbf{d}}\ \ \boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{\Lambda}_{\mathbf{d}}\ \ \boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{Q}^{\boldsymbol{T}}\right] + (1 - d)^2\boldsymbol{Q}\boldsymbol{\Lambda}_{\mathbf{L}}^{-\mathbf{1}}\ \ \gamma\ \ \gamma^{\boldsymbol{T}}\boldsymbol{\Lambda}_{\mathbf{L}}^{-\mathbf{1}}\boldsymbol{Q}^{\boldsymbol{T}}, \tag{2.19}$$

where $\boldsymbol{\Lambda}_{\mathbf{L}} = diag(\lambda_1 + 1, \lambda_2 + 1, \ldots, \lambda_p + 1)$ and $\boldsymbol{\Lambda}_{\mathbf{d}} = diag(\lambda_1 + d, \lambda_2 + d, \ldots, \lambda_p + d)$.
Thus, the TMSE for the Liu estimator is:

$$TMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}}) = Trace\left(MMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}})\right)$$

$$= \frac{\mathbf{1}}{\varphi}\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{p}}\left[\frac{(\lambda_j + d)^2}{\lambda_j(\lambda_j + 1)^2} + \frac{\gamma_j^{\mathbf{2}}\varphi(\boldsymbol{d} - \mathbf{1})^{\mathbf{2}}}{(\lambda_j + 1)^2}\right]. \tag{2.20}$$

The Liu type bias ridge estimators of Algamal and Abonazel [23] modified the Liu estimator for the BRM, defined as:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTRE}} = (\boldsymbol{A} + k\boldsymbol{I})^{-\mathbf{1}}(\boldsymbol{A} - d\boldsymbol{I})^{-\mathbf{1}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{MLE}},\ k > 0;\ \ -\infty < d < \infty. \tag{2.21}$$

The Liu-type estimator was originally proposed for the Poisson regression model, but its application generalizes to any generalized linear model (GLM) estimated via iteratively weighted least squares (IWLS). The shrinkage modifications depend only on $\mathbf{X}^T\mathbf{WX}$ and $\widehat{\boldsymbol{\beta}}_{\text{MLE}}$ quantities common to all GLMs and not on the specific distribution. Hence, the bias and TMSE expressions Eqs (2.18)–(2.20) and (2.22) hold for Beta regression, with λ_j as eigenvalues of $\mathbf{X}^T\mathbf{WX}$ at the Beta MLE.

The Modified Liu type bias ridge estimator vector is followed as:

$$\boldsymbol{Bias}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTRE}}) = -(d + k)(\boldsymbol{V} + k\boldsymbol{I})^{-\mathbf{1}}\boldsymbol{\beta},$$

where $\boldsymbol{\Lambda}_{-\mathbf{d}} = diag(\lambda_1 - d, \lambda_2 - d, \ldots, \lambda_p - d)$.

The MMSE of the modified Liu type bias ridge estimator is:

$$TMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTRE}}) = Trace\left(MMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{LTRE}})\right)$$

$$= \frac{\mathbf{1}}{\varphi}\sum_{\boldsymbol{j=1}}^{\boldsymbol{p}}\left[\frac{(\lambda_j - d)^2}{\lambda_j(\lambda_j + k)^2} + \frac{\gamma_j^2\varphi(d+k)^2}{(\lambda_j + k)^2}\right]. \tag{2.22}$$

2.5. Two-parameter ridge estimators

The two-parameter ridge-types estimators were initially modified in [23] for the BRM. The estimators mitigate the multicollinearity issues and improve the model parameter estimation accuracy. The two-parameter ridge regression model expression is:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{TRE}} = (\boldsymbol{A} + k\boldsymbol{I})^{-\mathbf{1}}(\boldsymbol{A} + d\boldsymbol{I})^{-\mathbf{1}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{MLE}} \quad ,k > 0;\ \ 0 < d < 1. \tag{2.23}$$

$$\boldsymbol{Bias}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{TRE}}) = k(d - 1)(\boldsymbol{A} + k\boldsymbol{I})^{-\mathbf{1}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{MLE}} \quad ,0 < d < 1. \tag{2.24}$$

The MMSE can be expressed as:

$$MMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{TRE}}) = \boldsymbol{Cov}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{TRE}}) + \boldsymbol{Bias}(\widehat{\boldsymbol{\beta}}_{\boldsymbol{TRE}})\boldsymbol{Bias}(\widehat{\boldsymbol{\beta}}_{TRE})^T\ \ \widehat{\boldsymbol{\beta}}_{\boldsymbol{LRE}}$$

$$= \frac{1}{\varphi}\left[\boldsymbol{Q}\ \ \boldsymbol{\Lambda}_{\mathbf{k}}^{-\mathbf{1}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}_{\mathbf{kd}}\boldsymbol{Q}^{\boldsymbol{T}}\right] + k^2\boldsymbol{Q}\boldsymbol{\Lambda}_{\mathbf{k}}^{-\mathbf{1}}\ \gamma\ \ \gamma^{\boldsymbol{T}}\boldsymbol{\Lambda}_{\mathbf{k}}^{-\mathbf{1}}\boldsymbol{Q}^{\boldsymbol{T}}, \tag{2.25}$$

where $\boldsymbol{\Lambda}_{\mathbf{kd}} = diag(\lambda_1 + kd, \lambda_2 + kd, \ldots, \lambda_p + kd)$

Thus, the TMSE is:

$$TMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{TRE}}) = Trace\left(MMSE(\widehat{\boldsymbol{\beta}}_{\boldsymbol{TRE}})\right)$$

$$= \frac{\mathbf{1}}{\varphi}\sum_{\boldsymbol{j=1}}^{\boldsymbol{p}}\left[\frac{(\lambda_j + kd)^2}{\lambda_j(\lambda_j + k)^2} + \frac{\gamma_j^2\varphi\boldsymbol{k}^{\mathbf{2}}(\boldsymbol{d}-\mathbf{1})^{\mathbf{2}}}{(\lambda_j + k)^2}\right]. \tag{2.26}$$

2.6. Overview of existing and newly proposed ridge type bias estimators

Initially, the authors in [1] introduced the ridge estimator to address the collinearity issues and obtained reliable parameter estimates of the BRM. It is defined as:

$$k_1 = \frac{1}{max(\hat{\beta}_j)\ )}, \tag{2.27}$$

where k_1 is referred to as the LRR estimator in this study.

Asar and Genc [24] developed the ridge-type bias estimator of estimates the parameters of the Poisson regression model to addressed the multicollinearity issues, referred to as LLE in this study, and it is defined as:

$$d_1 = \ min\left(\frac{\lambda_j}{1+\lambda_i\ \hat{\beta}_j}\right). \tag{2.29}$$

In this work, we propose two-parameter ridge-type estimators (MTRT) for the BRM to mitigate multicollinearity more effectively compared to existing estimators, and our proposed estimators are referred to as MTRT estimators, defined as:

$$k_2 = \sqrt{\sum_{j=1}^{p}\left(\frac{\lambda_j}{1+\frac{1}{\min(\lambda_j)}}\right)}\ \times \frac{1+\kappa(X)}{\sqrt{p}}, \text{ with} \tag{2.30}$$

$$d_2 = \frac{\sum_{j=1}^{p}\frac{\lambda_j}{1+\lambda_j}}{p\times(1+\kappa(X))}. \tag{2.31}$$

Here, $\kappa(X) = \frac{\lambda_{max}}{\lambda_{min}}$ is the condition number of $\boldsymbol{X^TWX}$, p is the number of predictors, λ_{max} is the maximum eigenvalue, and λ_{min} is the minimum eigenvalue. $k_2 > 0$ and $0 < d_2 < 1$ hold under conditions.

3. Monte Carlo simulation

In this section, we conduct a simulation study to compare the existing and proposed estimators' efficiency based on the estimated MSE creation.

3.1. Simulation technique

The predictors are generated using the following equation, as employed by other researchers [25,26]:

$$xij = (1-\rho^2)^{0.5}Uij + \rho Ui,p+1\ \ ;\ j = 1,2,\ldots,p \text{ and } i = 1,2,\ldots,n. \tag{3.1}$$

In Eq (3.1), $Uij \sim N(0,1)$. This yields a multivariate normal distribution with mean zero and a Toeplitz covariance structure, $\mathrm{Cov}(X_i, X_j) = \rho^{|i-j|}$(with $\rho = 0.5$), enabling correlations to decay with distinct between predictors. Four correlation levels are considered: $\rho = 0.80, 0.90, 0.95, 0.99$, representing moderate to severe multicollinearity.

The true coefficient vector β is set such that $\sum_{j=1}^{p}\beta_j^2 = 1$ and $\beta\ = \left(\beta_1, \beta_2, \ldots, \beta_p\right)^T$. The linear predictor $\boldsymbol{\eta_i} = \boldsymbol{X_i^T}\beta$ and the mean of the Beta distribution is defined as

$$\mu_i = \frac{\exp(\boldsymbol{\eta_i})}{1+\exp(\boldsymbol{\eta_i})}. \tag{3.3}$$

The response variable is simulated from a Beta distribution conditional for the mean μ_i and precision φ:

$$y_i \sim Beta(\alpha_i, \beta_i), \text{ with } \alpha_i = \mu_i\varphi,\ \beta_i = (1-\mu_i)\varphi.$$

We consider precision parameters $\varphi = 2$ and 4. The number of the Monte Carlo simulation is $R = 1000$. The Estimated MSE is obtained as follows:

$$MSE(\widehat{\boldsymbol{\beta}}) = \frac{1}{R}\sum_{i=1}^{R}(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})^T(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}).$$

The simulations are performed for different sample sizes $n = 20, 30, 50,$ and 100 and numbers of predictors ($p = 4, 6, 10,$ and 50). The estimated MSEs of the estimators are presented in Tables 1–4.

Table 1. Estimated MSEs of the estimators for $n = 20$ and $4, 6, 10$.

$\boldsymbol{n}$	$\boldsymbol{p}$	$\boldsymbol{\rho}$	$\boldsymbol{\varphi}$	**MLE**$(\boldsymbol{k_1} = \mathbf{0})$	**LRR** $(\boldsymbol{k_1})$	**LLE**$(\boldsymbol{d_1})$	MTRTE $(\boldsymbol{k_2}, \boldsymbol{d_2})$
20	4	0.80	2	26.6831847	14.4023879	2.26237915	0.71882547
		0.90		35.3055844	14.8457211	3.70021618	1.21394162
		0.95		72.7939754	20.5619458	6.54337935	1.01694099
		0.99		458.13819	57.1440224	25.8354487	0.99094186
		0.80	4	13.188359	7.56440113	2.00584681	1.35678911
		0.90		35.954689	13.7661816	3.30340563	0.67826487
		0.95		65.2872992	17.605764	6.47748194	1.25088751
		0.99		192.626599	14.5617592	17.6235419	1.01287574
	6	0.80	2	58.5559461	21.689451	3.79634733	0.87771309
		0.90		73.1575345	21.6625306	5.94543375	0.95623928
		0.95		132.309495	21.524576	9.76771287	1.01666448
		0.99		1194.19987	86.2340024	43.0601436	0.99204517
		0.80	4	23.7130184	10.1422458	2.93190248	1.14261296
		0.90		50.4630909	12.7397395	4.16212414	0.91841341
		0.95		84.9973126	13.4964374	8.65175725	0.99807772
		0.99		390.470206	15.5367069	17.8121932	0.99892164
	10	0.80	2	101.8566	36.8589392	6.47858501	1.02485066
		0.90		216.925663	48.3232564	9.34282354	0.98939901
		0.95		879.073008	126.882528	17.6175312	0.98343692
		0.99		2322.50808	105.698294	41.727555	0.99855703
		0.80	4	52.0298452	14.8936751	5.49587542	1.0437408
		0.90		101.051507	24.9635248	7.53911612	1.01318367
		0.95		275.230021	36.8873602	10.2247584	0.98286378
		0.99		1330.66371	61.7580225	28.8608176	1.00272643

Table 2. Estimated MSEs of the estimators for $n = 30$ and $4, 6,10$.

$\boldsymbol{n}$	$\boldsymbol{p}$	$\boldsymbol{\rho}$	$\boldsymbol{\varphi}$	**MLE**$(\boldsymbol{k_1 = 0})$	**LRR** $(\boldsymbol{k_1})$	**LLE**$(\boldsymbol{d_1})$	MTRTE $(\boldsymbol{k_2}, \boldsymbol{d_2})$
30	4	0.80	2	15.7550367	11.6498951	1.49702875	1.81002333
		0.90		29.6345393	17.9207236	2.81047675	1.32433792
		0.95		64.3209425	16.257827	4.14076022	0.94382219
		0.99		323.139624	57.055393	22.4738448	0.96454536
		0.80	4	12.5327528	7.904778	1.29784896	1.18454924
		0.90		13.4416031	3.8623196	2.57594477	0.9334279
		0.95		25.9502575	5.81676265	4.27330946	0.94458608
		0.99		151.945704	20.4925658	10.7943287	1.05715455
	6	0.80	2	19.0800763	8.297069	2.63952779	1.07956238
		0.90		41.9305642	15.6357304	4.5219737	1.11985941
		0.95		82.0797897	21.7269065	6.47920033	0.95249649
		0.99		386.46349	32.0532191	18.1160128	1.00385489
		0.80	4	16.4701639	5.97239328	2.33825026	0.81308944
		0.90		24.062802	7.93275221	3.21638268	0.90125284
		0.95		36.8869551	10.9113126	5.48202699	1.03345928
		0.99		209.182126	13.1787576	13.6212021	1.00004566
	10	0.80	2	65.174264	26.585506	4.40737061	1.23948737
		0.90		91.2605176	29.4122479	6.72874762	0.97341407
		0.95		193.256065	47.8316866	10.4605623	1.01773424
		0.99		1129.95831	84.8621217	26.7699362	1.00558081
		0.80	4	26.5606174	11.8842667	3.71431657	0.96053271
		0.90		44.3115243	13.575066	5.34106086	1.01034298
		0.95		107.402176	22.2098236	7.20861457	0.97157531
		0.99		538.362015	34.7922704	16.135746	0.9975552

Table 3. Estimated MSEs of the estimators for $n = 50$ and $4, 6, 10, 50$.

n	p	ρ	φ	MLE$(k_1 = 0)$	LRR (k_1)	LLE(d_1)	MTRTE (k_2, d_2)
50	4	0.80	2	7.2729099	5.04090415	1.02300507	1.05191788
		0.90		17.068314	13.548379	1.63193965	1.54575441
		0.95		21.8935266	8.73922055	3.13484482	0.96914534
		0.99		94.6473871	11.0199058	10.7263197	1.00186703
		0.80	4	10.6150021	4.88352103	0.79662668	0.39940703
		0.90		10.1170542	7.51663686	2.00349392	1.4020165
		0.95		17.6902334	9.20493432	2.51576588	1.2070408
		0.99		74.335078	20.3129472	9.55902529	1.06447094
	6	0.80	2	17.8988587	10.005479	1.3282665	0.62390353
		0.90		28.430637	22.8358192	2.54457994	1.32524512
		0.95		48.2094716	22.9594853	4.12385417	1.15895811
		0.99		278.288454	43.8530339	13.9475184	0.9814267
		0.80	4	12.5834774	3.88813681	1.43188056	0.64346911
		0.90		20.181274	10.7521346	2.53597918	1.3146343
		0.95		22.9352111	8.68842423	4.35328356	1.04613447
		0.99		165.272178	25.1944104	8.15327163	1.02755069
	10	0.80	2	23.0343184	12.3818998	2.57886865	1.00736162
		0.90		47.1705387	22.2031833	4.78035852	0.92432693
		0.95		94.2856761	34.8419944	6.87683097	1.04089811
		0.99		433.948959	50.1955119	13.7946248	1.00499012
		0.80	4	15.9392972	9.73720533	2.10271629	1.06795372
		0.90		25.6347223	10.6808246	3.76622851	0.99597135
		0.95		41.0460843	9.91642598	5.62625645	0.99836367
		0.99		266.118721	26.8947922	11.3432954	0.99493142
	50	0.80	2	29.6042101	23.0392102	3.90068678	1.08934526
		0.90		57.2078912	37.0630892	10.7835852	0.94500389
		0.95		104.809420	65.841994	17.9463011	1.89035261
		0.99		529.450569	68.4578012	24.8632230	1.73015910
		0.80	4	35.0039872	9.90402341	2.92001672	1.04036780
		0.90		45.3782793	17.4903221	8.09356721	0.99045672
		0.95		56.5049321	12.9940028	10.7803627	0.99256026
		0.99		322.509812	46.2309712	22.6720341	0.997356212

Table 4. Estimated MSEs of the Estimators for $n = 100$ and $4, 6, 10, 50$.

$\boldsymbol{n}$	$\boldsymbol{p}$	$\boldsymbol{\rho}$	$\boldsymbol{\varphi}$	**MLE**$(\boldsymbol{k_1 = 0})$	**LRR** $(\boldsymbol{k_1})$	**LLE**$(\boldsymbol{d_1})$	MTRTE $(\boldsymbol{k_2}, \boldsymbol{d_2})$
100	4	0.80	2	6.20431415	4.6352606	0.61665943	1.41524298
		0.90		8.3137785	2.66044831	1.08087556	0.72475188
		0.95		12.010263	5.80089611	1.7416001	0.8921656
		0.99		49.7716925	12.6598857	5.60111701	1.01223798
		0.80	4	7.96301006	9.66786082	0.47038921	1.93477035
		0.90		10.4362549	4.47163406	0.80454824	0.4697328
		0.95		12.5196925	5.10250336	1.53754939	0.71782417
		0.99		40.3401932	12.6039829	5.84732692	1.05337109
	6	0.80	2	8.35616229	4.48163462	0.9266536	1.41396964
		0.90		13.6890351	9.93135221	1.49156341	1.39393526
		0.95		29.1063297	9.19891871	2.90414044	0.91892774
		0.99		82.2204282	16.8268055	7.91501424	1.0051155
		0.8	4	10.6076394	7.72542332	0.6563953	0.87096148
		0.9		9.84172416	5.69489697	1.388497	1.17883261
		0.95		16.1373471	14.2307641	2.40921731	1.07748809
		0.99		78.2540816	16.3068802	7.02347474	0.96597222
	10	0.80	2	12.360563	4.88261399	1.37613999	1.08744172
		0.90		20.0817063	11.5073724	2.66765337	1.0472057
		0.95		29.2281038	11.5857684	4.11731417	1.01895202
		0.99		170.893074	33.9596321	12.2267072	1.00336043
		0.80	4	9.2398478	4.55801019	1.32215704	0.88593602
		0.90		13.6412452	6.01836555	2.34601427	1.12782515
		0.95		25.5994966	11.5660212	3.4915085	0.9931443
		0.99		92.9268104	17.56461	8.00697771	0.99662448
	50	0.80	2	12.3892383	8.26103782	2.70379213	1.20947812
		0.90		28.9830713	19.4038921	4.53782901	1.70468231
		0.95		56.8390201	27.9045629	7.30267183	1.19064738
		0.99		225.490236	43.7048921	20.8976154	1.82346278
		0.80	4	10.3902781	7.90456731	2.00198732	0.89745623
		0.90		18.8903810	10.6703710	3.3802178	1.13670512
		0.95		37.1890389	15.4035198	5.84036721	0.98402674
		0.99		138.903578	23.0036781	11.6067832	0.996408342

3.2. Discussion of the simulation results

In this study, we compare the newly proposed MTRT estimator with the MLE, LRR, and LLE methods using the MSE as the criterion, across sample sizes, numbers of predictor variables, correlations, and precision terms. Below are some remarks from the simulations.

i. **Effect of Sample Size:** As the sample size (n) increases, the MSEs of all estimators generally decrease. MTRTE consistently demosntrates lower MSE compared to MLE, LRR, and LLE, especially for larger sample sizes $(n = 50, 100)$.

ii. **Effect of Predictors:** With a higher number of predictor variables (p), the MSEs increases, indicating that all estimators are more sensitive to the number of predictors. MTRT, however, remain relatively stable across values of p.

iii. **Effect of Correlations and Precision:** Increasing correlations (ρ) and precision parameter (φ) tends to inflate the MSEs, particularly for MLE. MTRTE shows more robustness in handling higher correlations and precision terms, with relatively smaller MSEs compared to other methods, especially under high correlation and precision conditions.

Overall, the MTRT estimator outperforms the other methods across scenarios, demonstrating its effectiveness in reducing MSE, particularly when multicollinearity is present.

4. Practical applications

In this section, we use an empirical example to assess the performance of the proposed estimators relative to MLE and other estimators. The analysis is based on the Spanish La Liga football season (2016–2017) dataset described in [23]. The dataset covers 12 teams, and the outcome of interest is each team's proportion of matches won. Six covariates are included in the model: Counts of yellow cards, red cards, and substitutions, along with measures related to scoring matches with an average of 2.5 goals, matches that end with goals, and the ratio of goals scored to total matches.

Since strong dependence among predictors can reduce the stability and reliability of coefficient estimates, multicollinearity is investigated using standard diagnostic tools: the correlation matrix, variance inflation factors (VIF), and the condition number.

Figure 1 shows that the correlation matrix indicates several predictors that are highly correlated with one another, indicating notable linear dependence among the covariates. The condition number, computed as the largest-to-smallest eigenvalue ratio, is 806.63 [23]. Such a large value is well above commonly used benchmarks and signals substantial multicollinearity in the dataset.

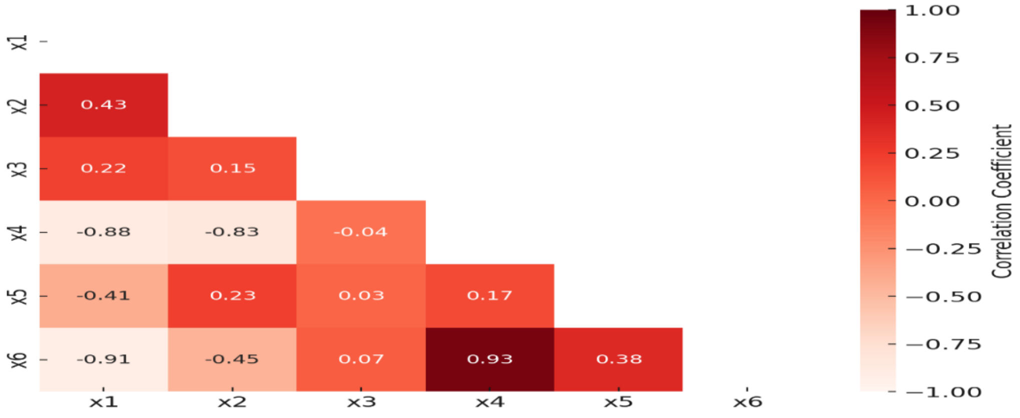

Figure 1. Correlation Matrix of the Spanish Liga Football Season (2016–2017).

Table 5 summarizes the estimated MSEs for the MLE and the alternative estimators considered. The results show that the proposed estimators yield smaller estimated MSEs than MLE and improve upon the ridge and Liu-type estimators, supporting their practical advantage when predictors are highly correlated.

Table 5 demonstrates that the results from the practical application are consistent with the findings observed in the simulation study.

Table 5. Estimated MSEs and coefficients of the estimators.

Variable	MLE($k_1 = 0$)	LRR (k_1)	LLE(d_1)	MTRTE (k_2, d_2)
MSE	1.3071	1.0045	0.9415	0.7108
$\hat{\beta}_0$	1.3083	1.1504	1.1290	0.8703
$\hat{\beta}_1$	–0.0189	-0.01871	-0.02191	-0.0019
$\hat{\beta}_2$	0.22405	0.214399	0.20097	0.10354
$\hat{\beta}_3$	0.04284	0.030276	0.00947	0.00701
$\hat{\beta}_4$	–0.06514	-0.06113	0.03908	0.01060
$\hat{\beta}_5$	–0.16877	-0.1592	-0.0962	0.00195
$\hat{\beta}_6$	1.84197	1.82706	0.67013	0.01984

5. Conclusions

In this study, we proposed a MTRT estimator for the BRM to alleviate multicollinearity and compare its performance with the MLE and other commonly used estimators, including LRR and LLE, using the estimated MSE as the evaluation criterion. Based on simulation experiments and real-data analysis, the MTRT estimator generally performed as well as or better than the MLE, LRR, and LLE under certain levels of multicollinearity.

While the MTRT estimator often yielded the lowest estimated MSE, its advantage was not uniform across all combinations of sample sizes, number of predictors, and correlation levels. A common challenge for ridge-type methods remains the selection of an optimal penalty parameter to balance bias-variance trade-off, as no single ridge estimator consistently performs best in all conditions. Nevertheless, the MTRT estimator demonstrated favorable performance in most scenarios, and application to the football dataset further confirmed its practical effectiveness with notably lower MSE. Overall, the MTRT estimator is recommended as a robust and efficient option for parameter estimation in Beta regression when multicollinearity is present.

Author contributions

Muteb Faraj Alharthi: Conceptualization, methodology, formal analysis, software, validation, visualization, writing original draft, writing review and editing; Nadeem Akhtar: Conceptualization, methodology, investigation, data curation, software, validation, supervision, project administration, writing review and editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors confirm that no Artificial Intelligence (AI) tools were used in the development of this article.

Funding

Deanship of Graduate Studies and Scientific Research, Taif University.

Conflict of interest

The authors declare that they have no conflict of interest regarding this research.

References

1. A. E. Hoerl, R. W. Kennard, Ridge regression: Applications to nonorthogonal problems, *Technometrics*, **12** (1970a), 69–82. https://doi.org/10.1080/00401706.1970.10488635
2. A. E. Hoerl, R. W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, **12** (1970b), 55–67, 1970, https://doi.org/10.1080/00401706.1970.10488634
3. M. Amin, M. N. Akram, B. M. G. Kibria, H. M. Alshanbari, N. Fatima, A. Elhassanein, On the estimation of the Binary response model, *Axioms*, **12** (2023), 175. https://doi.org/10.3390/axioms12020175
4. A. F. Lukman, K. Ayinde, A. S. Ajiboye, Monte Carlo study of some classification-based ridge parameter estimators, *J. Mod. Appl. Stat. Meth.*, **16** (2017), 428–451. https://doi.org/10.22237/jmasm/1493598240
5. Y. Li, R. Yang, X. Wang, J. Zhu, N. Song, Carbon price combination forecasting model based on Lasso regression and optimal integration, *Sustainability*, **15** (2023), 9354. https://doi.org/10.3390/su15129354
6. A. Zeileis, C. Kleiber, S. Jackman, Regression models for count data in R, *J. Stat. Softw.*, **27** (2008). https://doi.org/10.18637/jss.v027.i08
7. H. S. Norman R. Draper, *Applied regression analysis*, 3 Eds., WILLEY, 1998.
8. R. Cavoretto, Adaptive radial basis function partition of unity interpolation: A bivariate algorithm for unstructured data, *J. Sci. Comput.*, **87** (2021), 41. https://doi.org/10.1007/s10915-021-01432-z
9. R. Cavoretto, A. De Rossi, An adaptive residual sub-sampling algorithm for kernel interpolation based on maximum likelihood estimations, *J. Comput. Appl. Math.*, **418** (2023), 114658. https://doi.org/10.1016/j.cam.2022.114658
10. M. Qasim, K. Månsson, B. M. G. Kibria, On some beta ridge regression estimators: Method, simulation and application, *J. Stat. Comput. Sim.*, **91** (2021), 1699–1712. https://doi.org/10.1080/00949655.2020.1867549
11. S. Seifollahi, H. Bevrani, James-Stein type estimators in beta regression model: Simulation and application, *Hacet. J. Math. Stat.*, **52** (2023), 1046–1065. https://doi.org/10.15672/hujms.1122207
12. A. Erkoc, E. Ertan, Z. Y. Algmal, K. U. Akay, The beta Liu-type estimator: Simulation and application, *Hacet. J. Math. Stat.*, **52** (2023), 828–840. https://doi.org/10.15672/hujms.1145607
13. Y. Asar, M. Arashi, J. Wu, Restricted ridge estimator in the logistic regression model, *Commun. Stat.-Simul. C.*, **46** (2017), 6538–6544. https://doi.org/10.1080/03610918.2016.1206932
14. D. Dai, D. Wang, A generalized Liu-type estimator for logistic partial linear regression model with multicollinearity, *AIMS Math.*, **8** (2023), 11851–11874. https://doi.org/10.3934/math.2023600
15. A. T. Hammad, I. Elbatal, E. M. Almetwally, M. M. A. El-Raouf, M. A. El-Qurashi, A. M. Gemeay, A novel robust estimator for addressing multicollinearity and outliers in Beta regression: Simulation and application, *AIMS Math.*, **10** (2025), 21549–21580. https://doi.org/10.3934/math.2025958
16. A. A. Alzahrani, Modified two-parameter ridge estimator for the Beta logistic model to mitigate multicollinearity, *AIMS Math.*, **11** (2026), 543–557. https://doi.org/10.3934/math.2026023

17. S. Ferrari, F. C. Neto, Beta regression for modelling rates and proportions, *J. Appl. Stat.*, **31** (2004), 799–815. https://doi.org/10.1080/0266476042000214501
18. P. L. Espinheira, S. L. P. Ferrari, F. C. Neto, On beta regression residuals, *J. Appl. Stat.*, **35** (2008), 407–419. https://doi.org/10.1080/02664760701834931
19. G. H. A. Pereira, On quantile residuals in beta regression, *Commun. Stat.-Simul. C.*, **48** (2019), 302–316. https://doi.org/10.1080/03610918.2017.1381740
20. W. B. Altukhaes, M. Roozbeh, N. A. Mohamed, Feasible robust Liu estimator to combat outliers and multicollinearity effects in restricted semiparametric regression model, *AIMS Math.*, **9** (2024), 31581–31606. https://doi.org/10.3934/math.20241519
21. R. A. Farghali, M. Qasim, B. M. G. Kibria, M. R. Abonazel, Generalized two-parameter estimators in the multinomial logit regression model: Methods, simulation and application, *Commun. Stat.-Simul. C.*, **52** (2023), 3327–3342. https://doi.org/10.1080/03610918.2021.1934023
22. K. Månsson, B. M. G. Kibria, G. Shukur, On Liu estimators for the logit regression model, *Econ. Model.*, **29** (2012), 1483–1488. https://doi.org/10.1016/j.econmod.2011.11.015
23. Z. Y. Algamal, M. R. Abonazel, Developing a Liu-type estimator in beta regression model, *Concurr. Comp.-Pract. E.*, **34** (2022). https://doi.org/10.1002/cpe.6685
24. Y. Asar, A. Genç, A new two-parameter estimator for the Poisson regression model, *Iran. J. Sci. Technol.*, **42** (2018), 793–803. https://doi.org/10.1007/s40995-017-0174-4
25. Z. Y. Algamal, M. R. Abonazel, A. F. Lukman, Modified Jackknife ridge estimator for Beta regression model with application to chemical data, *Int. J. Math. Stat. Comput. Sci.*, **1** (2023), 15–24. https://doi.org/10.59543/ijmscs.v1i.7713
26. N. Akhtar, M. F. Alharthi, Enhancing accuracy in modelling highly multicollinear data using alternative shrinkage parameters for ridge regression methods, *Sci. Rep.*, **15** (2025), 1–13. https://doi.org/10.1038/s41598-025-94857-7