Economic Computation and Economic Cybernetics Studies and Research, Issue 4/2015

Mohammed Alguraibawi*, PhD
E-mail: mohammedam23@yahoo.com
Professor Habshah Midi, PhD
E-mail: habshahmidi@gmail.com
Lecturer Sohel Rana, PhD
E-mail: sranastat@yahoo.com
Faculty of Science and Institute for Mathematical Research
43400 UPM, Serdang, Selangor, Malaysia
***ATU, Al-Dewanyia Technical Institute, Iraq**

ROBUST JACKKNIFE RIDGE REGRESSION TO COMBAT MULTICOLLINEARITY AND HIGH LEVERAGE POINTS IN MULTIPLE LINEAR REGRESSIONS

Abstract. *In this paper, we modified the ordinary jackknife ridge regression (JRR) to be more resistant for high leverage points (outliers in X direction). The procedure for modification is by combining the JRR with a high breakdown point and high efficiency robust methods such as MM-estimator and modified generalized M-estimator (GM2). Here, two methods are suggested, the first one is JRR based on the MM-estimator (JRMM) and the second is JRR based on the GM2-estimator (JRGM2). The biggest advantages of the proposed methods are that they have less bias and higher efficiency than existing methods to overcome the combined problem of multicollinearity and high leverage points.*

Keywords*: Multicollinearity, outliers, high leverage point, ridge regression, Jackknife ridge regression, MM-estimator, GM2-estimator.*

JEL Classification: 62J05; 62J07

1. Introduction

Multicollinearity is a major problem in multiple regression, this issues occurs when two or more regressors are highly correlated. There are many bad consequences for the problem of multicollinearity, such as increased standard error and decreased reliability of coefficients, and often the results are confusing and give misleading conclusions [Groβ (2003)]. Another important problem in regression analysis is the existence of outliers. Rousseeuw and Leory (1987) defined an outlier as an observation that seems inconsistent with the bulk of the data. Hekimoglu and Erenoglu (2013) classified outliers into three categories namely, outliers in the *Y*-direction (small outliers), gross errors in the *Y*-direction (vertical outliers) and leverage points in the *X*-direction. Outliers also have significant impacts in regression, such as causing model failure and misleading conclusions. Under Gaussian Markov assumptions, the Ordinary Least Squares (OLS) estimation method is widely used in multiple linear regression because it has excellent properties and simplifies the computation. However, in the presence of multicollinearity and outliers, the OLS estimators become very unstable and may have large variance [Rousseeuw and Leroy (2003)], which leads to poor predictions.

To address this problem of multicollinearity, one alternative approach is ridge regression (RR), which was introduced by Hoerl and Kenard (1970). Although RR has optimal properties in the case of multicollinearity, its estimators are significantly biased [see, Batah et al. (2008), Esra and Fikri (2012)]. Singh et al. (1986) suggested an almost unbiased ridge estimator depending on the Jackknife technique to reduce the biasness in RR estimators.

They showed that the Jackknife ridge estimator has smaller bias and lower mean square error (MSE) than the classical RR under some conditions. However, RR and the Jackknife techniques are not robust to outliers and leverage point. As a remedial technique, many robust methods have been proposed [Huber (2003, Maronna (2006)], such as the least median of squares, the M-estimator, the MM-estimator and the generalized M (GM-estimator). Unfortunately, neither robust methods nor the RR technique alone is sufficient to address the complicated problem of multicollinearity and outliers [see, Habshah and Marina (2007)]. To circumvent this combined problem, significant works have been done by integrating RR with the robust method to get an estimator that is much less influenced by multicollinearity and unusual data.

Askin and Montgomery (1980) suggested using weighted RR to remedy this complicated problem. Habshah and Marina (2007) suggested a new robust ridge regression by incorporating RR with the MM-estimator to remedy the problem of multicollinearity in the present of outliers. Jadhav and Kashid (2011) suggested using a Jackknife ridge M-estimator to overcome multicollinearity and outliers in the Y direction. However most of the suggested methods do not focus on the combined problem of multicollinearity and high leverage points (HLPs) when the outliers lie in the X direction. In this article, we propose to integrate Jackknife Ridge Regression (JRR) with two robust methods, namely the MM-estimator and the GM2-estimator, to overcome the multicollinearity and HLPs and to obtain estimators that are much less biased than robust RR estimators.

The paper is organized as follows: Section 2 presents the model and briefly explains it for OLS method. Section 3 gives the structure and estimators of generalized ridge regression. Section 4 gives the structure and estimators of Jackknife ridge regression. Section 5 presents the procedure for GM2-estimator and briefly explained some robust methods. The estimators, bias, and variance for the suggested method are given in Section 6. Section 7 presents the simulation study. The discussion is presented in Section 8. Finally, Section 9 gives some concluding remakes.

2. Models and Estimators

Consider the following standard multiple linear regression model:

$$y = X\beta + u \qquad (1)$$

where it is assumed that y is an ($n \times 1$) vector of the dependent variable, X is an ($n \times p$) and full rank matrix of regressor variables, β is a ($p \times 1$) vector of an unknown regression parameters and u is an ($n \times 1$) vector of the error term with elements are assumed to be independently and identically normally such that $E(u) = 0$ and the dispersion matrix $E(uu') = \sigma^2 I$. For the purpose of convenience, it is assumed that all variables are standardized so that the design matrix $X'X$ ($'$ denotes transpose) is in correlation form. The OLS estimator, namely

$$\hat{\beta}_{LS} = (X'X)^{-1}X'y \qquad (2)$$

has optimal properties under Gaussian-Markov assumptions. Let $\Lambda = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p)$ be the matrix of eigenvalues for $X'X$ and γ is a ($p \times p$) matrix of corresponding eigenvectors whose columns are normalized. According to Singh et al. (1986), the equivalent formula for (1) with canonical form is

$$y = Z\alpha + u \qquad (3)$$

where $Z = X\gamma$ and $\alpha = \gamma'\beta$; hence, $\Lambda = Z'Z = \gamma'X'X\gamma$. The OLS estimator for α is given by

$$\hat{\alpha}_{LS} = (Z'Z)^{-1}Z'y = \Lambda^{-1}Z'y \qquad (4)$$

since $\alpha = \gamma'\beta$, then $\hat{\beta}_{LS}$ can be written as:

$$\hat{\beta}_{LS} = \gamma\,\hat{\alpha}_{LS} \qquad (5)$$

The MSE for the OLS estimator is given by:

$$MSE(\hat{\beta}_{LS}) = MSE(\hat{\alpha}_{LS}) = \sigma^2\Lambda^{-1} \qquad (6)$$

Hoerl and Kennard showed that a solution to the OLS does not always exist and there is no unique solution when the matrix $X'X$ is ill-conditioned (not invertible) due to the multicollinearity problem. In this situation the OLS estimators tend to become very large and may have large variance.

3. Generalized Ridge Regression

In order to handle the difficulties of OLS, Hoerl and Kennard (1970) proposed a biased estimation method called Generalized Ridge Regression (GRR) as an alternative technique in the case of the existence of multicollinearity. This suggested technique is based on adding some bias into the estimators to reduce their variance. They showed that the quantity $X'X + K$, where K is a positive constant, unlike $X'X$ in (2), is always invertible, so there is always a unique solution in GRR. Here, $K = \text{diag}(k_1, k_2, ..., k_p),\ k_i \geq 0\ \ , i = 1,2,..., p,$ is called the ridge or shrinkage parameter. The generalized ridge regression estimate is obtained by minimizing the penalized sum of squares:

$$\sum_{i=1}^{n}(y_i - x_i'\beta)^2 + k_i \sum_{j=1}^{p}\beta_j^2 \qquad (7)$$

In (7) we can see that GRR penalizes the size of the regression coefficients to be more resistant to multicollinearity. When $k_1 = k_2 = \cdots = k_p = k, k > 0$, k is fixed, the solution of GRR, namely ordinary RR, is given by

$$\hat{\beta}_{RR} = (X'X + k\,I_n)^{-1} X'y \qquad (8)$$

Several methods of identifying k have been proposed in the literature. The most popular approach for choosing the optimal k was suggested by Horel and Kennard (1970) as follows:

$$k = k_{HK} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2} \qquad (9)$$

where $\hat{\sigma}^2$ and $\hat{\alpha}$ are obtained by using OLS.

If k is equal to zero, $\hat{\beta}_{LS}$ and $\hat{\beta}_{RR}$ are equivalent. The shrinkage parameter k has the impact of shrinking the estimates toward zero, which leads to the introduction of bias but reduces the variance of the estimate [Horel and Kennard (1970)]. The canonical form of the RR estimator in (8) is given by [see, Bastlevsky (1994), Batah et al. (2008), Esra and Fikri (2012)]

$$\begin{aligned} \hat{\alpha}_{RR}(k) &= (\Lambda + kI_p)^{-1} Z'y \\ &= B^{-1}Z'y \\ &= (I - kB^{-1})\,\hat{\alpha} \end{aligned} \qquad (10)$$

where $B = (\Lambda + kI_p)$, and hence the RR coefficients can be formulated as:

$$\hat{\beta}_{RR} = \gamma\,\hat{\alpha}_{RR}(k) \qquad (11)$$

The bias, variance, and MSE of the RR estimator are given as follows:

$$\begin{aligned} Bias(\hat{\alpha}_{RR}(k)) &= kB^{-1}\hat{\alpha} \\ Var(\hat{\alpha}_{RR}(k)) &= \sigma^2 (I - kB^{-1})\Lambda^{-1}(I - kB^{-1})' \\ MSE(\hat{\alpha}_{RR}(k) &= Var(\hat{\alpha}_{RR}(k)) + [Bias(\hat{\alpha}_{RR}(k))][Bias(\hat{\alpha}_{RR}(k))]' \\ &= \sigma^2 (I - kB^{-1})\Lambda^{-1}(I - kB^{-1})' + k^2 B^{-1}\hat{\alpha}\hat{\alpha}'B^{-1} \end{aligned} \qquad (12)$$

where $Var(\hat{\alpha}) = \sigma^2\Lambda^{-1}$ and $\hat{\alpha} = \hat{\alpha}_{LS}$.

4. Jackknife Ridge Regression (JRR)

The Jackknife technique was originally proposed by Quenouille (1956) as a technique for reducing the bias of an estimator. Singh et al. (1986) suggested an approach to circumvent the biasing in RR depending on the Jackknife technique, formulated as:

$$y_{(-i)} = X_{(-i)}\beta + u^* \qquad (13)$$

where u^* is an error term with the ith coordinate deleted such that $u^* = 0$ and $Cov[u^*] = \sigma^2 I_{n-1}$ and $y_{(-i)}$ and $X_{(-i)}$ are the vector y with its ith value deleted and the matrix X with its ith row deleted respectively. The matrix $X_{(-i)}$ does not necessarily have full column rank. Then, the RR estimator in the reduced model is given as

$$\hat{\beta}_{RR,(-i)} = (X'_{(-i)} X_{(-i)} + kI_p)^{-1} X'_{(-i)} y_{(-i)} \tag{14}$$

The model in (14) is called ordinary Jackknife Ridge Regression (JRR). The JRR solution is given by [see, Hinkley (1977), Singh et al. (1986) and, Groβ (2003)]

$$\hat{\alpha}_{RR(-i)} = (Z'_{(-i)} Z_{(-i)} + kI)^{-1} Z'_{(-i)} y_{(-i)}$$

Let z_i and y_i being the ith column vector of Z and the ith coordinate of y, then:

$$\hat{\alpha}_{RR(i)} = (Z'Z - z'_i z_i + kI)^{-1} (Z'y - z_i y_i) \tag{15}$$

we can simplified (12) as following

$$\hat{\alpha}_{RR(-i)} = \hat{\alpha}_{RR} - \frac{B^{-1} z_i e_i}{1 - h_i}, \tag{16}$$

where $e_i = (y_i - z'_i \hat{\alpha})$ and $h_i = z'_i (Z'Z)^{-1} z_i$. Then, the pseudo values P_i defined as follows: [see, Groβ (2003)]

$$P_i = n\hat{\alpha}_{RR} - (n-1)\hat{\alpha}_{RR(-i)} \tag{17}$$

from (16) and (17), we obtain,

$$\overline{P} = \frac{1}{n}\sum_{i=1}^{n} P_i = \hat{\alpha}_{RR} + \frac{(n-1)}{n} B^{-1} \sum_{i=1}^{n} \frac{z_i e_i}{(1-h_i)} \tag{18}$$

since the variance of $(\hat{\alpha}_{RR} - \hat{\alpha}_{RR(-i)})$ is an increasing function of h_i, then we can define the pseudo values as

$$Q_i = \hat{\alpha}_{RR} + n(1-h_i)(\hat{\alpha}_{RR} - \hat{\alpha}_{RR(-i)}) \tag{18}$$

Hence, the corresponding JRR estimators, namely $\hat{\alpha}_{JRR}(k)$, are given by

$$\hat{\alpha}_{JRR}(k) = \overline{Q} = \frac{1}{n}\sum Q_i = \hat{\alpha}_{RR} + B^{-1} \sum z_i e_i \tag{19}$$

The simplified model for JRR can be written as

$$\begin{aligned} \hat{\alpha}_{JRR}(k) &= (I + kB^{-1})\,\hat{\alpha}_{RR}(k) \\ &= (I - k^2 B^{-2})\,\hat{\alpha} \end{aligned} \tag{20}$$

From (20) we can clearly see that the JRR estimators are obtained by shrinking the least squares estimator $\hat{\alpha}$ by the amount $k^2 B^{-2}$. The $\hat{\beta}_{JRR}$ is obtained by:

$$\hat{\beta}_{JRR} = \gamma\,\hat{\alpha}_{JRR}(k) \tag{21}$$

The bias, variance, and MSE for the JRR estimator are given as follows:

$$Bias(\hat{\alpha}_{JRR}(k)) = -k^2 B^{-2}\hat{\alpha}$$
$$Var(\hat{\alpha}_{JRR}(k)) = \sigma^2 (I - k^2 B^{-2})\Lambda^{-1}(I - k^2 B^{-2})' \quad (22)$$
$$MSE(\hat{\alpha}_{JRR}(k)) = Var(\hat{\alpha}_{JRR}(k)) + [Bias(\hat{\alpha}_{JRR}(k))][Bias(\hat{\alpha}_{JRR}(k))]'$$
$$= \sigma^2 (I - k^2 B^{-1})\Lambda^{-1}(I - k^2 B^{-1})' + k^4 B^{-2}\hat{\alpha}\hat{\alpha}' B^{-2}$$

The confidence interval for $\hat{\beta}_{JRR}$ is given by [see, Groβ (2003)

$$\hat{\beta}_{(JRR)i} \mp t(1 - \frac{SL}{2}, n - p).\sqrt{S} \ , \quad (23)$$

where *SL* is a significant level and *S* is given by

$$S = \frac{1}{n(n-p)}\sum_{i=1}^{n}(Q_i - \hat{\beta}_{JRR})(Q_i - \hat{\beta}_{JRR})' \quad (24)$$

5. Robust Regression Estimators

In regression analysis, the existence of outliers in the data set is a serious problem. This problem occurs when the distribution of the error term comes from a heavy tail distribution, and when the location or the scale is contaminated by outliers [see, Askin and Montgomery (1980)]. Outliers can have an affect in both *X* and *Y* directions and can have a significant adverse impact on the regression estimators. The classical methods are very sensitive to outliers because they rely on the least square estimation. For instance, even one outlier can destroy the least square estimation [see, Rousseeuw and Leroy (1987), Maronna (2006)]. The robust regression technique is an alternative method when the normal assumptions are unfulfilled by the given data. Several methods have been suggested to detect outliers in the data.

Huber (2003) proposed the most popular general technique of robust regression with a high breakdown point called the M-estimator. The M-estimator technique is based on replacing the sum squares of residuals in the least squares method by another robust function to cope with the problem of outliers. The objective function of the M-estimator is given by

$$min\sum_{i=1}^{n}\rho(\frac{r_i}{\hat{\sigma}}) = min\sum_{i=1}^{n}\rho(\frac{y_i - x_i'\hat{\beta}}{\hat{\sigma}}) \quad (25)$$

where $\hat{\sigma}$ is an estimate of the unknown scale parameter and ρ is a function that assigns the contribution of the individual residual in the objective function. Huber (2003) showed that the M-estimator is robust for outliers in *Y* direction but not robust to HLPs because it has an unbounded influence function. To circumvent the shortcomings of the M-estimator, Yohai developed a special class of M-estimator, namely the MM-estimator [see, Rousseeuw and Leroy (2003), Maronna (2006)].

The MM-estimator has good properties due to it combines high breakdown point (0.5) and excellent efficiency (0.95 of OLS efficiency under the normal assumptions). Another high breakdown robust method, the GM-estimator, was originally proposed by Mallows [Hekimoglu and Erenoglu (2013)]. The GM-estimator aims to down weight outliers in both the *X* and *Y* coordinates to ensure that HLPs get lower weights than small leverage points. Unfortunately, it is breakdown point does not exceed the amount [1/(p+1)] and also it is robust only with a small fraction of outliers [Hekimoglu and Erenoglu (2013)]. Multi-stage GM-estimators were then introduced in an attempt to remedy this shortcomings.

The procedure of multi-stage GM-estimators is to perform an initial estimation method that has good properties and also merge different techniques in different stages to satisfy a desirable property of the GM-estimator. GM1 and GM6 are two of the most practical modified GM-estimators. However, both GM1 and GM6 are not very successful in rectifying this problem. In this respect, Bagheri (2011) suggested the GM2 as a modification for the GM-estimator to be more efficient for outliers and high leverage points. The procedure for GM2 is summarized as follows.

Step 1: Calculate the initial residuals $r_i, i = 1,2,\ldots,n$ from the S-estimator and then find the scale of residuals $\hat{\tau}$, as follows:

$$r_i = y_i - \hat{y}_i\,,\;\; i = 1,2,\ldots,n$$

$$\hat{\tau} = 1.4826\,(1 + 5/(n-p))\,\text{Median}\left|r_i\right|$$

Step 2: Find the diagonal weights matrix *W*, with elements w_i given by:

$$w_{\mathbf{i}} = min[1, \{\frac{\chi^2_{(0.95,p+1)}}{\text{RMD}^2}\}],\quad i = 1,2,\ldots,n$$

where RMD is the robust mahalanobis distance based on the minimum volume ellipsoid (MVE) [Rousseeuw and Leroy (1987)]

Step 3: Calculate the influence ψ function for standardized residuals and then compute

$$A = \text{diag}\;\psi^*\,(\frac{r_i}{\hat{\tau} \times w_i})$$

where ψ^* is the derivative of Huber's influence function.

Step 4: Finally, the GM2-estimator can be obtained by deriving a one-step Newton Raphson as:

$$\hat{\beta}_{GM\,2} = \hat{\beta}_0 + (X'AX)^{-1}X'W\,\psi(\frac{r_{\mathbf{i}}}{w_i\hat{\tau}})\,\hat{\tau} \tag{26}$$

The RMD based on MVE can improve the ability for GM-estimator to detect the high leverage precisely.

6. Robust Ridge Regression and Robust Jackknife Ridge Regression

6.1 Robust Ridge Regression (RRR)

Assume that $\tilde{\beta}$ and $\tilde{\alpha}$ are robust coefficients obtained by using M-estimator, MM-estimator or any robust estimator, then the general form of robust ridge regression (RRR) version based on these robust methods can be written as follows:

$$\hat{\alpha}_{RRR}(k) = (I - k\,B^{-1})\,\tilde{\alpha} \tag{27}$$

The robust ridge regression coefficient is presented as

$$\hat{\beta}_{RRR} = \gamma\,\hat{\alpha}_{RRR}(k) \tag{28}$$

The robust shrinking parameter k corresponding with these robust coefficients can be calculated by using different approaches. For instance, by employing the Hoerl and Kennard technique,

$$\hat{k} = \frac{p\,\tilde{\sigma}^2}{\tilde{\beta}'\,\tilde{\beta}} \tag{29}$$

where

$$\tilde{\sigma}^2 = \frac{(y - X\tilde{\beta})'(y - X\tilde{\beta})}{\tilde{\beta}'\tilde{\beta}}$$

The bias, variance, and MSE of the RRR estimator are given by

$$\begin{aligned}
Bias\,(\hat{\alpha}_{RRR}(k)) &= E[\hat{\alpha}_{RRR}(k) - \alpha] = -k\,B^{-1}\alpha \\
Var\,(\hat{\alpha}_{RRR}(k)) &= E[(\hat{\alpha}_{RRR}(k) - \alpha)(\hat{\alpha}_{RRR}(k) - \alpha)'] \\
&= (I - k\,B^{-1})\Omega(I - k\,B^{-1})' \\
MSE\,(\hat{\alpha}_{RRR}(k)) &= Var\,(\hat{\alpha}_{RRR}(k)) + [Bias\,(\hat{\alpha}_{RRR}(k))][Bias\,(\hat{\alpha}_{RRR}(k))]' \\
&= (I - k\,B^{-1})\Omega\,(I - k\,B^{-1})' + k^2 B^{-1}\alpha\alpha' B^{-1}
\end{aligned} \tag{30}$$

where Ω is a finite covariance matrix of any robust component such as $\tilde{\alpha}$.

6.2 Robust Jackknife Ridge Regression (RJRR)

The robust Jackknife estimator based on the robust estimation method is given by [see, Batah et al. (2008), Jadhav and Kashid (2011) and Esra and Fikri (2012)],

$$\begin{aligned}
\hat{\alpha}_{RJRR.}(k) &= [I + k\,B]\,\tilde{\alpha}_{RRR} \\
&= [I + k\,B^{-1}][I - k\,B^{-1}]\,\tilde{\alpha} \\
&= (I - k^2 B^{-2})\,\tilde{\alpha}
\end{aligned}$$

The RJRR coefficient is given as

$$\hat{\beta}_{RJRR} = \gamma\,\hat{\alpha}_{RJRR}(k)$$

The bias is defined as follows:

$$\begin{aligned} Bias(\hat{\alpha}_{RJRR}(k)) &= E[\hat{\alpha}_{RJRR}(k)] - \alpha \\ &= E[(I - k^2 B^{-2})\tilde{\alpha}] - \alpha \\ &= (I - k^2 B^{-2}) E[\tilde{\alpha}] - \alpha \\ &= (I - k^2 B^{-2})\alpha - \alpha \\ &= -k^2 B^{-2}\alpha \end{aligned} \tag{31}$$

The variance of RJRR estimator is given by

$$\begin{aligned} Var(\alpha_{RJRR}(k)) &= E[(\hat{\alpha}_{RJRR}(k) - E(\hat{\alpha}_{RJRR}(k))][(\hat{\alpha}_{RJRR}(k) - E(\hat{\alpha}_{RJRR}(k))]' \\ &= E[\hat{\alpha}_{RJRR}(k) - \alpha][\hat{\alpha}_{RJRR}(k) - \alpha]' \\ &= (I - k^2 B^{-2})\Omega(I - k^2 B^{-2})' \end{aligned} \tag{32}$$

The MSE of RJRR estimator is given by

$$\begin{aligned} MSE(\hat{\alpha}_{RJRR}(k)) &= Var(\hat{\alpha}_{RJRR}(k) + [bias(\hat{\alpha}_{RJRR}(k)][bias(\hat{\alpha}_{RJRR}(k)]' \\ &= (I - k^2 B^{-2})\Omega(I - k^2 B^{-2})' + k^4 B^{-2}\alpha\alpha' B^{-2} \end{aligned} \tag{33}$$

7. Simulation Study

In this part, we discuss a simulation study to assess the performance of the proposed methods (RJMM and RJGM2) in the case of the simultaneous presence of the multicollinearity problem and HLPs in a data set. To generate simulated data with a different degree of multicollinearity, we apply a simulation approach given by Lawrence and Arthur (1990) and McDonald and Galarneau (1975). We consider the multivariate linear regression model as:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i \tag{34}$$

where ε is the error term distributed as $N(0,\sigma^2 I)$. The explanatory variables are generated by

$$x_{ij} = \rho v_{i4} + (1-\rho^2)^{1/2} v_{ij}, \quad i = 1,2,\cdots,n; \; j = 1,2, \text{and } 3.$$

where v_{i1}, v_{i2}, v_{i3}, and v_{i4} are independent standard normal pseudo random numbers, and p = 3 is the number of explanatory variables. The explanatory variables are standardized so that the design matrix $X'X$ is in the canonical form. The character ρ^2 denotes the degree of collinearity between x_{ij} and x_{im} for $j \neq m$. In addition, three different values of high collinearity are selected corresponding to $\rho = 0.90, 0.95$ and 0.95, and four different sets of observations are considered corresponding to n = 20, 30, 50 and 100.

The contamination is done by replacing a clean datum in the explanatory variables with HLPs corresponding to various ratios of the HLP, namely $\tau =$ 0.05, 0.10 and 0.15. Moreover, seven estimation methods are applied in this study, namely:

- Ordinary Least Squares (OLS)
- Ridge Regression (RR)
- Jackknife Ridge Regression (JRR)
- Robust Ridge Regression based on M-estimator (RRM)
- Robust Jackknife Ridge Regression based on M-estimator (RJM)
- Robust Jackknife Ridge Regression based on MM-estimator (RJMM)
- Robust Jackknife Ridge Regression based on GM2-estimator (RJGM2).

To compare the performance of the above estimation methods we used the following criteria:

- **Rote Mean Square Error (RMSE)**:
The RMSE is given as follows [see, Lawrence and Arthur (1990)]

$$\text{RMSE}(\beta,\hat{\beta}) = \sqrt{E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)']} \quad (35)$$

$$\text{RMSE}(\beta_j) = \sqrt{\frac{1}{R}\sum_{i=1}^{R}(\hat{\beta}_{ij}-\beta_j)^2},\ j=1,2,\ldots,p$$

here, R = 1500 is a replication of Monte Carlo simulation experiments, $\hat{\beta}_{ij}$ is the *i*th estimate of the *j*th parameter in the *i*th replication, and β_j, j = 1, 2, and 3, are the true coefficients of the regression model chosen as $\beta_1=\beta_2=\beta_3=1$

- **Square Loss Function (Loss):**
The Loss criteria is given as follows [see, Groβ (2003)]

$$\text{Loss}(\beta,\hat{\beta}) = (\hat{\beta}-\beta)'(\hat{\beta}-\beta) = \sum_{j=1}^{p}(\beta_j^*-\beta)^2 \quad (36)$$

where

$$\beta_j^* = E(\hat{\beta}) = \frac{1}{R}\sum_{i=1}^{R}\hat{\beta}_{ij},\ j=1,2 \text{ and } 3$$

- **Comparison of MSE**
Comparison of The MSE ratios of RJMM and RJGM2 over OLS, RR, JRR, RRM, and RJRM estimators are computed for all possible combinations of n,ρ and τ. If the ratio is less than one, the numerator is more efficient than the denominator, while if the ratio is greater than one, the denominator is more efficient than the numerator. If the ratio is exactly one, the numerator and denominator have the same efficiency.

8. Discussion

The simulation experiment is replicated 1500 times for all possible combinations of *n*, *p* and τ and the comparison criterias of MSE, RMSE, and Loss are computed for all methods of this study. The results for the simulation study are summarized in Tables 1 to 9. When the simulated data have multicollinearity and HLPs we can clearly observe that the values of RMSE and Loss for OLS, RR, and JRR are larger than the other robust estimator methods for all possible combinations of *n*, *p* and τ.

The values of RMSE and Loss for RRM and RJRM are smaller than those for the classical estimator (OLS, RR and RR) but they are less efficient than RJMM and RJGM2 because RRM and RJRM depend on the M-estimator, which is known to be less efficient with HLPs, while the MM-estimator and the GM2-estimator can do well with HLPs. RJMM and RJGM2 are the best methods in the presence of multicollinearity and HLPs. However, the performance of RJGM2 is better than that of RJMM in all possible cases except in the case of a small sample size, not very strong multicollinearity, and low and moderate HLP ratios (n = 20, ρ = .90 and τ = 0.05 and 0.10).

From Figs. 1 and 2 we can see that the curve of RMSE increases with increases in the degree of multicollinearity and also with increases in the ratio of HLPs, whereas RMSE decreases with increases in the number of observations. The estimators for RJGM2 are more reliable because they depend on RMD (MVE) for detecting multiple HLPs, so it can correctly identify the HLPs and then successfully down weight them. The comparison of ratios of MSE for the estimator methods are exhibited in Tables 4 to 9 shows that:

- The efficiency of the RJMM and RJGM2 estimators increases compared to the other estimators as the size of the sample increases (see Figs. 1and 2).
- The efficiency of the RJMM and RJGM2 estimators increases compared to the other estimators as the degree of multicollinearity increases (see Fig. 1).
- The efficiency of the RJMM and RJGM2 estimators increases compared to the other estimators as the ratio of HLPs increases (see Fig. 2).

In general, we can say that the RJGM2 is the best estimation method and is more efficient than the others to overcome the multicolinearity problem in the presence of HLPs.

Table 1: RMSE and Loss for estimation methods with τ = 0.05 (ratio of HLPs)

	n	20		30		50		100	
	ρ	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**
OLS		0.3029	0.0171	0.2103	0.0159	0.1991	0.0162	0.1850	0.0164
RR		0.2155	0.0168	0.1707	0.0160	0.1682	0.0163	0.1661	0.0165
JRR		0.2537	0.0168	0.1850	0.0159	0.1796	0.0162	0.1737	0.0164
RRM	**0.90**	0.1815	0.0172	0.1280	0.0076	0.1265	0.0070	0.1453	0.0115
RJRM		0.2070	0.0171	0.1327	0.0072	0.1281	0.0067	0.1447	0.0111
RJMM		0.0835	0.0001	0.0522	0.0004	0.0420	0.0001	0.0310	0.0001
RJGM2		0.0964	0.0001	0.0524	0.0003	0.0406	0.0001	0.0281	0.0001

OLS		0.3902	0.0174	0.2513	0.0159	0.2313	0.0161	0.2067	0.0163
RR		0.2537	0.0168	0.1833	0.0159	0.1773	0.0161	0.1724	0.0163
JRR		0.3141	0.0169	0.2089	0.0159	0.1978	0.0161	0.1864	0.0163
RRM	**0.95**	0.1996	0.0171	0.1183	0.0055	0.1255	0.0068	0.1413	0.0103
RJRM		0.2440	0.0170	0.1294	0.0053	0.1307	0.0065	0.1426	0.0100
RJMM		0.1019	0.0001	0.0593	0.0005	0.0476	0.0002	0.0356	0.0001
RJGM2		0.1313	0.0001	0.0557	0.0004	0.0426	0.0002	0.0269	0.0001
OLS		0.7946	0.0206	0.4657	0.0166	0.4068	0.0164	0.3334	0.0163
RR		0.4520	0.0173	0.2677	0.0161	0.2418	0.0161	0.2190	0.0162
JRR		0.6062	0.0183	0.3481	0.0165	0.3088	0.0162	0.2681	0.0162
RRM	**0.99**	0.2988	0.0173	0.1433	0.0052	0.1289	0.0054	0.1410	0.0093
RJRM		0.4323	0.0179	0.1916	0.0052	0.1563	0.0052	0.1530	0.0091
RJMM		0.1788	0.0001	0.0844	0.0006	0.0614	0.0003	0.0413	0.0001
RJGM2		0.2004	0.0001	0.0743	0.0005	0.0537	0.0002	0.0255	0.0001

Table 2: RMSE and Loss for estimation methods with τ = 0.10 (ratio of HLPs)

	n	**20**		**30**		**50**		**100**	
	ρ	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**
OLS		0.3990	0.0182	0.3011	0.0176	0.2532	0.0176	0.2107	0.0174
RR		0.2614	0.0179	0.2083	0.0175	0.1898	0.0176	0.1772	0.0174
JRR		0.3220	0.0180	0.2462	0.0176	0.2147	0.0176	0.1908	0.0174
RRM	**0.90**	0.2142	0.0177	0.1727	0.0175	0.1674	0.0178	0.1635	0.0175
RJRM		0.2662	0.0178	0.1896	0.0174	0.1745	0.0176	0.1650	0.0173
RJMM		0.1449	0.0039	0.1132	0.0031	0.0936	0.0028	0.0768	0.0026
RJGM2		0.1572	0.0037	0.0970	0.0030	0.0881	0.0025	0.0738	0.0026
OLS		0.5315	0.0188	0.3908	0.0177	0.3157	0.0176	0.2469	0.0174
RR		0.3237	0.0179	0.2450	0.0175	0.2117	0.0175	0.1891	0.0173
JRR		0.4160	0.0183	0.3065	0.0177	0.2538	0.0176	0.2129	0.0173
RRM	**0.95**	0.2386	0.0176	0.1790	0.0174	0.1695	0.0174	0.1635	0.0173
RJRM		0.3161	0.0177	0.2088	0.0173	0.1830	0.0173	0.1676	0.0172
RJMM		0.1691	0.0033	0.1257	0.0025	0.0995	0.0022	0.0779	0.0020
RJGM2		0.1782	0.0034	0.0917	0.0023	0.0821	0.0025	0.0666	0.0020
OLS		1.1238	0.0236	0.8092	0.0197	0.6209	0.0184	0.4402	0.0176
RR		0.6242	0.0191	0.4431	0.0184	0.3417	0.0177	0.2669	0.0173
JRR		0.8486	0.0210	0.6040	0.0194	0.4607	0.0181	0.3416	0.0174
RRM	**0.99**	0.4058	0.0178	0.2415	0.0174	0.1937	0.0172	0.1708	0.0171
RJRM		0.6022	0.0185	0.3475	0.0179	0.2486	0.0172	0.1919	0.0171
RJMM		0.2873	0.0028	0.1939	0.0022	0.1368	0.0019	0.0935	0.0017
RJGM2		0.2560	0.0026	0.1047	0.0021	0.0869	0.0017	0.0610	0.0017

Table 3: RMSE and Loss for estimation methods with τ = 0.15 (ratio of HLPs)

	n	**20**		**30**		**50**		**100**	
	ρ	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**	**RMSE**	**Loss**
OLS		0.4772	0.0186	0.3400	0.0181	0.2837	0.0181	0.2337	0.0179
RR		0.2999	0.0183	0.2227	0.0180	0.2038	0.0180	0.1872	0.0179
JRR		0.3802	0.0185	0.2713	0.0181	0.2365	0.0180	0.2065	0.0179
RRM	**0.90**	0.2395	0.0182	0.1758	0.0181	0.1714	0.0181	0.1663	0.0180
RJRM		0.3123	0.0184	0.1961	0.0180	0.1830	0.0180	0.1695	0.0179
RJMM		0.1968	0.0100	0.1427	0.0064	0.1282	0.0066	0.1189	0.0073
RJGM2		0.1935	0.0099	0.1234	0.0068	0.1200	0.0073	0.1143	0.0076

OLS		0.6431	0.0190	0.4489	0.0183	0.3625	0.0183	0.2827	0.0179
RR		0.3801	0.0184	0.2679	0.0181	0.2351	0.0180	0.2049	0.0178
JRR		0.4998	0.0188	0.3454	0.0183	0.2888	0.0181	0.2379	0.0178
RRM	**0.95**	0.2879	0.0182	0.1870	0.0179	0.1774	0.0180	0.1673	0.0179
RJRM		0.3990	0.0185	0.2258	0.0180	0.2000	0.0180	0.1739	0.0178
RJMM		0.2240	0.0087	0.1561	0.0055	0.1342	0.0056	0.1191	0.0062
RJGM2		0.2017	0.0088	0.1172	0.0057	0.1121	0.0062	0.1056	0.0065
OLS		1.3739	0.0230	0.9464	0.0212	0.7352	0.0200	0.5313	0.0181
RR		0.7539	0.0197	0.5040	0.0194	0.4089	0.0185	0.3118	0.0178
JRR		1.0372	0.0217	0.7006	0.0207	0.5521	0.0192	0.4093	0.0179
RRM	**0.99**	0.5626	0.0201	0.2601	0.0181	0.2195	0.0180	0.1783	0.0177
RJRM		0.8331	0.0229	0.3865	0.0187	0.3007	0.0181	0.2076	0.0177
RJMM		0.3726	0.0084	0.2365	0.0049	0.1807	0.0050	0.1402	0.0055
RJGM2		0.2696	0.0082	0.1286	0.0050	0.1123	0.0054	0.0985	0.0056

9. Conclusions

In this study, we proposed new estimation methods called RJMM and RJGM2, by integrating the JRR method with MM-estimator and GM2-estimator respectively, to remedy the combined problem of multicollinearity and HLPs. In order to examine the performance of the suggested methods, we compared them with existing methods by using a variety of simulation data based on RMSE, Loss and ratio of MSE.

The results indicate that the classical methods, RRM and RJM have bad performance compared with proposed methods when the correlated data has HLPs. So, we can say that the proposed methods are the best methods for solving multicollinearity in the presence of HLPs and for producing estimates with lower RMSE and less bias.

Table 4: Ratio of MSE of RJGM2 comparison with the other estimation methods of the study when τ = 0.05

n	ρ	$\frac{RJMGM}{OLS}$	$\frac{RJMGM}{RR}$	$\frac{RJMGM}{JRR}$	$\frac{RJMGM}{RRR}$	$\frac{RJMGM}{RJRM}$	$\frac{RJMGM}{RJMM}$
20	**0.90**	0.3182	0.4473	0.3799	0.5310	0.4656	1.1542
30		0.2492	0.3070	0.2833	0.4093	0.3949	1.0032
50		0.2041	0.2416	0.2262	0.3213	0.3172	0.9675
100		0.1518	0.1691	0.1616	0.1933	0.1941	0.9068
20	**0.95**	0.3366	0.5176	0.4181	0.6579	0.5382	1.2881
30		0.2218	0.3041	0.2669	0.4712	0.4306	0.9405
50		0.1843	0.2404	0.2155	0.3397	0.3261	0.8958
100		0.1302	0.1561	0.1443	0.1905	0.1887	0.7559
20	**0.99**	0.2523	0.4435	0.3306	0.6707	0.4637	1.1209
30		0.1595	0.2775	0.2134	0.5184	0.3877	0.8806
50		0.1319	0.2220	0.1738	0.4163	0.3434	0.8744
100		0.0765	0.1164	0.0951	0.1807	0.1666	0.6165

Table 5: Ratio of MSE of RJGM2 comparison with the other estimation methods of the study when τ = 0.10

n	ρ	$\frac{RJMGM}{OLS}$	$\frac{RJMGM}{RR}$	$\frac{RJMGM}{JRR}$	$\frac{RJMGM}{RRR}$	$\frac{RJMGM}{RJRM}$	$\frac{RJMGM}{RJMM}$
20	**0.90**	0.3939	0.6013	0.4881	0.7338	0.5905	1.0846
30		0.3223	0.4659	0.3941	0.5620	0.5120	0.8574
50		0.3478	0.4640	0.4102	0.5261	0.5048	0.9411
100		0.3501	0.4162	0.3866	0.4512	0.4469	0.9606
20	**0.95**	0.3354	0.5506	0.4285	0.7470	0.5640	1.0540
30		0.2347	0.3745	0.2993	0.5123	0.4394	0.7296
50		0.2600	0.3877	0.3235	0.4844	0.4485	0.8251
100		0.2699	0.3523	0.3130	0.4075	0.3975	0.8556
20	**0.99**	0.2278	0.4102	0.3017	0.6309	0.4252	0.8912
30		0.1294	0.2363	0.1734	0.4337	0.3014	0.5402
50		0.1399	0.2542	0.1885	0.4484	0.3494	0.6349
100		0.1386	0.2286	0.1786	0.3572	0.3179	0.6528

Table 6: Ratio of MSE of RJGM2 comparison with the other estimation methods of the study when τ = 0.15

n	ρ	$\frac{RJMGM}{OLS}$	$\frac{RJMGM}{RR}$	$\frac{RJMGM}{JRR}$	$\frac{RJMGM}{RRR}$	$\frac{RJMGM}{RJRM}$	$\frac{RJMGM}{RJMM}$
20	**0.90**	0.4054	0.6451	0.5090	0.8078	0.6196	0.9833
30		0.3628	0.5541	0.4547	0.7016	0.6291	0.8647
50		0.4229	0.5889	0.5074	0.6999	0.6559	0.9358
100		0.4893	0.6107	0.5536	0.6877	0.6746	0.9614
20	**0.95**	0.3137	0.5307	0.4036	0.7007	0.5055	0.9005
30		0.2611	0.4375	0.3393	0.6266	0.5192	0.7507
50		0.3092	0.4766	0.3881	0.6318	0.5604	0.8352
100		0.3734	0.5152	0.4438	0.6312	0.6069	0.8862
20	**0.99**	0.1962	0.3576	0.2599	0.4792	0.3236	0.7235
30		0.1359	0.2551	0.1836	0.4944	0.3327	0.5436
50		0.1527	0.2746	0.2033	0.5115	0.3733	0.6214
100		0.1854	0.3159	0.2407	0.5525	0.4746	0.7027

Table 7: Ratio of MSE of RJMM comparison with the other estimation methods of the study when τ = 0.05

n	ρ	$\frac{RJMM}{OLS}$	$\frac{RJMM}{RR}$	$\frac{RJMM}{JRR}$	$\frac{RJMM}{RRR}$	$\frac{RJMM}{RJRM}$	$\frac{RJMM}{RJMGM}$
20	**0.90**	0.2757	0.3875	0.3292	0.4601	0.4034	0.8664
30		0.2484	0.3060	0.2824	0.4080	0.3936	0.9968
50		0.2109	0.2497	0.2338	0.3320	0.3279	1.0336
100		0.1674	0.1864	0.1782	0.2132	0.2141	1.1027
20	**0.95**	0.2613	0.4018	0.3246	0.5107	0.4178	0.7763
30		0.2359	0.3233	0.2837	0.5010	0.4579	1.0633
50		0.2057	0.2683	0.2405	0.3792	0.3640	1.1163
100		0.1722	0.2065	0.1909	0.2520	0.2496	1.3229
20	**0.99**	0.2250	0.3956	0.2950	0.5984	0.4137	0.8921
30		0.1811	0.3151	0.2423	0.5887	0.4403	1.1356
50		0.1509	0.2538	0.1987	0.4761	0.3927	1.1436
100		0.1240	0.1888	0.1542	0.2932	0.2702	1.6222

Table 8: Ratio of MSE of RJMM comparison with the other estimation methods of the study when τ = 0.15

n	ρ	$\frac{RJMM}{OLS}$	$\frac{RJMM}{RR}$	$\frac{RJMM}{JRR}$	$\frac{RJMM}{RRR}$	$\frac{RJMM}{RJRM}$	$\frac{RJMM}{RJMGM}$
20	**0.90**	0.3631	0.5544	0.4500	0.6765	0.5444	0.9220
30		0.3759	0.5434	0.4597	0.6555	0.5971	1.1664
50		0.3696	0.4930	0.4359	0.5590	0.5364	1.0626
100		0.3645	0.4333	0.4025	0.4697	0.4652	1.0411
20	**0.95**	0.3182	0.5224	0.4065	0.7088	0.5351	0.9488
30		0.3218	0.5133	0.4102	0.7023	0.6023	1.3707
50		0.3151	0.4699	0.3920	0.5872	0.5436	1.2120
100		0.3154	0.4118	0.3658	0.4763	0.4646	1.1688
20	**0.99**	0.2556	0.4603	0.3386	0.7079	0.4771	1.1221
30		0.2396	0.4375	0.3210	0.8029	0.5579	1.8513
50		0.2204	0.4004	0.2969	0.7062	0.5503	1.5750
100		0.2123	0.3501	0.2736	0.5472	0.4870	1.5318

Table 9: Ratio of MSE of RJMM comparison with the other estimation methods of the study when τ = 0.15

n	ρ	$\frac{RJMM}{OLS}$	$\frac{RJMM}{RR}$	$\frac{RJMM}{JRR}$	$\frac{RJMM}{RRR}$	$\frac{RJMM}{RJRM}$	$\frac{RJMM}{RJMGM}$
20	**0.90**	0.4123	0.6561	0.5176	0.8216	0.6302	1.0170
30		0.4196	0.6408	0.5259	0.8114	0.7276	1.1565
50		0.4519	0.6293	0.5422	0.7479	0.7008	1.0686
100		0.5090	0.6352	0.5758	0.7153	0.7017	1.0402
20	**0.95**	0.3483	0.5893	0.4482	0.7781	0.5614	1.1105
30		0.3478	0.5828	0.4520	0.8348	0.6916	1.3321
50		0.3702	0.5706	0.4647	0.7564	0.6710	1.1973
100		0.4214	0.5814	0.5007	0.7122	0.6849	1.1284
20	**0.99**	0.2712	0.4943	0.3592	0.6623	0.4473	1.3822
30		0.2499	0.4693	0.3377	0.9095	0.6120	1.8394
50		0.2457	0.4418	0.3272	0.8231	0.6007	1.6092
100		0.2639	0.4496	0.3425	0.7863	0.6754	1.4231

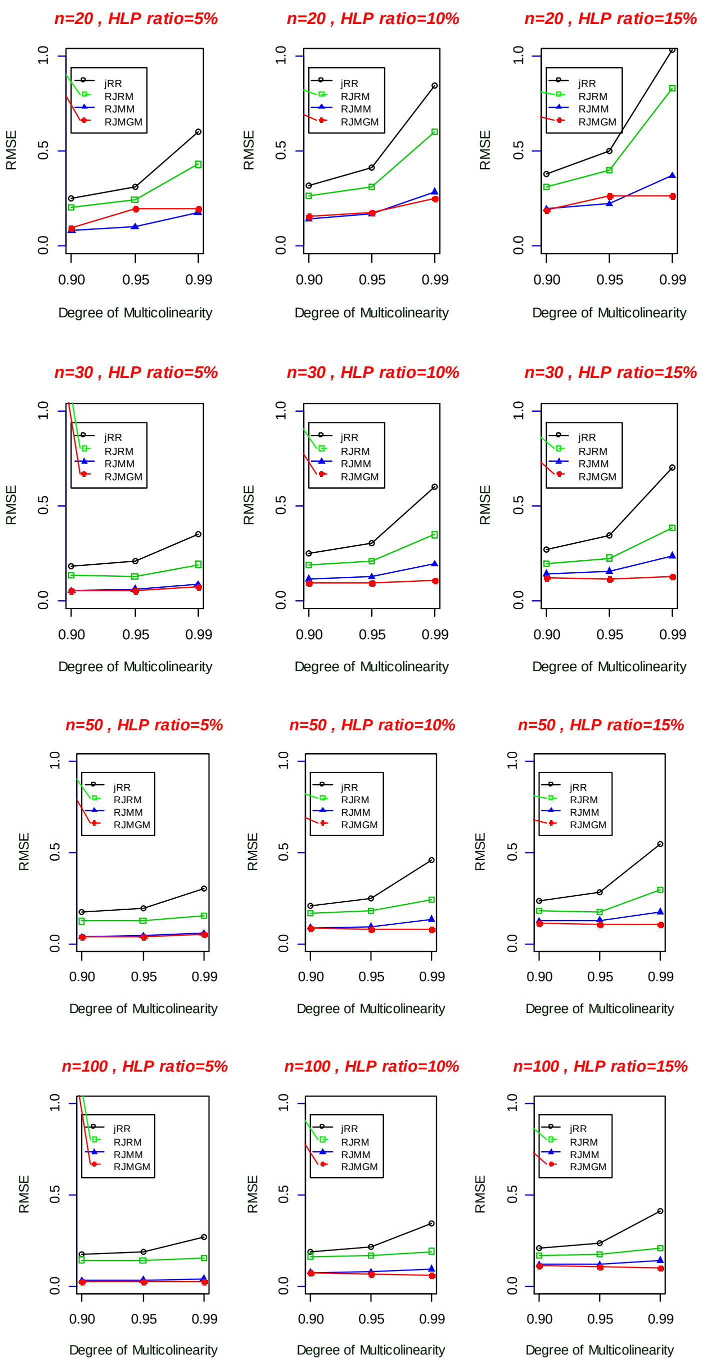

Figure 1. Degree of Multicollinearity against the RMSE for the robust estimation methods

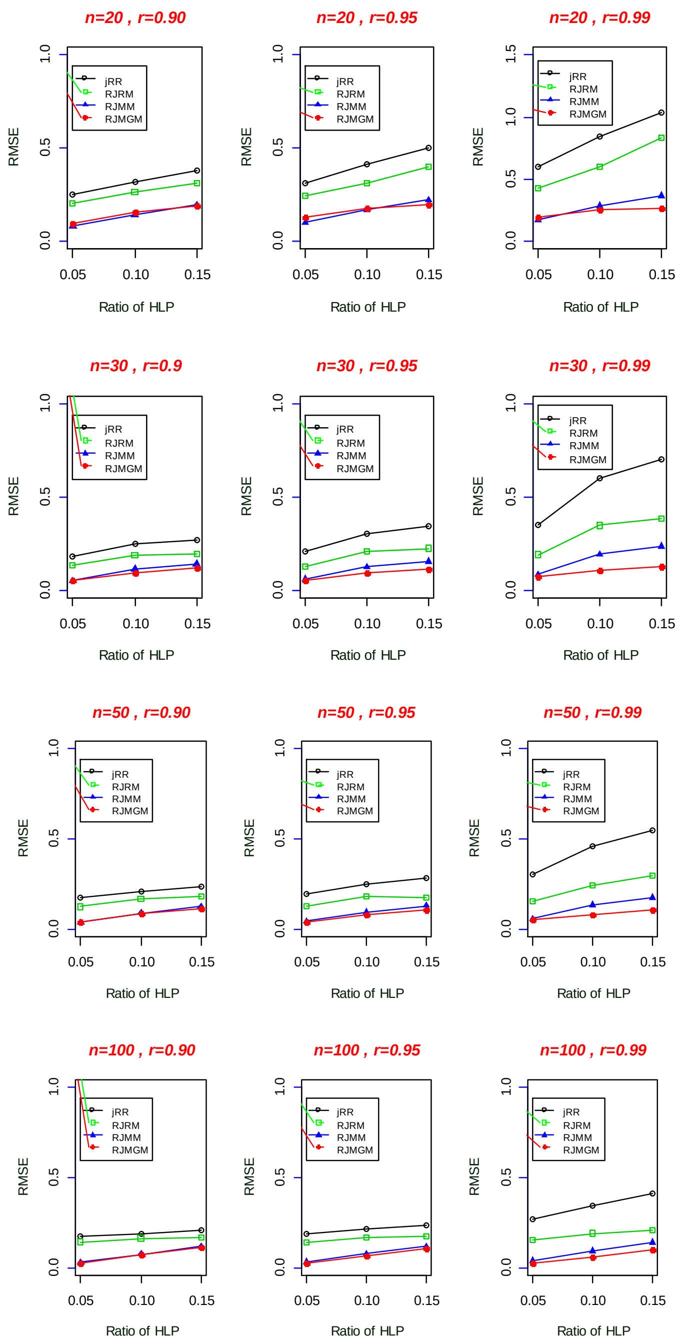

Figure 2. Ratio of HLPs against the RMSE for the robust estimation methods

REFERENCES

[1] **Askin R. G. and Montgomery D. C. (1980),** ***Augmented Robust Estimators***; Technometrics, 22, 333-341;
[2] **Baghri A. (2011),** ***Robust Estimation Methods and Robust Multicollinearity Diagnostics For Multiple Regression Model in the Presence of High Leverage Collinearity -Influential Observation***; Thesis submitted to the School of Graduate Studies, UPM;
[3] **Bastlevsky A. (1994),** ***Statistical Factor Analysis And Related Methods****; John Wiley*, INC.;
[4] **Batah F. S., Ramanathan T. V. and Gore S. D. (2008),** ***The Efficiency of Modified Jackknife and Ridge Type Regression Estimators: A Comparison***; Surv Math Appl, 3 (2008) 111-122;
[5] **Esra A. and Fikri A. (2012),** ***Efficiency of the Modified Jackknifed Liu-type Estimator****;* Stat Papers, 53, 265-28;.
[6] **Groβ J. (2003),** ***Linear Regression- Lecturer Notes in Statistics****; Springer Verlag Berlin Heidelberg;*
[7] **Habshah M. and Marina Z. (2007),** ***A Simulation Study on Ridge Regression Estimation in the Presence of Outliers and Multicollinearity***; Journal Teknologi, 47, 59-74;
[8] **Hekimoglu S. and Erenoglu R.C. (2013),** ***A New GM-estimate with High Breakdown Point****;* Acta Geod Geophys, 48, 419- 43;
[9] **Hinkley D. V.**1977) **,** ***Jackknifing in Unbalanced Situations***; Technometrics, 19, 285-292;
[10] **Hoerl E. and R. Kennard W. (1970),** ***Biased Estimation for Non-orthogonal Problems****;* Technometrics, 12, 69-82;
[11] **Huber, P.J (2003),** ***Robust Statistics*** *; Wiley*, New York, USA;
[12] **Jadhav N. H. and Kashid D. N. (2011),** ***A Jackknifed Ridge M-Estimator for Regression Model with Multicollinearity and Outliers***; Journal of Statistical Theory and Practice, 5, 659-673;
[13] **Lawrence K. D. and Arthur J. L. (1990),** ***Robust Regression: Analysis and Applications****; Marcel Dekker*, New York;
[14] **Maronna R. (2003),** ***Robust Statistics****; Wiley*, New York, USA;
[15] **McDonald G. C. and Galarneau D. I. (1975),** ***A Monte Carlo Evaluation of some Ridge-type Estimators****;* JASA, 20, 407-41;
[16] **Quenouille M. H. (1970),** ***Trust Notes on Bias in Estimation***; *Biometrika*, 53, 353-360;
[17] **Rousseeuw P. J. and Leroy A. M. (1987),** ***Robust Regression and Outlier Detection****; Wiley* New York;
[18] **Rousseeuw P.J. and Leory A.M. (2003),** ***Regression and Outlier Detection***; *Wiley*, New York, USA;
[19] **Singh B., Chaubey Y. P. and Dwivedi T. D. (1986),** ***An almost Unbiased Ridge Estimator****; The Indian Journal of Statistics*, 48, 342-346.