Chapter 2

Parameter Estimation in the Presence of Multicollinearity and Outliers

2.1. Introduction

This chapter is focused on the problem of simultaneous occurrence of outliers and multicollinearity. In Chapter 1, we have mentioned, the ordinary least squares estimation method performs well under some basic assumptions. In real life situation, multicollinearity, outliers and departure from the normality assumption are common problems in regression. These produce undesirable effects on the ordinary least squares estimator. This fact is pointed out by many researchers.

The problem of multicollinearity has received considerable attention from several researchers. Some of them have developed alternative parameter estimation methods when the multicollinearity is severe. Among these, Hoerl and Kennard (1970) proposed the ordinary ridge regression (ORR) estimator and generalized ridge regression (GRR) estimator. ORR estimator is widely used because of its optimality properties (see Vinod and Ullah, 1981). Various biased estimators like Liu Estimator (Liu, 1993), Liu-Type Estimators (LTE, Liu, 2003), $k-d$ class estimator (Sakallioglu and Kaciranlar, 2008) etc are available in the literature to deal with the problem of multicollinearity.

Another important problem which needs attention in regression parameter estimation is the presence of outliers in the data. It is well-known that (Montgomery, Peck and Vining, 2006), there are three types of outliers that may be present in the data. One type is Y-outlier, another one is X-outlier (outliers present in regressor variables) and the third is XY-outlier (outliers present in both response and regressor variables). Outliers affect the efficiency of ordinary least squares estimates (see Montgomery, Peck and Vining, 2006). Huber's landmark book on Robust Statistics (1998) focused on Y-outlier and

proposed M-estimator as a robust estimator of regression parameters. Thereafter, major work in this area was done by Rousseeuw and Leroy (1987). They proposed several robust estimators like least median squares (LMS), least trimmed squares (LTS), MM estimator and so on which are robust to Y as well as to X-outliers.

Several procedures are available to deal with multicollinearity and outliers separately. However, there is not much significant work reported in the literature which takes into account the presence of both multicollinearity and outlier problems simultaneously. To overcome the combined problem of multicollinearity and outliers, we propose a new estimator namely, Jackknifed Ridge M-estimator. This estimator is obtained by shrinking an M-estimator instead of the ordinary least squares estimator.

The outline of this chapter is as follows. Section 2.2 introduces various estimators of regression parameters. In Section 2.3, we define canonical form of model and parameter estimation methods with their Mean Square Error (MSE). In Section 2.4, we propose a new estimator designated as Jackknifed Ridge M-estimator (JRM) and discuss its properties. The performance of this estimator is studied in Section 2.5. In Section 2.6, a simulation study is carried for justifying the superiority of the suggested estimator. The chapter ends with some concluding remarks in Section 2.7.

2.2. Parameter Estimation Methods

In this section, we discuss the M-estimator and ORR estimator for the model given in Eq. (1.1.2). Also, estimation procedure of ridge parameter is presented.

2.2.1. Ordinary Ridge Regression (ORR) Estimator

The ORR estimator is defined as

$$\hat{\beta}_R(r) = \left(X'X + rI\right)^{-1} X'Y \qquad (2.2.1)$$

where r is the ridge parameter (known as shrinkage parameter or biasing constant).

2.2.2. Choice of Ridge Parameter

Hoerl, Kennard and Baldwin (1975) proposed the choice of the ridge parameter r to obtain the ORR estimator and is given by

$$r = \frac{(k-1)\hat{\sigma}^2}{\hat{\beta}'_{OLS}\hat{\beta}_{OLS}} \tag{2.2.2}$$

where $\hat{\sigma}^2 = Y'Y - \hat{\beta}'_{OLS}X'Y/(n-k)$.

2.2.3. M-estimator

The Huber M-estimator (Huber, 1973) for β given in (1.1.2) is obtained by minimizing a sum of the function of the scaled residuals

$$\sum_{i=1}^{n} \rho\left(\frac{\varepsilon_i}{s}\right) \tag{2.2.3}$$

where s is scale parameter which is replaced by its estimate. Differentiating (2.2.3) partially with respect to each parameter and equate to zero we get

$$\sum_{i=1}^{n} \psi\left(\frac{Y_i - X_i'\beta}{s}\right) x_{ij} = 0 \tag{2.2.4}$$

where $\psi(\cdot)$ is partial derivative of $\rho(\cdot)$ and x_{ij} is j^{th} entry in the i^{th} row of matrix X. There are various popular robust functions $\rho(\cdot)$ available in the literature. Some of them are listed in the Table 2.2.1 (Montgomery, Peck and Vining, 2006). Solution to Eq. (2.2.4) is obtained by iterative reweighted least squares method (Draper and Smith, 2003). At convergence, M-estimator may be written as

$$\hat{\beta}_M = (X'WX)^{-1}X'W\,Y \tag{2.2.5}$$

where W is diagonal weight matrix.

Table 2.2.1
Robust Criterion Functions

Criterion	$\rho(z)$	$\psi(z)$	$W(z)$	Range
Least squares	$\frac{1}{2}z^2$	z	1.0	$\lvert z\rvert < \infty$
Huber's t function t =1.345	$\frac{1}{2}z^2$	z	1.0	$\lvert z\rvert \le t$
	$\lvert z\rvert\, t - \frac{1}{2}t^2$	$t\,sign(z)$	$t/\lvert z\rvert$	$\lvert z\rvert > t$

Hampel's 17A function $a = 1.7$ $b = 3.4$ $c = 8.5$	$\frac{a}{\lvert z\rvert}$ $a\lvert z\rvert - \frac{1}{2}a^2$ $\frac{a(c\lvert z\rvert - \frac{1}{2}z^2)}{c-b} - (7/6)a^2$ $a(b+c-a)$	z $a\,sign(z)$ $\frac{a\,sign(z)(c-\lvert z\rvert)}{c-b}$ 0	1.0 $a/\lvert z\rvert$ $\frac{a\,(c\,-\,\lvert z\rvert)}{\lvert z\rvert(c\,-\,b)}$ 0	$\lvert z\rvert \le a$ $a < \lvert z\rvert \le b$ $b < \lvert z\rvert \le c$ $\lvert z\rvert > c$

Below we discuss the above estimators for regression parameters in the canonical form.

2.3. Canonical Form of Regression Model and Estimators

Consider the multiple linear regression model given in Eq. (1.1.2). We rewrite this model as

$$Y = \beta_0 1 + X^*\beta^* + \varepsilon \tag{2.3.1}$$

where Y is an $n \times 1$ response vector, 1 is column of ones, X^* is an $n \times (k-1)$ matrix of regressor variables, β_0 is intercept parameter, β^* is $(k-1) \times 1$ vector of unknown regression coefficients, ε is an $n \times 1$ error vector with mean 0 and variance $\sigma^2 I$. For convenience, it is assumed that the X^* variables are standardized so that $X^{*\prime}X^*$ has the form of correlation matrix.

Let $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ be the eigen values of $X^{*\prime}X^*$ and $q_1, q_2, \ldots, q_{k-1}$ be the corresponding eigen vectors. Let $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_{k-1})$ and $Q = (q_1, q_2, \ldots, q_{k-1})$ such that $X^{*\prime}X^* = Q\Lambda Q'$. The regression model (2.3.1) can be written in the canonical form as

$$Y = \beta_0 1 + Z\alpha + \varepsilon \tag{2.3.2}$$

where $Z = X^*Q$ and $\alpha = Q'\beta^*$.

The ordinary least squares (OLS) estimator of (2.3.2) is given by

$$\hat{\alpha}_{OLS} = \Lambda^{-1}Z'Y \tag{2.3.3}$$

Since $\alpha = Q'\beta^*$ and $Q'Q = I$, the OLS estimator of β^* is given by

$$\hat{\beta}^*_{OLS} = Q\hat{\alpha}_{OLS} \tag{2.3.4}$$

Note that, because of the relation $\alpha = Q'\beta^*$, any estimator $\hat{\alpha}$ of α has a corresponding $\hat{\beta}^* = Q\hat{\alpha}$ and MSE $(\hat{\alpha})$ = MSE $(\hat{\beta}^*)$. Hence, it is sufficient to consider only the canonical form. Various estimators of regression parameters in the canonical form of model are obtained and reported in the following subsections.

2.3.1. Ordinary Ridge Regression (ORR) Estimator in the Presence of Multicollinearity

In the canonical form, the Ordinary Ridge Regression (ORR) estimator proposed by Hoerl and Kennard (1970) is defined as

$$\begin{aligned}\hat{\alpha}_R(r) &= (\Lambda + rI)^{-1}Z'Y \\ &= (I - rA^{-1})\hat{\alpha}_{OLS}\end{aligned} \tag{2.3.5}$$

where $r > 0$ is the shrinkage or ridge parameter and $A = (\Lambda + rI)$.

The Mean Square Error (MSE) matrix of ORR estimator is given by

$$\begin{aligned}MSE\ (\hat{\alpha}_R(r)) &= \text{Var}\ (\hat{\alpha}_R(r)) + [\text{bias}\ (\hat{\alpha}_R(r))][\text{bias}\ (\hat{\alpha}_R(r))]' \\ &= \sigma^2(I - rA^{-1})\Lambda^{-1}(I - rA^{-1}) + r^2A^{-1}\alpha\alpha'A^{-1},\end{aligned} \tag{2.3.6}$$

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

2.3.2. Ordinary Jackknifed Ridge Regression (OJR) Estimator

Singh, Chaubey and Dwivedi (1986) extended the Jackknife procedure to ridge regression and proposed Ordinary Jackknifed Ridge Regression (OJR) estimator as

$$\begin{aligned}\hat{\alpha}_J(r) &= (I - rA^{-1})\hat{\alpha}_R(r) \\ &= (I - r^2A^{-2})\hat{\alpha}_{OLS}\end{aligned} \tag{2.3.7}$$

The OJR estimator is obtained by shrinking the OLS estimator $\hat{\alpha}_{OLS}$ using the matrix$(I - r^2A^{-2})$. The matrix of MSE value of this estimator is

$$\begin{aligned}MSE\ (\hat{\alpha}_J(r)) &= \text{Var}\ (\hat{\alpha}_J(r)) + [\text{bias}\ (\hat{\alpha}_J(r))][\text{bias}\ (\hat{\alpha}_J(r))]' \\ &= \sigma^2(I - r^2A^{-2})\Lambda^{-1}(I - r^2A^{-2})' + r^4A^{-2}\alpha\alpha'A^{-2},\end{aligned} \tag{2.3.8}$$

2.3.3. M-Estimator when Data Contains Outliers

M-estimator for parameters of regression model given in Eq. (2.3.2) is obtained by solving simultaneously k equations of the following form

$$\sum_{i=1}^{n} v_{ij}\, \psi\left(\frac{Y_i-\beta_0-Z_i'\alpha}{s}\right) = 0, \qquad j = 0, 1, \ldots, k-1, \tag{2.3.9}$$

where v_{ij} is the jth entry in the ith row of matrix V= [1, Z]. The solution is obtained by iterative reweighted least squares method (Huber, 1973, Draper and Smith, 2003). Huber (1973) has shown that asymptotically $\hat{\alpha}_M$ has an approximate normal distribution with covariance matrix

$$cov(\hat{\alpha}_M) = \Omega = \sigma^2 \frac{E\psi^2(\varepsilon/s)}{[E\psi'(\varepsilon/s)]^2}(V'V)^{-1} \tag{2.3.10}$$

Notice that OJR estimator is robust to multicollinearity and M-estimator is robust to outliers. However, if the multicollinearity and outliers occur simultaneously in the data, individual performance of these estimators is unsatisfactory. To deal with the combined problem of multicollinearity and outliers, we propose therefore a new estimator called the Jackknifed Ridge M-estimator (JRM).

2.3.4. Choice of Ridge Parameter r

The theoretical choice of r suggested by Hoerl, Kennard and Baldwin (1975) is

$$r = \frac{(k-1)\sigma^2}{\alpha'\alpha}, \tag{2.3.11}$$

which minimizes the MSE of$\hat{\alpha}_R(r)$. Usually, the value of r is obtained by substituting σ^2 and α by their OLS estimates. To develop the proposed method, we estimate the value of r by, say $\hat{r}$ by replacing σ^2 and α in Eq. (2.3.11) with their robust estimates. More specifically, the estimate of r is obtained by

$$\hat{r} = \frac{(k-1)s^2}{\hat{\alpha}_M'\hat{\alpha}_M} \tag{2.3.12}$$

Where $s = 1.4826\, median|e_i - median(e_i)|$ and e_i is ith residual obtained by using OLS estimator. We use s as a robust estimate of σ.

BARR. B. [illegible] & BEK LIBRARY
SHIVAJI UNIVERSITY, KOLHAPUR.

2.4. Proposed Estimator - Jackknifed Ridge M-estimator

Let $(\hat{\beta}_{0M}, \hat{\alpha}_M)$ be an M-estimator of (β_0, α). It is obtained by solving the equations $\sum_{i=1}^{n} v_{ij}\, \psi\left(\frac{Y_i - \beta_0 - Z_i'\alpha}{s}\right) = 0$, where $\psi(\cdot)$ is robust criterion function. Thus, Jackknifed Ridge M-estimator (JRM) is defined as,

$$\hat{\alpha}_{JRM}(r) = (I - r^2 A^{-2})\hat{\alpha}_M \tag{2.4.1}$$

Now we proceed to obtain the expressions for bias, variance and MSE of JRM estimator.

2.4.1. Bias

The bias in $\hat{\alpha}_{JRM}(r)$ is given by

$$\begin{aligned}
\text{Bias}\,(\hat{\alpha}_{JRM}(r)) &= E\left[\hat{\alpha}_{JRM}(r) - \alpha\right] \\
&= E[(I - r^2 A^{-2})\hat{\alpha}_M - \alpha] \\
&= (I - r^2 A^{-2})E[\hat{\alpha}_M] - \alpha \\
&= (I - r^2 A^{-2})\alpha - \alpha \\
&= -r^2 A^{-2}\alpha
\end{aligned} \tag{2.4.2}$$

2.4.2. Variance

The variance of $\hat{\alpha}_{JRM}(r)$ is as follows:

$$\begin{aligned}
\text{Var}\,(\hat{\alpha}_{JRM}(r)) &= \text{Var}\,((I - r^2 A^{-2})\hat{\alpha}_M) \\
&= (I - r^2 A^{-2})\ \text{Var}(\hat{\alpha}_M)\ (I - r^2 A^{-2})' \\
&= (I - r^2 A^{-2})\ \Omega\ (I - r^2 A^{-2})'
\end{aligned} \tag{2.4.3}$$

where Ω is covariance matrix of $\hat{\alpha}_M$, and we assume that it is finite.

2.4.3. Mean Square Error (MSE)

$$\begin{aligned}
\text{MSE}\,(\hat{\alpha}_{JRM}(r)) &= \text{Var}\,((\hat{\alpha}_{JRM}(r)) + \\
&\qquad [\text{Bias}\,(\hat{\alpha}_{JRM}(r))]\ [\text{Bias}\,(\hat{\alpha}_{JRM}(r))]' \\
&= (I - r^2 A^{-2})\ \Omega\ (I - r^2 A^{-2})' + r^4 A^{-2}\alpha\alpha' A^{-2}
\end{aligned} \tag{2.4.4}$$

2.5. Performance of JRM Estimator

In this section we compare the performance of proposed estimator with ORR, OJR and M-estimator in the MSE sense. We need a scalar function of the MSE matrix to compare meaningfully the performance of the above estimators. To do this, note that the diagonal elements of MSE matrix represent the mean square error of individual estimates. Further, the Trace of MSE measures the 'total variation' in the estimator. We use the Trace of MSE as a measure for evaluation of performance of various estimators. What it means is that suppose $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are two estimators of α with $\mathrm{m}(\hat{\alpha}_1)$ = trace(MSE($\hat{\alpha}_1$)) and $\mathrm{m}(\hat{\alpha}_2)$ = trace(MSE($\hat{\alpha}_2$)) respectively where MSE($\hat{\alpha}_1$) and MSE($\hat{\alpha}_2$) are the matrices of MSE. We prefer $\hat{\alpha}_2$ to $\hat{\alpha}_1$, if $\mathrm{m}(\hat{\alpha}_1)$–$\mathrm{m}(\hat{\alpha}_2) \geq 0$ (see Sakallioglu and Kaciranlar, 2008). Using the Trace criterion, we compare the JRM estimator with ORR estimator, OJR estimator, and M-estimator.

2.5.1. Comparison between the JRM and ORR Estimator

Theorem 2.5.1. If

$$\frac{\sigma^2}{\lambda_i} > \left(1+\frac{r}{\lambda_i+r}\right)^2 \Omega_{ii} \;\forall\, \lambda_i, i = 1,2,\dots,k \text{ and } r > 0 \qquad (2.5.1)$$

then $\mathrm{m}(\hat{\alpha}_R(r)) > \mathrm{m}(\hat{\alpha}_{JRM}(r))$.

Proof:

We write the difference between the MSE matrices as

$$\text{MSE}\,(\hat{\alpha}_R(r)) - \text{MSE}\,(\hat{\alpha}_{JRM}(r)) = \sigma^2(I - rA^{-1})\Lambda^{-1}(I - rA^{-1}) + r^2A^{-1}\alpha\alpha'A^{-1} - (I - r^2A^{-2})\,\Omega\,(I - r^2A^{-2})' - r^4A^{-2}\alpha\alpha'A^{-2} \qquad (2.5.2)$$

Equivalently, we can write

$$\mathrm{m}(\hat{\alpha}_R(r)) - \mathrm{m}(\hat{\alpha}_{JRM}(r)) = \sum_{i=1}^{k} \frac{\sigma^2}{\lambda_i}\left(1-\frac{r}{\lambda_i+r}\right)^2 + \left(\frac{r}{\lambda_i+r}\right)^2 \alpha_i^2 \left[\left(1+\frac{r}{\lambda_i+r}\right)^2 \Omega_{ii}\right]$$

$$= \sum_{i=1}^{k}\left(1-\frac{r}{\lambda_i+r}\right)^2\left[\frac{\sigma^2}{\lambda_i} - \left(1+\frac{r}{\lambda_i+r}\right)^2\Omega_{ii}\right] + \sum_{i=1}^{k}\left(\frac{r}{\lambda_i+r}\right)^2\left(1-\frac{r^2}{(\lambda_i+r)^2}\right)\alpha_i^2, \qquad (2.5.3)$$

where Ω_{ii} is the ith diagonal element of Ω.

Hence

$$m(\hat{\alpha}_R(r)) - m(\hat{\alpha}_{JRM}(r)) > 0,$$

whenever,

$$\frac{\sigma^2}{\lambda_i} > \left(1 + \frac{r}{\lambda_i + r}\right)^2 \Omega_{ii} \; \forall \; i = 1, 2, \ldots, k$$

Hence the proof.

2.5.2. Comparison between the JRM and OJR Estimator

Theorem 2.5.2. If

$$\frac{\sigma^2}{\lambda_i} \geq \Omega_{ii} \qquad \forall \; \lambda_i, \; i = 1, 2, \ldots, k, \tag{2.5.4}$$

then $m(\hat{\alpha}_J(r)) \geq m(\hat{\alpha}_{JRM}(r))$.

Proof:

Using equations (2.3.8) and (2.4.4), we write the difference of MSE matrices as

$$\begin{aligned} MSE(\hat{\alpha}_J(r)) - MSE(\hat{\alpha}_{JRM}(r)) &= \sigma^2(I - r^2A^{-2})\Lambda^{-1}(I - r^2A^{-2})' + r^4A^{-2}\alpha\alpha'A^{-2} \\ &\quad -(I - r^2A^{-2})\,\Omega\,(I - r^2A^{-2})' - r^4A^{-2}\alpha\alpha'A^{-2} \end{aligned}$$

$$\begin{aligned} &= \sigma^2(I - r^2A^{-2})\Lambda^{-1}(I - r^2A^{-2})' \\ &\quad -(I - r^2A^{-2})\,\Omega\,(I - r^2A^{-2})' \end{aligned} \tag{2.5.5}$$

Equivalently,

$$m(\hat{\alpha}_J(r)) - m(\hat{\alpha}_{JRM}(r)) = \sum_{i=1}^{k} \left(1 - \frac{r^2}{(\lambda_i + r)^2}\right)^2 \left[\frac{\sigma^2}{\lambda_i} - \Omega_{ii}\right] \tag{2.5.6}$$

The MSE of JRM is less than OJR if $\frac{\sigma^2}{\lambda_i} > \Omega_{ii} \; \forall \; i = 1, 2, \ldots, k$.

2.5.3. Comparison between the JRM and M-Estimator

Theorem 2.5.3. If

$$r < \frac{\lambda_i}{\sqrt{\frac{\Omega_{ii} + \alpha_i^2}{\Omega_{ii}} - 1}} \qquad \forall \; i = 1, 2, \ldots, k, \tag{2.5.7}$$

then $m(\hat{\alpha}_M) \geq m(\hat{\alpha}_{JRM}(r))$.

Proof:

The MSE expressions of $\hat{\alpha}_M$ and $\hat{\alpha}_{JRM}(r)$ are

$$\text{MSE}\,(\hat{\alpha}_M) = \sum_{i=1}^{k} \Omega_{ii} \tag{2.5.8}$$

$$f(r) = \text{MSE}(\hat{\alpha}_{JRM}(r)) = \sum_{i=1}^{k} \left(\left(1 - \frac{r^2}{(\lambda_i + r)^2}\right)^2 \Omega_{ii} - \frac{r^4}{(\lambda_i + r)^4} \right) \alpha_i^2, r > 0 \tag{2.5.9}$$

Since $f(0) = \text{MSE}\,(\hat{\alpha}_M)$, we find the condition on r such that $f(r)$ is decreasing function in r. If we take derivative of $f(r)$ with respect to r, we obtain

$$\frac{\partial f}{\partial r} = f'(r) = \sum_{i=1}^{k} \left[\frac{-4\, r \lambda_i}{(\lambda_i + r)^3} \Omega_{ii} + \frac{4\, r^3 \lambda_i}{(\lambda_i + r)^5} (\Omega_{ii} + \alpha_i^2) \right] \tag{2.5.10}$$

Hence, $f'^{(r)} < 0$ if

$$\frac{-4\, r \lambda_i}{(\lambda_i + r)^3} \Omega_{ii} + \frac{4\, r^3 \lambda_i}{(\lambda_i + r)^5} (\Omega_{ii} + \alpha_i^2) < 0 \qquad \forall\, i = 1, 2, \ldots, k.$$

The function $f(r)$ will be a decreasing function of r if

$$r < \frac{\lambda_i}{\sqrt{\frac{\Omega_{ii} + \alpha_i^2}{\Omega_{ii}} - 1}} \qquad \forall\, i = 1, 2, \ldots, k. \tag{2.5.11}$$

Hence the proof is completed.

It can be seen from the above results that in case of more severe multicollinearity and more influential outliers, the values of last few λ_i's will be small and consequently the $\text{m}(\hat{\alpha}_{JRM}(r))$. It indicates that in the presence of multicollinearity and outlier, performance of JRM estimator is better than ORR, OJR and M-estimator.

Remark 2.5.1 When there is no multicollinearity and no outliers in the data, the JRM, M-estimator, OJR and ORR estimators reduce to the ordinary least squares estimator. Theoretically, the MSE of each estimator is equal.

2.6. A Simulation Study

In this section, we present a simulation study to evaluate the performance of JRM estimator. To achieve required degree of multicollinearity, we used a simulation design given by McDonald and Galarneau (1975). The regressor variables are generated by

$$x_{ij} = (1-\rho^2)^{1/2} Z_{ij} + \rho Z_{i5}, i = 1, 2, \ldots, n, \; j = 1, 2, \ldots, p \quad (2.6.1)$$

where $p = 4$, Z_{ij} are independent standard normal pseudo random numbers, ρ^2 is the correlation between any two regressor variables and error term is generated from $N(0, \sigma^2)$. Response variable Y is generated using the regression model

$$Y = 2 + 5X_1 + 3X_2 + 4X_3 + 8X_4 + \varepsilon.$$

Note that the choice of regression model given above is arbitrary and for the sake of illustration, it is used here. Experiments with some other choices of regression coefficients have produced similar results. In this study, ρ = 0.9, 0.99, 0.999, sample size (n) = 20, 30, 50, 100 and error variance (σ^2) = 1, 9, 25, 100, 400, 900 are used.

Let $e_i^*, i = 1, 2, \ldots, n$ be the absolute value of i^{th} residual obtained from the least squares fit. Next, arrange these values in increasing order of their magnitude such that $e_{(1)}^* \leq e_{(2)}^* \leq \cdots \leq e_{(n)}^*$. We introduce one outlier by multiplying actual value of Y corresponding to $e_{(n)}^*$ by twenty. Similarly, two outliers are introduced in response variable Y corresponding to $e_{(n)}^*$, $e_{(n-1)}^*$ and so on.

2.6.1. Frequency of Smaller MSE

The above simulation experiment is repeated 10000 times and the MSE of JRM, M-estimator, OJR, ORR and OLS are computed. The frequency with which each estimator has lowest MSE is reported in Table 2.6.1. This experiment is replicated for all possible combinations of ρ, n and σ^2 for two cases

Case I: Multicollinearity without outlier
Case II: Multicollinearity with outlier

Table 2.6.1
Smaller MSE frequencies of JRM, M-estimator, OJR, ORR and OLS estimators in 10000 replications.

ρ	σ^2	JRM	M-est*	OJR	ORR	OLS	JRM	M-est	OJR	ORR	OLS
		Case I					Case II				
		n = 20									
0.9	1	1875	1755	202	3978	2190	6146	3788	2	33	31
	9	1568	986	215	6052	1179	7845	1926	6	175	48
	25	1340	543	218	7341	558	8730	822	4	416	28
	100	904	132	194	8611	159	8869	205	18	898	10
	400	571	34	230	9125	40	9049	62	9	878	2
	900	471	18	184	9308	19	9282	22	8	686	2
0.99	1	1630	977	174	6100	1119	8194	1724	2	65	5
	9	858	154	167	8662	159	9436	223	3	333	5
	25	604	58	133	9157	48	9371	57	6	565	1
	100	440	12	116	9419	13	8947	15	5	1033	0
	400	302	2	63	9633	0	8930	5	4	1061	0
	900	273	0	58	9666	3	9237	1	0	762	0
0.999	1	872	135	146	8685	162	9711	184	1	102	2
	9	435	18	110	9417	20	9601	14	2	382	1
	25	314	7	94	9577	8	9348	4	1	647	0
	100	264	1	49	9686	0	8930	0	0	1070	0
	400	243	0	29	9728	0	8956	0	0	1044	0
	900	213	0	16	9771	0	9255	0	0	745	0
		n = 30									
0.9	1	1984	1815	191	3715	2295	5914	4026	3	38	19
	9	1648	1208	230	5471	1443	7319	2416	10	212	43
	25	1422	755	215	6785	823	8217	1302	5	437	39
	100	934	202	246	8382	236	8763	326	16	883	12
	400	580	49	253	9066	52	9097	79	11	807	6
	900	464	32	224	9250	30	9435	46	5	514	0
0.99	1	1713	1139	216	5476	1456	7554	2365	1	67	13
	9	981	262	162	8304	291	9292	350	5	345	8
	25	621	96	177	8993	113	9225	98	6	666	5
	100	364	26	130	9451	29	8841	34	8	1115	2
	400	300	7	104	9585	4	8997	6	5	991	1
	900	256	4	71	9669	0	9377	2	0	621	0
0.999	1	921	220	202	8405	252	9570	291	2	135	2
	9	402	28	155	9387	28	9522	22	7	448	1
	25	297	12	113	9567	11	9225	12	6	756	1
	100	209	6	71	9710	4	8790	3	1	1206	0
	400	163	0	38	9798	1	8987	0	2	1011	0
	900	152	0	24	9823	1	9308	0	1	691	0
		n = 50									
0.9	1	2067	1928	165	3514	2326	5583	4343	3	40	31
	9	1884	1398	191	4732	1795	6792	2904	13	208	83
	25	1621	943	230	6009	1197	7613	1840	16	435	96
	100	1093	324	245	7939	399	8510	574	23	857	36
	400	546		280	8969	108	9116	146	18	716	4
	900	520	47	283	9098	52	9437	77	2	483	1
0.99	1	1865	1375	223	4823	1714	6876	3016	1	78	29
	9	1117	402	211	7784	486	8905	681	6	387	21
	25	676	157	180	8820	167	9077	178	10	722	13
	100	359	52	191	9357	41	8618	47	18	1316	1
	400	277	6	147	9558	12	8996	15	5	983	1
	900	223	3	117	9654	3	9380	7	2	611	0
0.999	1	1123	347	197	7919	414	9294	562	1	137	6
	9	423	47	188	9295	47	9351	60	5	583	1
	25	310	20	163	9484	23	9047	23	9	920	1
	100	216	8	108	9660	8	8616	3	8	1372	1
	400	153	0	56	9790	1	8887	3	3	1107	0
	900	138	0	44	9818	0	9360	1	0	639	0

		n = 100									
0.9	1	2092	1922	135	3289	2562	5383	4514	2	61	40
	9	1972	1649	187	4112	2080	6059	3585	9	223	124
	25	1825	1306	193	5042	1634	6808	2604	30	419	139
	100	1336	593	234	7060	777	7889	1123	37	866	85
	400	768	183	259	8543	247	9032	254	12	692	10
	900	619	86	299	8899	97	9352	135	4	504	5
0.99	1	1972	1610	171	4167	2080	6265	3599	2	85	49
	9	1458	696	222	6790	834	8197	1296	18	442	47
	25	979	276	230	8176	339	8646	468	16	834	36
	100	521	86	197	9101	95	8375	128	21	1464	12
	400	327	22	179	9438	34	8875	32	10	1082	1
	900	290	10	152	9540	8	9255	10	1	734	0
0.999	1	1411	661	203	6894	831	8686	1163	5	132	14
	9	546	84	201	9064	105	9157	121	15	697	10
	25	367	28	176	9389	40	8739	52	11	1197	1
	100	244	14	152	9582	8	8296	8	14	1682	0
	400	201	2	95	9700	2	8894	2	1	1103	0
	900	165	1	69	9764	1	9287	1	1	711	0

*M-est denotes M-Estimator.

It is observed that from Table 2.6.1;

- Smaller MSE frequency of ORR estimator is greater than that of the OLS, OJR, M-estimator and JRM estimators in the presence of only multicollinearity.
- In the presence of multicollinearity and outlier, the performance of JRM estimator is better than OLS, ORR, OJR and M-estimators.
- Smaller MSE frequency of JRM estimator increases with increase in degree of multicollinearity and error variance (σ^2).
- More than 80 % times MSE of JRM estimator is small as compare to other estimators in the presence of severe multicollinearity and large variance (σ^2) for outlier data.

2.6.2. Comparison of MSE

The simulation experiment given in 2.6.1 is repeated 10,000 times for above two cases and the MSE of each estimator is computed as

$$MSE(\hat{\alpha}_i) = \frac{1}{10000}\sum_{j=1}^{10000}\left(\hat{\alpha}_{ij} - \alpha_i\right)^2 \qquad (2.6.2)$$

where, $\hat{\alpha}_{ij}$ denote the estimate of the ith parameter in jth replication and α_i is a true parameter value. MSE ratio of JRM estimator over M-estimator, OJR, ORR and OLS estimator is tabulated in Table 2.6.2.

Table 2.6.2

Ratio of MSE of JRM, M-estimator, OJR, ORR and OLS estimators.

ρ	σ^2	$\frac{JRM}{M-est}$	$\frac{JRM}{OJR}$	$\frac{JRM}{ORR}$	$\frac{JRM}{OLS}$	$\frac{JRM}{M-est}$	$\frac{JRM}{OJR}$	$\frac{JRM}{ORR}$	$\frac{JRM}{OLS}$
		Case I				Case II			
					n = 20				
0.9	1	0.9996	1.0697	1.0891	1.0693	0.9996	0.0002	0.0002	0.0002
	9	0.9828	1.0700	1.2079	1.0510	0.9844	0.0017	0.0020	0.0017
	25	0.9395	1.0731	1.3387	1.0063	0.9472	0.0043	0.0056	0.0039
	100	0.8288	1.0688	1.5526	0.8784	0.8528	0.0114	0.0181	0.0084
	400	0.7416	1.0768	1.7004	0.7855	0.7607	0.0206	0.0360	0.0118
	900	0.7188	1.0879	1.7508	0.7686	0.7329	0.0233	0.0412	0.0126
0.99	1	0.9801	1.0662	1.2157	1.0441	0.9831	0.0002	0.0002	0.0002
	9	0.8396	1.0805	1.5643	0.8998	0.8682	0.0018	0.0027	0.0014
	25	0.7641	1.0814	1.6891	0.8144	0.8099	0.0047	0.0078	0.0031
	100	0.7140	1.0931	1.7726	0.7649	0.7587	0.0141	0.0246	0.0076
	400	0.6994	1.0918	1.7999	0.7470	0.7201	0.0219	0.0394	0.0109
	900	0.6958	1.0933	1.8024	0.7432	0.7004	0.0251	0.0456	0.0122
0.999	1	0.8309	1.0766	1.5707	0.8867	0.8669	0.0002	0.0003	0.0002
	9	0.7135	1.0947	1.7828	0.7652	0.7744	0.0021	0.0034	0.0012
	25	0.6963	1.0891	1.7985	0.7416	0.7604	0.0053	0.0090	0.0030
	100	0.6921	1.0957	1.8149	0.7401	0.7399	0.0136	0.0242	0.0070
	400	0.6901	1.0993	1.8175	0.7388	0.7103	0.0222	0.0404	0.0108
	900	0.6949	1.0917	1.7975	0.7413	0.7022	0.0243	0.0443	0.0120
					n = 30				
0.9	1	0.9999	1.0585	1.0701	1.0583	0.9999	0.0003	0.0003	0.0003
	9	0.9932	1.0645	1.1586	1.0572	0.9933	0.0021	0.0024	0.0021
	25	0.9663	1.0578	1.2682	1.0215	0.9688	0.0050	0.0063	0.0048
	100	0.8641	1.0653	1.5212	0.9161	0.8749	0.0132	0.0207	0.0106
	400	0.7420	1.0731	1.7342	0.7864	0.7517	0.0223	0.0404	0.0134
	900	0.7004	1.0822	1.8223	0.7441	0.7085	0.0259	0.0484	0.0134
0.99	1	0.9922	1.0634	1.1639	1.0549	0.9924	0.0002	0.0003	0.0002
	9	0.8767	1.0663	1.5100	0.9307	0.8904	0.0022	0.0033	0.0018
	25	0.7780	1.0719	1.6926	0.8263	0.8048	0.0057	0.0099	0.0039
	100	0.6914	1.0779	1.8437	0.7319	0.7191	0.0161	0.0302	0.0084
	400	0.6659	1.0842	1.8929	0.7073	0.6764	0.0244	0.0469	0.0116
	900	0.6556	1.0877	1.9132	0.6982	0.6594	0.0275	0.0530	0.0127
0.999	1	0.8646	1.0622	1.5234	0.9139	0.8827	0.0003	0.0004	0.0002
	9	0.6979	1.0877	1.8469	0.7469	0.7464	0.0027	0.0049	0.0015
	25	0.6730	1.0808	1.8708	0.7139	0.7109	0.0066	0.0119	0.0034
	100	0.6586	1.0819	1.9021	0.6979	0.6874	0.0164	0.0312	0.0079
	400	0.6575	1.0854	1.9035	0.6982	0.6613	0.0248	0.0483	0.0113
	900	0.6571	1.0854	1.9080	0.6972	0.6538	0.0278	0.0537	0.0126
					n = 50				
0.9	1	1.0000	1.0590	1.0660	1.0589	1.0000	0.0004	0.0004	0.0004
	9	0.9977	1.0535	1.1098	1.0511	0.9977	0.0030	0.0032	0.0030
	25	0.9869	1.0557	1.1942	1.0417	0.9873	0.0069	0.0081	0.0068
	100	0.9132	1.0557	1.4278	0.9622	0.9169	0.0161	0.0238	0.0141
	400	0.7753	1.0698	1.7102	0.8237	0.7751	0.0261	0.0472	0.0173
	900	0.7027	1.0690	1.8198	0.7422	0.7057	0.0322	0.0621	0.0177
0.99	1	0.9973	1.0548	1.1175	1.0520	0.9974	0.0004	0.0004	0.0004
	9	0.9257	1.0612	1.4119	0.9811	0.9295	0.0030	0.0043	0.0027
	25	0.8208	1.0640	1.6392	0.8688	0.8328	0.0072	0.0124	0.0054
	100	0.6962	1.0717	1.8618	0.7367	0.7143	0.0196	0.0381	0.0108
	400	0.6471	1.0798	1.9497	0.6860	0.6453	0.0314	0.0641	0.0141
	900	0.6318	1.0825	1.9875	0.6703	0.6374	0.0356	0.0723	0.0160
0.999	1	0.9175	1.0574	1.4300	0.9685	0.9211	0.0004	0.0005	0.0003
	9	0.7053	1.0703	1.8439	0.7453	0.7261	0.0035	0.0068	0.0020
	25	0.6554	1.0763	1.9402	0.6934	0.6812	0.0088	0.0172	0.0045
	100	0.6312	1.0736	1.9709	0.6651	0.6464	0.0213	0.0431	0.0095
	400	0.6271	1.0805	1.9879	0.6636	0.6314	0.0333	0.0681	0.0146
	900	0.6211	1.0797	2.0016	0.6565	0.6275	0.0367	0.0750	0.0158

	n = 100							
0.9	1	1.0000	1.0559	1.0593	1.0559	1.0000	0.0006	0.0006	0.0006
	9	0.9995	1.0486	1.0763	1.0481	0.9995	0.0052	0.0054	0.0052
	25	0.9965	1.0521	1.1270	1.0484	0.9965	0.0113	0.0124	0.0113
	100	0.9653	1.0596	1.3011	1.0224	0.9653	0.0241	0.0315	0.0229
	400	0.8452	1.0592	1.5891	0.8915	0.8423	0.0381	0.0647	0.0293
	900	0.7503	1.0616	1.7547	0.7904	0.7545	0.0449	0.0847	0.0282
0.99	1	0.9994	1.0546	1.0847	1.0539	0.9994	0.0006	0.0006	0.0006
	9	0.9713	1.0515	1.2718	1.0211	0.9722	0.0050	0.0064	0.0048
	25	0.8981	1.0556	1.4864	0.9463	0.9019	0.0114	0.0177	0.0098
	100	0.7415	1.0637	1.8023	0.7824	0.7435	0.0291	0.0564	0.0179
	400	0.6472	1.0728	1.9759	0.6843	0.6433	0.0467	0.0973	0.0220
	900	0.6190	1.0697	2.0193	0.6530	0.6207	0.0532	0.1123	0.0233
0.999	1	0.9671	1.0564	1.2945	1.0211	0.9681	0.0006	0.0008	0.0006
	9	0.7519	1.0621	1.7827	0.7923	0.7640	0.0055	0.0104	0.0036
	25	0.6639	1.0659	1.9371	0.6998	0.6802	0.0144	0.0293	0.0075
	100	0.6136	1.0754	2.0454	0.6490	0.6225	0.0335	0.0707	0.0145
	400	0.6008	1.0706	2.0533	0.6320	0.6018	0.0503	0.1070	0.0207
	900	0.6030	1.0754	2.0633	0.6365	0.5996	0.0534	0.1138	0.0221

From Table 2.6.2,

- When there is no outlier in the data, MSE ratio of JRM over OJR and ORR estimator is greater than one.
- When there is no outlier in the data, efficiency of ORR estimator over the JRM estimator increases when the degree of multicollinearity and error variance increases.
- For outlier data, JRM estimator is more efficient than OLS, ORR and OJR estimators.
- For fixed sample size, efficiency of JRM estimator over M-estimator increases with increase in the degree of multicollinearity and error variance.

2.6.3. Comparison of MSE for More Than One Outlier

The same simulation experiment is repeated 10,000 times for all possible combinations of ρ and σ^2 for sample size 30 by introducing two and three outliers. The MSE ratio of JRM over M-estimator, OJR, ORR and OLS estimators is computed and reported in the following table.

Table 2.6.3

Ratio of MSE of estimators in presence of multicollinearity and outliers.

ρ	σ^2	With 2 outliers				With 3 outliers			
		$\frac{JRM}{M-est}$	$\frac{JRM}{OJR}$	$\frac{JRM}{ORR}$	$\frac{JRM}{OLS}$	$\frac{JRM}{M-est}$	$\frac{JRM}{OJR}$	$\frac{JRM}{ORR}$	$\frac{JRM}{OLS}$
0.9	1	0.9999	0.0001	0.0001	0.0001	0.9999	0.0001	0.0001	0.0001
	9	0.9933	0.0012	0.0013	0.0012	0.9933	0.0009	0.0010	0.0009
	25	0.9712	0.0030	0.0037	0.0028	0.9719	0.0022	0.0027	0.0021
	100	0.8859	0.0081	0.0124	0.0066	0.8951	0.0064	0.0096	0.0054
	400	0.7629	0.0147	0.0260	0.0090	0.7766	0.0121	0.0212	0.0076
	900	0.7135	0.0172	0.0320	0.0092	0.7146	0.0141	0.0258	0.0077
0.99	1	0.9926	0.0001	0.0002	0.0001	0.9923	0.0001	0.0001	0.0001
	9	0.8974	0.0012	0.0018	0.0010	0.9057	0.0009	0.0013	0.0008
	25	0.8218	0.0031	0.0052	0.0022	0.8382	0.0024	0.0039	0.0018
	100	0.7480	0.0095	0.0172	0.0054	0.7702	0.0073	0.0128	0.0044
	400	0.6957	0.0163	0.0307	0.0080	0.7130	0.0132	0.0242	0.0068
	900	0.6704	0.0181	0.0348	0.0086	0.6865	0.0156	0.0295	0.0075
0.999	1	0.8923	0.0001	0.0002	0.0001	0.8988	0.0001	0.0002	0.0001
	9	0.7680	0.0014	0.0024	0.0008	0.7926	0.0011	0.0018	0.0007
	25	0.7476	0.0036	0.0063	0.0021	0.7626	0.0026	0.0045	0.0016
	100	0.7121	0.0097	0.0178	0.0050	0.7449	0.0077	0.0137	0.0043
	400	0.6867	0.0165	0.0313	0.0079	0.6990	0.0135	0.0250	0.0067
	900	0.6642	0.0188	0.0365	0.0085	0.6795	0.0156	0.0296	0.0074

From Table 2.6.3 it is clear that

- JRM estimator becomes more and more efficient as compared to OJR, ORR and OLS estimators as the number of outliers in the data increases.
- JRM estimator is better than M-estimator, OJR, ORR and OLS estimators for different degrees of multicollinearity and number of outliers in the data.

The same fact is also depicted graphically. Figure 2.6.1 to 2.6.6 show MSE ratio of JRM to OJR, ORR and OLS for σ^2 = 1 and 9, with different degrees of multicollinearity and number of outliers.

Figure 2.6.7 to 2.6.12 present the graph of MSE ratio of JRM estimator with M-estimator versus degree of multicollinearity for one, two and three outliers. It shows that for various values of σ^2, efficiency of JRM estimator over M-estimator increases as degree of multicollinearity increases. It is also clear that, as error variance increases, JRM estimator becomes more efficient than M-estimator.

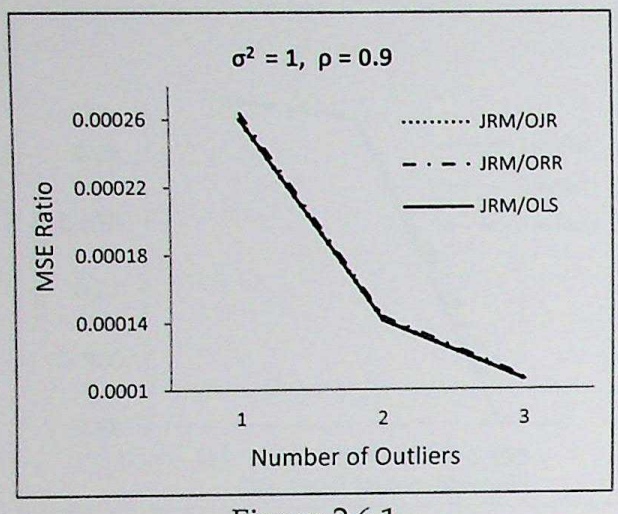


Figure 2.6.1

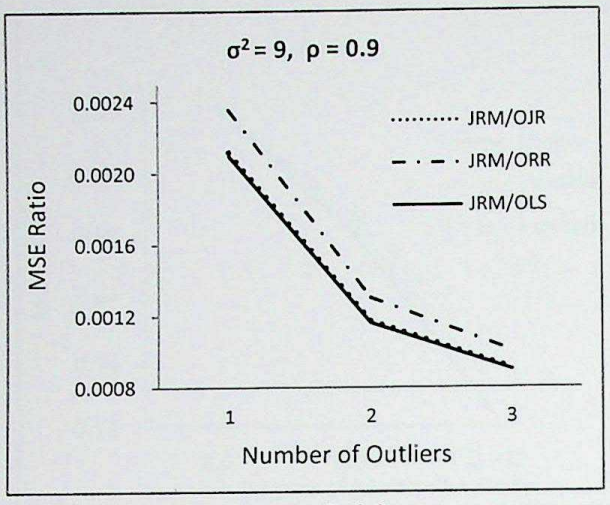


Figure 2.6.2

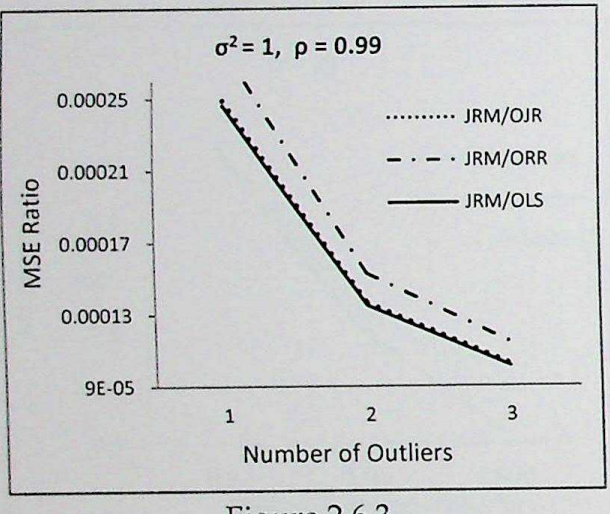


Figure 2.6.3

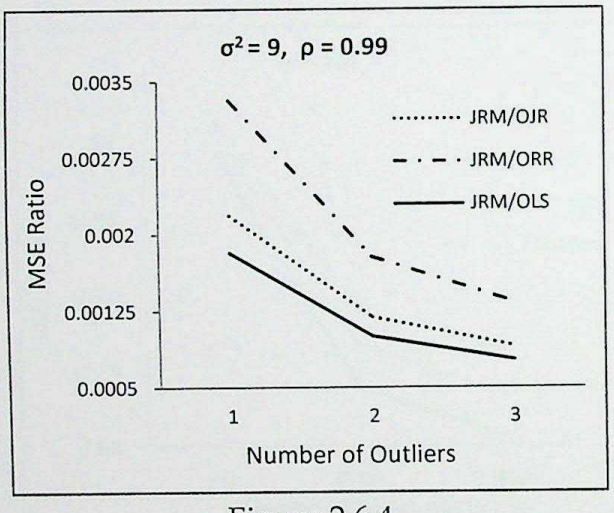


Figure 2.6.4

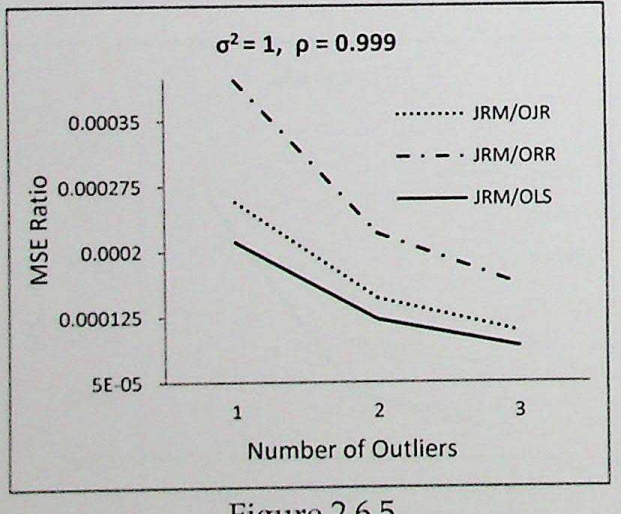


Figure 2.6.5

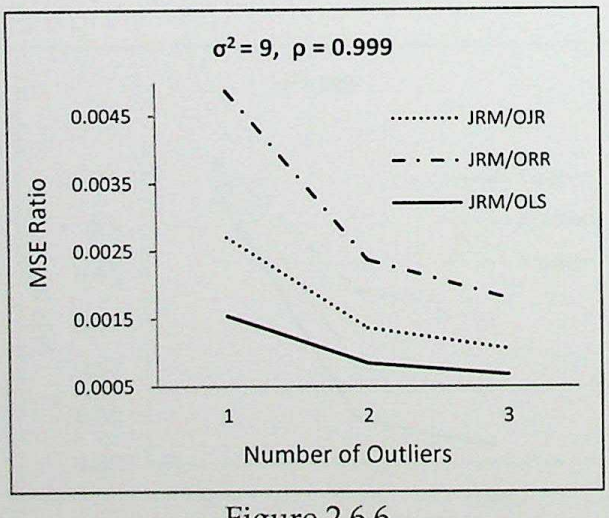


Figure 2.6.6

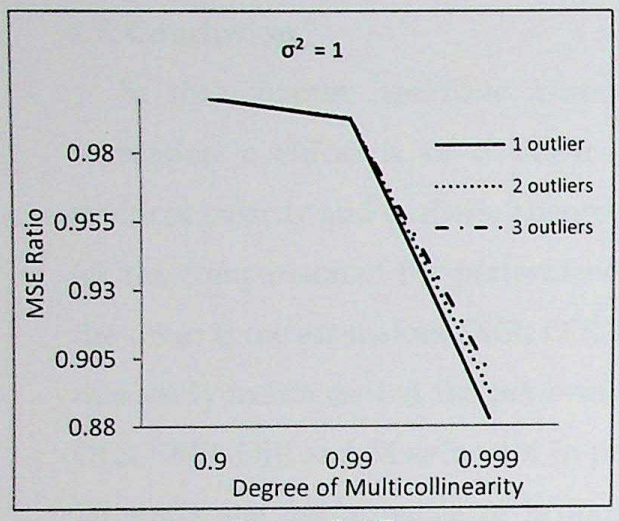


Figure 2.6.7

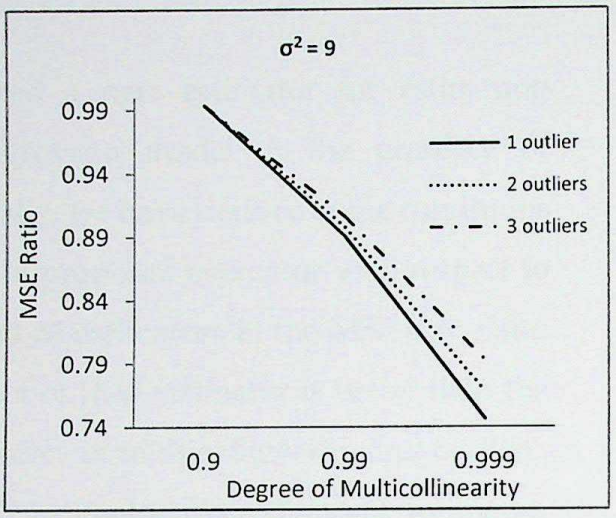


Figure 2.6.8

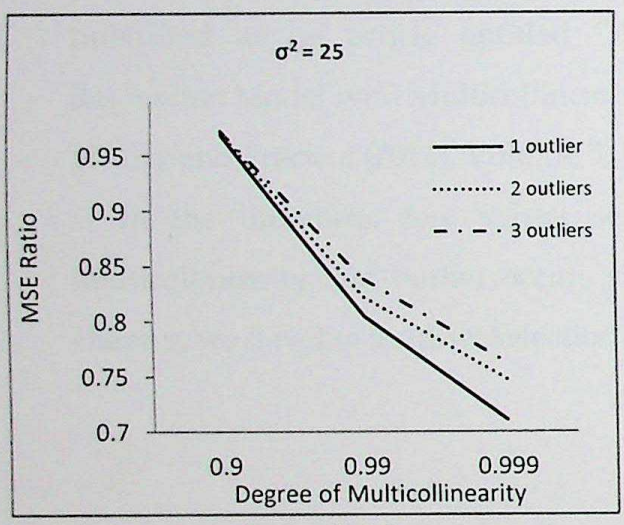


Figure 2.6.9

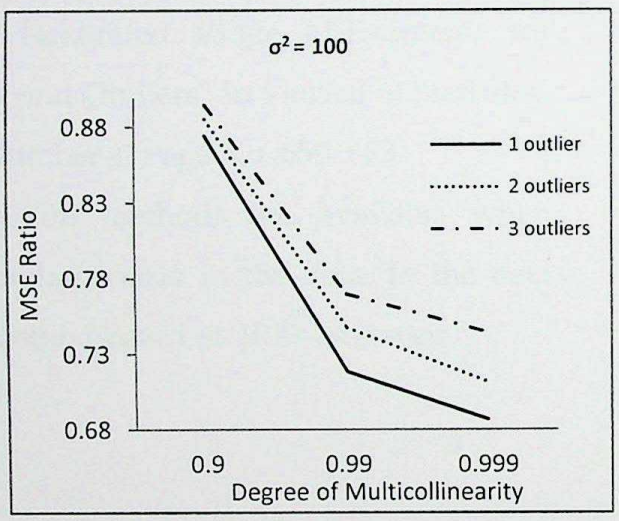


Figure 2.6.10

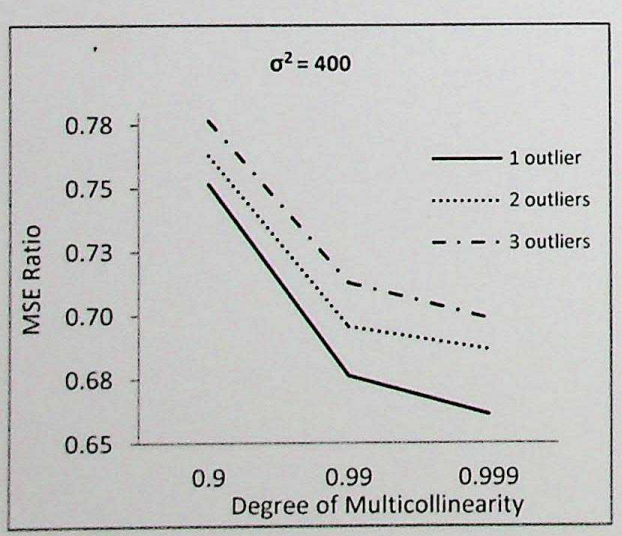


Figure 2.6.11

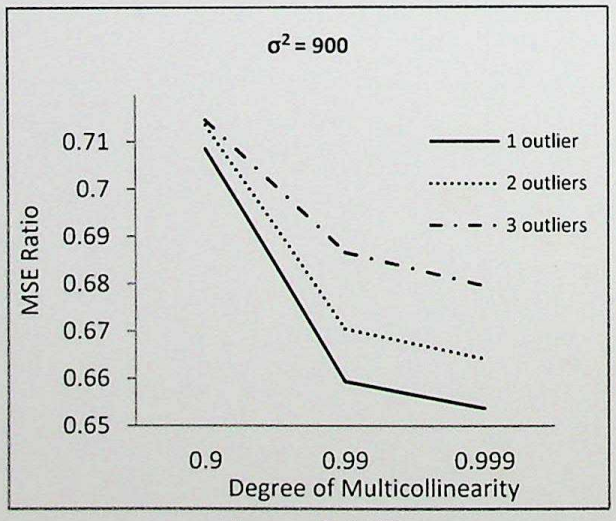


Figure 2.6.12

2.7. Conclusion

In this chapter, we have introduced a new estimator for estimating regression coefficients of a linear regression model in the presence of multicollinearity and outliers. Theoretically, we have derived some conditions for the comparison of the performance of proposed estimator with respect to the other three estimators ORR, OJR and M-estimators in the MSE sense and this study indicates that the performance of JRM estimator is better than the OLS, ORR, OJR and M-estimator in presence of multicollinearity and outliers. To study the performance of proposed estimator over different estimator, Matlab codes are used and are given in Appendix - I. A part of this chapter is published as an article entitled "A Jackknifed Ridge M-Estimator for Regression Model with Multicollinearity and Outliers" in Journal of Statistical Theory and Practice (2011), Volume 5, Number 4, Page No. 659- 673.

In the literature, few subset selection methods are available when multicollinearity and outlier occurs simultaneously in the data. In the next chapter, we develop a subset selection method based on JRM estimator.