Thailand Statistician
July 2010; 8(2) : 183-205
http://statassoc.or.th
Contributed paper

Robust Estimation of Regression Coefficients with Outliers

Pimpan Ampanthong * [a] and Prachoom Suwattee [b]

[a] Department of Mathematics, Faculty of Science and Technology, Rajamangala University of Technology Suvarnabhumi, Supan Buri Campus, 72130, Thailand.

[b] School of Applied Statistics, National Institute of Development Administration, Bangkapi, Bangkok, 10240, Thailand.

***Author for correspondence**; e-mail: pim_pimpan@hotmail.com

Received: 26 February 2010
Accepted: 24 June 2010

Abstract

This study concentrates on the construction of weights for the estimation of regression coefficients in multiple linear regression with outliers using a new proposed influence function. Set of weights, modified weights one (MW1) are obtained from newly modified influence function. The proposed estimates are applied in the M-estimator of the regression coefficients with outliers and compared to ordinary least-squares (OLS) and other M-estimates by simulation. Results of the estimates indicate that the new weights out perform the least squares estimates and the other M-estimates. As for X-outliers and XY-outliers, it is found that the proposed estimates using MW out perform the least squares estimates for all sample sizes. It also gives high values of R^2 and low MSE at different percentages of outliers as well.

Keywords: influence functions, M-estimates, multiple linear regression, outliers, robust regression, weighted least-squares.

1. Introduction

Linear regression models are commonly used to study the relationship between a response variable and independent variables. A linear model is one of the form $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$ where $\underline{y}$ is an $n\times 1$ vector of observed values of the dependent variable, $X = \left(\underline{x}_1, \underline{x}_2, ..., \underline{x}_n\right)'$, an $n\times p$ matrix with row vector $\underline{x}_i' = \left(x_{i1}, x_{i2}, ..., x_{ip}\right)'$ of components x_{ij} of the regressors, $i = 1,2,...,n.$, $j = 1,2,...,p$, $\underline{\beta}$ a $p\times 1$ vector of unknown parameters, and $\underline{\varepsilon}$ an $n\times 1$ vector of errors. If $\underline{\varepsilon}$ follows a normal $N(\underline{0},\sigma^2 I)$ distribution, then the ordinary least-squares (OLS) estimate of $\underline{\beta}$ turn-out to be the best linear unbiased estimates (BLUE) of $\underline{\beta}$ according to the Gauss-Markov theorem. The OLS estimate is BLUE also without residual value of $\underline{\beta}$ which minimizes the sum of squares of the residuals, $\underset{\hat{\beta}}{Min}\sum_{i=1}^{n}\left(y_i - \underline{x}'_i\underline{\hat{\beta}}\right)^2$. In many situations, the sample data violate the above assumptions. In particular, outliers have effects on the OLS estimate. In several cases, the OLS estimate may not be appropriate and not sensitive to the presence of outliers. When the sample data contain some outliers, they may have strong influence on regression analysis. Robust regression is an alternative to find estimates of the coefficient $\underline{\beta}$ for data with outliers. Robust regression reduces the effects of outliers instead of ignoring them. In addition to insensitivity to outliers, a robust estimation procedure should produce essentially the same results as OLS when the underlying assumption is true with no outliers. One of the three types of outliers may exist with the sample data in regression. There are X-outliers, Y-outliers and XY-outliers. They are explained in many textbooks on robust estimations such as, Thomas [1], Draper [2], Montgomery et al. [3]. It is important to have some special treatment for the analysis of the data with outliers. This study focuses on a modification of the $\rho - function$ and the estimation of the regression coefficients with outliers. Robust estimates of regression coefficients were introduced by many authors. They are in the class of M-estimators, high breakdown estimators, bounded-influence estimators and others. Some of the most widely known regression estimators are the least median squares estimator (LMS), the least trimmed square estimator (LTS) , the S-estimators [4], the MM-estimator [5], the generalized M-estimator (GM) [6], Krasker and Welsch [7] handled all X-outliers by replacing the least squares errors by studentized residuals, standardized residuals, or DEFFITi and PRESS residuals. Simpson Ruppert and Carroll [8] used the compound estimator. In this paper, we focus on finding proper weights from a nearly proposed $\rho-function$ for the estimation of regression coefficients with outliers using M-estimates as numerical example shows the idea of weight construction and the weight application in the estimates of regression coefficients with outliers. A simulation study is carried out to compare the R^2 and MSE with those from the OLS and other M-estimates.

2. Weight Construction

A $\rho-function$ is used to derive weights for the sample data so that the outliers will be less important. Several $\rho-functions$ have been proposed by many authors. Some of these are Andrew's, Hampel's, Huber's, Tukey's and Winsorized's functions. The $\rho-functions$ are symmetric, bounded and nondecreasing with unique minimum at zero. The $\rho-functions$ should have the following properties[1]: if $0 \le u \le v$, then $0 \le \rho(u) \le \rho(v)$, where u and v are some real numbers and if $a = \sup \rho(\underline{r})$ then $0 < a < \infty$, where $r_i = y_i - \underline{x}'_i \underline{\hat{\beta}}$, $\underline{x}'_i = (x_{i1}, x_{i2}, ..., x_{ip})'$ for $i = 1,2,...,n$ and y_i by least-squares based on the whole data set of n observations. The M-estimators use weights which minimize the sum of $\rho\left(\frac{y_i - \underline{x}'_i \underline{\hat{\beta}}}{\hat{\sigma}}\right)$, i.e. $\underset{\underline{\hat{\beta}}}{Min} \sum_{i=1}^{n} \rho\left(\frac{y_i - \underline{x}'_i \underline{\hat{\beta}}}{\hat{\sigma}}\right)$ where $\rho-function$ is a symmetric, bounded and nondecreasing function with minimum at 0. The most popular choice for $\hat{\sigma}$ is an estimate of the median absolute deviation defined as $\hat{\sigma} = \left(median\left|r_i - med(r_i)\right|\right)/0.6745$. The constant 0.6745 makes $\hat{\sigma}$ an approximately unbiased estimate of σ, if n is large and error distribution is normal. Taking the first and second partial derivatives of the sum of residuals with respect to the regression coefficients $\hat{\beta}_j$, for $j = 1,2,...,p$, setting them equal to zero and solving leads for $\underline{\beta}$ to desired results. This gives two systems of equations:

$$\frac{d}{d\hat{\beta}_j}\left(\sum_{i=1}^{n} \rho\left(\frac{y_i - \underline{x}'_i \underline{\hat{\beta}}}{\hat{\sigma}}\right)\right) = \frac{d}{dr_i}\left(\sum_{i=1}^{n} \rho\left(\frac{y_i - \underline{x}'_i \underline{\hat{\beta}}}{\hat{\sigma}}\right) . \frac{dr_i}{d\hat{\beta}_j}\right) = 0 .$$

[1] $\rho-functions$ with these properties yield the proper weights to be applied in the estimation in the class of M-estimators. See Yohai, 1987: 642.

Hence, $$\sum_{i=1}^{n}\rho'\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)\left(-x_{ij}\right)=0\,. \qquad \text{.......... (1)}$$

Also $$\frac{d}{d\hat{\beta}_j}\left(\sum_{i=1}^{n}\rho'\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)\right)=\frac{d}{dr_i}\left(\sum_{i=1}^{n}\rho'\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right).\frac{dr_i}{d\hat{\beta}_j}\right)>0\,.$$

.... (2)

Hence, $\sum_{i=1}^{n}\rho''\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)\left(-x_{ij}\right)>0$, where $\rho'\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)$ and $\rho''\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)$ are nonlinear, satisfying the properties is Montgomery et al. [3] and Maronna, Martin and Yohai [9]. The estimates of $\underline{\beta}$ is the solution of (1) and (2) obtained by iteratively reweighted least squares. Suppose that $\hat{\underline{\beta}}$ and $\hat{\sigma}$ are the initial estimates by least squares. Instead of (1) and (2), we can equivalently write

$$\sum_{i=1}^{n}\sum_{j=1}^{p}\rho'\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)\left(-x_{ij}\right)\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)/\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)=0 \qquad \text{, ... (3)}$$

or $$\sum_{i=1}^{n}\sum_{j=1}^{p}W\left(-x_{ij}\right)\left(\left(y_i-\underline{x}'_i\hat{\underline{\beta}}\right)/\hat{\sigma}\right)=0\,, \text{ for } j=1,2,\ldots,p \qquad \text{, ... (4)}$$

where $$w_{ii}=\begin{cases}\rho'\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right)/\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right) & ,\left|\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right|\neq 0,\\ \rho''\left(\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right) & ,\left|\frac{y_i-\underline{x}'_i\hat{\underline{\beta}}}{\hat{\sigma}}\right|=0.\end{cases} \qquad \text{, ... (5)}$$

which, matrix weight function is $W=\begin{bmatrix} w_{11} & 0 & \ldots & 0\\ 0 & w_{22} & \ldots & 0\\ \vdots & \vdots & \ldots & \vdots\\ 0 & 0 & \ldots & w_{nn}\end{bmatrix}$, $0\le w_{ii}\le 1$ and $w_{ij}=0$ with $i\neq j$ for $i=1,2,\ldots,n$ and $j=1,2,\ldots,p$. Construction of weights depends on the current residuals. The weights needed in the M-estimates can be derived by the following steps:

1) Choose a $\rho-function$ with the desirable properties where c is a constant chosen to be 3 as suggested by many authors.

2) Take the first partial derivatives of the $\rho - function$ with respect to β_j ($j = 1,2,\ldots,p$) and equate them to zero, giving $\rho'\left(\frac{y_i - \underline{x}'_i\underline{\hat{\beta}}}{\hat{\sigma}}\right) / \left(\frac{y_i - \underline{x}'_i\underline{\hat{\beta}}}{\hat{\sigma}}\right)$.

3) Take the second partial derivatives of the $\rho - function$ with respect to β_j $(j = 1,2,\ldots,p)$ and equate them to zero, giving $\rho''\left(\frac{y_i - \underline{x}'_i\underline{\hat{\beta}}}{\hat{\sigma}}\right)$.

4) Solve the equations for the estimation of regression coefficients with outliers by iteratively reweighted least squares and find a weight function from 2) and 3). We get the modified weights function.

3. Estimation of Regression Coefficients with Outliers

A new influence function ρ in (6) is introduced to obtain proper weights for the residuals so that the outliers will be less important. It is

$$\rho_1\left(\frac{r_i}{\hat{\sigma}}\right) = \begin{cases} \left(a\left(\frac{r_i}{\hat{\sigma}}\right)^2 / 2\right) + \left(\left(\frac{r_i}{\hat{\sigma}}\right)^6 / 6c^4\right), \left|\frac{r_i}{\hat{\sigma}}\right| \le c \quad , \\ \left(\left|\frac{r_i}{\hat{\sigma}}\right| / c\right) \qquad , \left|\frac{r_i}{\hat{\sigma}}\right| > c \quad , \end{cases} \qquad ,\ldots\ldots\ldots\ldots (6)$$

where the i^{th} residuals, $r_i = y_i - \underline{x}'_i\underline{\hat{\beta}}$, $\underline{x}'_i = (x_{i1}, x_{i2}, ..., x_{ip})'$ for $i = 1,2,...,n.$, a and c are constants. The constant c is chosen to be 3 as suggested by many authors[2]., when the estimates of regression coefficients with outliers approach these weighted least squares estimates of $\underline{\beta}$, and $\hat{\sigma}$ is an estimate of σ. The function in (6) is used to obtain a diagonal weight matrix W whose diagonal elements are nonincreasing. Outlying observations will receive smaller weights by this process. Construction of weights needed in the M-estimates depends on the current residuals r_i.

Theorem 1 The proposed $\rho_1 - function$ (6) gives a diagonal weight matrix, called modified weight one (MW1) with diagonal elements,

---

[2] Different values of c are suggested in the literature; Rousseeuw and Leroy used $c = 2.5$ (Rousseeuw and Leroy [4]), while Huber gave $c = 2.0$ (Montgomery et al. [3]), Andrews suggested $c = 3.14$ (Huber [10]) and Tukey gave $c = 1.0$ (Tukey [11]).

$$w_{ii}=\begin{cases}1/2+\left(\left(\dfrac{r_i}{\hat{\sigma}}\right)^4/c^4\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|\le c,\\ 1+\left(5\left(\dfrac{r_i}{\hat{\sigma}}\right)^4/c^4\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|=0,\\ \left(1/c\left|\dfrac{r_i}{\hat{\sigma}}\right|\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|>c.\end{cases} \quad ,\ldots\ldots\ldots\ldots (7)$$

Proof Take the first and second partial derivatives of the $\rho_1-function$ (6) with respect to the coefficients $\hat{\beta}_j$, for $j=1,2,...,p$. Setting the first partial derivatives equal to 0 and a is a constant 0.5, produces a system of p estimating equations for the estimates of the coefficients.

$$\rho_1'\left(\frac{r_i}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho_1\left(\frac{r_i}{\hat{\sigma}}\right)\right)=\begin{cases}\left((0.5)\left(\dfrac{r_i}{\hat{\sigma}}\right)\right)+\left(\left(\dfrac{r_i}{\hat{\sigma}}\right)^5/c^4\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|\le c \quad ,\\ \left(sign\left(\dfrac{r_i}{\hat{\sigma}}\right)/c\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|>c \quad .\end{cases} \quad ,\ldots\ldots\ldots\ldots (8)$$

The second derivative of $\rho_1-function$ with respect to $\hat{\beta}_j$, $j=1,2,...,p$ and a is a constant 1.0, as follow,

$$\rho_1''\left(\frac{r_i}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho_1'\left(\frac{r_i}{\hat{\sigma}}\right)\right)=\left\{1+\left(5\left(\frac{r_i}{\hat{\sigma}}\right)^4/c^4\right)\right. \quad ,\left|\frac{r_i}{\hat{\sigma}}\right|=0 \quad . \quad ,\ldots\ldots\ldots (9)$$

The weight function in (7) is obtained by iteratively reweighted least squares.

The weight function has elements is $=\begin{cases}1/2+\left(\left(\dfrac{r_i}{\hat{\sigma}}\right)^4/c^4\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|\le c,\\ 1+\left(5\left(\dfrac{r_i}{\hat{\sigma}}\right)^4/c^4\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|=0,\\ \left(1/c\left|\dfrac{r_i}{\hat{\sigma}}\right|\right) & ,\left|\dfrac{r_i}{\hat{\sigma}}\right|>c.\end{cases}$. So the MW1 gives the weight function (7) using (5).

Theorem 2 If the $\rho-function$ is symmetric, bounded and nondecreasing with unique minimum at zero, then weight matrix W has diagonal elements $w_{ii}=\rho'\left(\frac{r_i}{\hat{\sigma}}\right)/\left(\frac{r_i}{\hat{\sigma}}\right)$, which is bounded and monotone decreasing for $r_i>0$ and $\hat{\sigma}>0$.

Proof To simplify the proof we assume $\hat{\sigma}=1$.

Let $$U(r_i)=\sum_{i=1}^{n}U\left(y_i-\underline{x}'_i\underline{\hat{\beta}}\right)$$,………. (10)

where $U(r_i)$ is a quadratic function of a variable k_i as $\underline{x}'_i=\left(x_{i1},x_{i2},...,x_{ip}\right)'$ and replace $\rho-function$ by $$U(k_i)=a+ck_i+\frac{1}{2}k_i^2$$,………… (11)

with a and c to be determined such that, assume $U(k_i)\ge\rho(k_i)$, for all k_i, and $U(r_i)=\rho(r_i)$, where the vector of residual $r_i'=(r_1,r_2,...,r_n)'$, for all $i=1,2,...,n.$

These conditions imply that U and ρ have a common at $r_i=y_i-\underline{x}'_i\underline{\hat{\beta}}$.

$$U'(r_i)=c+r_i=\rho'(r_i).$$

Hence $\frac{c}{r_i}=\frac{\rho'(r_i)}{r_i}-\frac{r_i}{r_i}=w_{ii}-1$ and $a=\rho(r_i)-r_i\rho'(r_i)+\frac{1}{2}r_i^2$. The difference of functions

$$F(k_i)=U(k_i)-\rho(k_i)=\rho(r_i)-r_i\rho'(r_i)+\frac{1}{2}r_i^2+\left(\rho'(r_i)-r_i\right)k_i+\frac{1}{2}k_i^2-\rho(k_i)$$

Since element of weight function matrix $w_{ii}=\frac{\rho'(k_i)}{(k_i)}$ is bounded and monotone decreasing for $k_i\ge 0$ and the same holds for $k_i\le 0$ because satisfies the assumption symmetry, as $F(r_i)=F(-r_i)=0$ and $F'(r_i)=F'(-r_i)=0$.

Finally, $$F'(r_i)=\rho'(r_i)-r_i+k_i-\rho'(k_i).$$

It follows that $$F'(r_i)=\begin{cases}\le 0\\ \ge 0\end{cases}.$$,………. (12)

Finally, $F'(r_i)=\rho'(r_i)-r_i+k_i-\rho'(k_i)$. Since element of weight function matrix $w_{ii}=\frac{\rho'(k_i)}{(k_i)}$ is bounded and monotone decreasing for $k_i\ge 0$ as $F'(k_i)\ge F(r_i)=0$, the same holds for $k_i\le 0$ because symmetry.

4. Applying the Modified Influence Function

The robust estimate of $\underline{\beta}$ are the iteratively reweighted least squares. The weights will change the effects of outliers in a data set. The outlier detection methods in multiple linear regression models have been studied and compared by Ampanthong and

Suwattee ([12], [13], [14] and [15]). Applying the modified influence function, the normal equations may be written in matrix form, $(X'WX)\underline{\hat{\beta}}^* = X'W\underline{y}$, with W a diagonal matrix of weights. Given a data set, we can compute the estimated values of $\underline{\beta}$ by a one step regression following the steps:

1) Find initial estimates $\underline{\hat{\beta}}$ of $\underline{\beta}$ and $\hat{\sigma}^2$ of σ^2 by least-squares.

2) Compute the residuals $r_i = y_i - \underline{x}'_i \underline{\hat{\beta}}$, for $i = 1, 2, ..., n$, and find the weights as described earlier.

3) Find the estimates of the regression coefficients with outliers by the generated weighted least-squares method.

4) Continue 2), 3) until the maximum difference of two successive pairs of estimates are arbitrarily small, i.e. $\max_{1 \le i \le p} \left| \hat{\beta}_j^* - \hat{\beta}_j^{**} \right| < \delta$, where $\hat{\beta}_j^*$ and $\hat{\beta}_j^{**}$ are two successive estimates of β_j of all regression coefficients and $\delta > 0$. This guarantees that the two successive estimates almost agree with each other.

5. Properties of the New Estimators

The M-estimators applied as iteratively reweighted least squares transform the weight matrix of diagonal elements $w_{11}, w_{22}, ..., w_{nn}$. The new solution of weighted least-squares estimators of $\underline{\beta}$ are

$$\underline{\hat{\beta}}^* = \left(X'WX \right)^{-1} \left(X'W\underline{y} \right) \qquad \text{……….. (13)}$$

The weight function are calculated from the given $\rho-function$ in order to find the estimates $\underline{\hat{\beta}}^*$ of $\underline{\beta}$ as in (13). There are a number of possible estimates in the form of MW1 as in (7), which are equivalent to the best linear unbiased estimators of the regression coefficients. The estimator in (13) is in class of M-estimators. The M-estimator applied as iteratively reweighted least squares problem of set $\rho-function$ transform the weight function matrix of diagonal element $w_{11}, w_{22}, ..., w_{nn}$, then the new solution of weighted least-squares. A statistics is equivariance if it has the following three properties According to Rousseeuw, et. al.[4] 〈3, the M-estimators are regression, scale, and affine equivariant.

〈3 Equivariance. A statistics T is send to be equivariant if it has the following three properties

Lemma If a $\rho-function$ is location equivariant, scale equivariant and affine equivariant, then the estimating weight function has the same properties

Proof The estimating weight function matrix has the same properties

1) If $\rho-function$ is the location equivariant then

$$\rho\left(\frac{r^*_i}{\hat{\sigma}}\right)=\rho\left(\frac{y^*_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)=\rho\left(\frac{y_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)=\rho\left(\frac{y_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)+u \qquad \text{.......... (14)}$$

where $r^*_i=y^*_i-\underline{x}'_i\underline{\hat{\beta}}^*$, $y_i^*=y_i+u$ as u is any constant. The first and second derivative of $\rho-function$ with respect to $\hat{\beta}_j$, $j=1,2,...,p$ .are

$$\rho'\left(\frac{r^*_i}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho\left(\frac{r^*_i}{\hat{\sigma}}\right)\right)=\frac{d}{dr_i}\left(\rho\left(\frac{r^*_i}{\hat{\sigma}}\right)\right).\frac{dr_i}{d\hat{\beta}_j}=0$$

and $\rho''\left(\frac{r^*_i}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho'\left(\frac{r^*_i}{\hat{\sigma}}\right)\right)=\frac{d}{dr_i}\left(\rho'\left(\frac{r^*_i}{\hat{\sigma}}\right)\right).\frac{dr_i}{d\hat{\beta}_j}$, respectively.

The resulting weight function matrix W has diagonal elements

$$w_{ii}=\begin{cases}\rho'\left(\frac{y_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)\Big/\left(\frac{y_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right) & ,\left|\frac{y_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right|\neq 0,\\ \rho''\left(\frac{y_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right) & ,\left|\frac{y_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right|=0.\end{cases}$$

---

1) T is location equivariant, i.e. $T\left(\underline{x}'_i,y_i+u\right)=T\left(\left(x_{11},...,x_{1p},y_1+u\right)\right.$ $,\left(x_{21},...,x_{2p},y_2+u\right)$ $,...,\left(x_{n1},...,x_{np},y_n+u\right)$ for $i=1,2,...,n$ where $\underline{x}'_i=\left(\underline{x}_{i1},...,\underline{x}_{ip}\right)'$ and u is any constant.

2) T is scale equivariant, i.e. $T\left(\underline{x}'_i,cy_i\right)=T\left(\left(x_{11},...,x_{1p},cy_1\right)\right)$ $,\left(x_{21},...,x_{2p},cy_2\right)$ $,...,\left(x_{n1},...,x_{np},cy_n\right)$, for $i=1,2,...,n$ where $\underline{x}'_i=\left(\underline{x}_{i1},...,\underline{x}_{ip}\right)'$ and c is any constant.

3) T is affined equivariant, i.e. $T\left(\underline{x}'_i,cy_i+u\right)=T\left(\left(x_{11},...,x_{1p},cy_1+u\right)\right.$ $,\left(x_{21},...,x_{2p},cy_2+u\right)$ $,...,\left(x_{n1},...,x_{np},cy_n+u\right)$, for $i=1,2,...,n$ where $\underline{x}'_i=\left(\underline{x}_{i1},...,\underline{x}_{ip}\right)'$, c and u are any constants.

2) If $\rho-function$ is scale equivariant then

$$\rho\left(\frac{r_i^{**}}{\hat{\sigma}}\right)=\rho\left(\frac{y_i^{**}-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)=\rho\left(\frac{cy_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)=c\rho\left(\frac{y_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right), \qquad \text{……….} (15)$$

where $r_i^{**}=y_i^{**}-\underline{x}'_i\underline{\hat{\beta}}^*$, $y_i^{**}=cy_i$ as c is any constant.

The first and second derivative of $\rho-function$ with respect to $\hat{\beta}_j$, $j=1,2,...,p$ .are

$$\rho'\left(\frac{r_i^{**}}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho\left(\frac{r_i^{**}}{\hat{\sigma}}\right)\right)=\frac{d}{dr_i}\left(\rho\left(\frac{r_i^{**}}{\hat{\sigma}}\right)\right).\frac{dr_i}{d\hat{\beta}_j}=0 \text{ and}$$

$$\rho''\left(\frac{r_i^{**}}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho'\left(\frac{r_i^{**}}{\hat{\sigma}}\right)\right)=\frac{d}{dr_i}\left(\rho'\left(\frac{r_i^{**}}{\hat{\sigma}}\right)\right).\frac{dr_i}{d\hat{\beta}_j}\text{, respectively.}$$

The resulting weight function matrix W has diagonal elements

$$w_{ii}=\begin{cases}\rho'\left(\dfrac{cy_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)\Big/\left(\dfrac{cy_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right) & ,\left|\dfrac{cy_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right|\neq 0,\\ \rho''\left(\dfrac{cy_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right) & ,\left|\dfrac{cy_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right|=0.\end{cases}$$

and 3) If $\rho-function$ is the affined equivariant then

$$\rho\left(\frac{r_i^{***}}{\hat{\sigma}}\right)=\rho\left(\frac{y_i^{***}-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)=\rho\left(\frac{cy_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)=c\rho\left(\frac{y_i-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)+u, \qquad \text{……….} (16)$$

where $r_i^{***}=y_i^{***}-\underline{x}'_i\underline{\hat{\beta}}^*$, $y_i^{***}=cy_i+u$ as c and u are any constants.

The first and second derivative of $\rho-function$ with respect to $\hat{\beta}_j$, $j=1,2,...,p$ .are

$$\rho'\left(\frac{r_i^{***}}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho\left(\frac{r_i^{***}}{\hat{\sigma}}\right)\right)=\frac{d}{dr_i}\left(\rho\left(\frac{r_i^{***}}{\hat{\sigma}}\right)\right).\frac{dr_i}{d\hat{\beta}_j}=0 \text{ and}$$

$$\rho''\left(\frac{r_i^{***}}{\hat{\sigma}}\right)=\frac{d}{d\hat{\beta}_j}\left(\rho'\left(\frac{r_i^{***}}{\hat{\sigma}}\right)\right)=\frac{d}{dr_i}\left(\rho'\left(\frac{r_i^{***}}{\hat{\sigma}}\right)\right).\frac{dr_i}{d\hat{\beta}_j}\text{, respectively.}$$

The resulting weight function matrix W has diagonal elements

$$w_{ii}=\begin{cases}\rho'\left(\dfrac{cy_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right)\Big/\left(\dfrac{cy_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right) & ,\left|\dfrac{cy_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right|\neq 0,\\ \rho''\left(\dfrac{cy_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right) & ,\left|\dfrac{cy_i+u-\underline{x}'_i\underline{\hat{\beta}}^*}{\hat{\sigma}}\right|=0.\end{cases}$$

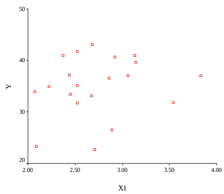
The weight function is calculated from the given $\rho-function$ in order to find the estimates $\underline{\hat{\beta}}^*$ of $\underline{\beta}$ as in (13). The estimates, depending on the normalized of residuals iterative method is needed. Initial values of regression coefficients are chosen as the first estimates. The residuals and scale estimate, $\hat{\sigma}$, are computed. This procedure is repeated using the residuals and scale estimate of regression coefficients with outliers from the previous iteration at each stage until stable convergence of the estimates is achieved. The weights are obtained from the data and the large residuals will be minimized. There are a number of possible estimates in the form of MW1 as in (7), which are equivalent to the best linear unbiased estimators of the regression coefficients. They also have the unbiased and the mean squares error for $\underline{\beta}$. So $\underline{\hat{\beta}}^*$ also have the above properties. Besides $\underline{\hat{\beta}}^*$ have the tendency to give small MSE and larger R^2 than other M-estimators and OLS estimators in multiple linear regression with outliers**.** The coefficient of determination is given by $R^2=\dfrac{(\underline{y}'WX)(X'WX)^{-1}}{(\underline{y}'W\underline{y})}$. Some the estimate of $\underline{\beta}$ in regression model builders prefer to use an adjusted R^2, denoted R^2_{adj} with

$$R^2_{adj}=\frac{(\underline{y}'WX)(X'WX)^{-1}\big/(p)}{(\underline{y}'W\underline{y})\big/(n-1)}$$

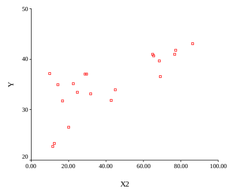
. Thus two way to assess the overall adequacy of the model are R^2 and the adjusted R^2_{adj}**.**

6. An Example

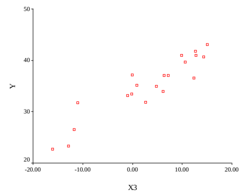
Data come from the Coleman data set given in Rousseeuw, P.J. and Leroy, A.M., (1987: 79). It is the six different variables and one response of verbal mean test score (Y). There are 20 cases from the Mid-Atlantic and New England states. There are five different independent variables, the staff salaries per pupil $(X1)$, the percent of white-collar fathers $(X2)$, the socioeconomic status composite deviation $(X3)$, the mean teacher's verbal test score $(X4)$ and the mean mother's educational level $(X5)$. In the data, so p=6, n=20, and it is well known that the observations 3, 11, 17 and 18 are outliers. Scatter plot is an important tool in analyzing the relationship between dependent against independent variables. That is show the scatter plot of Y against X_j, for $j = 1, 2, ..., 5$.



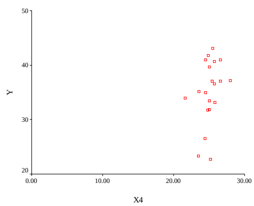
(a) Scatter Plot of Y vs. $X1$



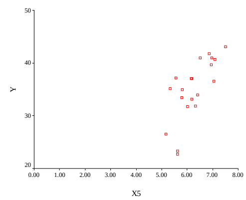
(b) Scatter Plot of Y vs. $X2$



(c) Scatter Plot of Y vs. $X3$



(d) Scatter Plot of Y vs. $X4$



(e) Scatter Plot of Y vs. $X5$

Figure 1. Scatter Plots of Dependent versus Independent Variables by Original Observations.

From the scatter plots in the Fig 1, there might be some outliers. There might be some outliers. The estimation of regression coefficients may depend on these points. When the sample contains outliers, the outliers may have large effects on the estimates and alternative approach to the problem should be applied to obtain better fit of the model or more precise estimate of $\underline{\beta}$. The robust estimates of $\underline{\beta}$ are the iteratively reweighted least squares. The weight functions will change the effects of outliers in a data set. To apply the proposed modified influence function**,** we first compute the estimated values of $\underline{\beta}$ and σ^2 by estimate least-squares. The least-squares fit of the regression model is

$$\hat{y}_i = 19.9485 - 1.7933\underline{x}_{i1} + 0.0436\underline{x}_{i2} + 0.5556\underline{x}_{i3} + 1.1101\underline{x}_{i4} - 1.8109\underline{x}_{i5}.$$

After obtaining the least-squares fit, we compute the standardized residuals $r_i/\hat{\sigma} = \left(y_i - \underline{x}'_i\underline{\hat{\beta}}\right)/\hat{\sigma}$, for $i = 1,2,...,20$. The most popular choice for $\hat{\sigma}$ is an estimate of the median of the absolute deviations of σ defined as, $\hat{\sigma} = \left(median\left|r_i - med(r_i)\right|\right)/0.6745$. From our data $\hat{\sigma} = \left(median\left|r_i - \text{-}106.4956\right|\right)/0.6745 \ = 0.0350$. Using MW1 the weights w_{ii} are found as follows:

$$w_{11} = \begin{cases} 1/2 + \left(\left(0.34877\right)^4/\left(3\right)^4\right) & , \\ 1 + \left(5\left(0.34877\right)^4/\left(3\right)^4\right) & , \\ \left(1/\left(3\right)*\left|0.34877\right|\right) & . \end{cases} = \begin{cases} 0.5001 & ,\left|0.34877\right| \le 3, \\ 1.0091 & ,\left|0.34877\right| = 0, \\ 1.7071 & ,\left|0.34877\right| > 3. \end{cases}$$

$$w_{22} = \begin{cases} 1/2 + \left(\left(0.1906\right)^4/\left(3\right)^4\right) & , \\ 1 + \left(5\left(0.1906\right)^4/\left(3\right)^4\right) & , \\ \left(1/\left(3\right)*\left|0.1906\right|\right) & . \end{cases} = \begin{cases} 0.5000 & ,\left|0.1906\right| \le 3, \\ 1.0000 & ,\left|0.1906\right| = 0, \\ 1.6961 & ,\left|0.1906\right| > 3. \end{cases}$$

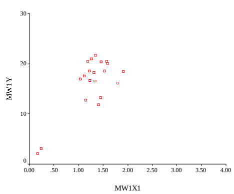
$$\ldots, w_{20} = \begin{cases} 1/2 + \left(\left(0.1476\right)^4/\left(3\right)^4\right) & , \\ 1 + \left(5\left(0.1476\right)^4/\left(3\right)^4\right) & , \\ \left(1/\left(3\right)*\left|0.1476\right|\right) & . \end{cases} = \begin{cases} 0.5000 & ,\left|0.1476\right| \le 3, \\ 1.0009 & ,\left|0.1476\right| = 0, . \\ 0.2582 & ,\left|0.1476\right| > 3. \end{cases}$$

The weights obtained from MW1is given in the Table 1 which also containing some other weights.

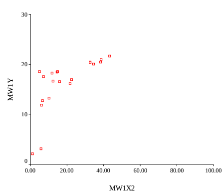
Table 1 illustrates the different weights function. It is interesting to compare the modified weight function and some M-estimators. The effects of outliers in the observation are downweight by proposed weights function. The value of reduce residuals is better than the other existing, but the other weight function does not differ from the value of weight MW1 in summary four of the downweight residuals in points 3, 11, 17 and 18 are not outliers. The value is 0.0844, 0.5088 and 0.0666 in point three, seventeen and eighteen, respectively. It is interesting to compare the modified weight function, MW1 with some others in term of MSE or R^2. The effects of outliers are downweighted by these weight functions. These 20 weights are used to transform the sample data before weighted least squares is applied to get the estimates of regression coefficients.

Table 1. Weights Obtained from the Coleman Data Set.

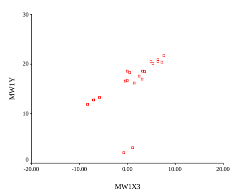
Observation	Weight Function					
	MW1	**AND**	**HAM**	**HUB**	**TUR**	**WIN**
1	**0.5000**	0.3326	0.2552	1.0000	0.0608	0.0405
2	**0.5000**	0.3326	0.2573	1.0000	0.0612	0.0408
3	**0.0844**	0.2450	0.9382	1.0000	0.6121	0.2001
4	**0.5001**	0.3320	0.4868	1.0000	0.1122	0.0748
5	**0.5008**	0.3296	0.0572	1.0000	0.3054	0.2033
6	**0.5000**	0.3333	0.2331	1.0000	0.0037	0.0024
7	**0.5005**	0.3302	0.9418	1.0000	0.2586	0.1722
8	**0.5001**	0.3322	0.4138	1.0000	0.0943	0.0629
9	**0.5003**	0.3309	0.7667	1.0000	0.1951	0.1300
10	**0.5000**	0.3331	0.0017	1.0000	0.0221	0.0148
11	**0.5489**	0.3040	0.7058	1.0000	0.6762	0.6251
12	**0.5191**	0.3148	0.8492	1.0000	0.5835	0.0167
13	**0.5025**	0.3266	0.5542	1.0000	0.5526	0.3666
14	**0.5000**	0.3326	0.2494	1.0000	0.0597	0.0398
15	**0.5206**	0.3141	0.9093	1.0000	0.6474	0.0548
16	**0.5058**	0.3230	0.0158	1.0000	0.8515	0.5611
17	**0.5088**	0.3207	0.2801	1.0000	0.0549	0.6910
18	**0.0666**	0.1990	0.8907	1.0000	0.7268	0.3392
19	**0.5033**	0.3256	0.6913	1.0000	0.6343	0.4201
20	**0.5000**	0.3329	0.0956	1.0000	0.0345	0.0230



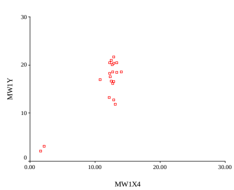
(a) Scatter Plot of $MW1Y$ vs. $MW1X1$



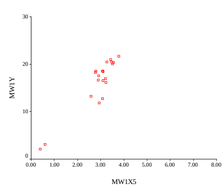
(b) Scatter Plot of $MW1Y$ vs. $MW1X2$



(c) Scatter Plot of $MW1Y$ vs. $MW1X3$



(d) Scatter Plot of $MW1Y$ vs. $MW1X4$



(e) Scatter Plot of $MW1Y$ vs. $MW1X5$

Figure 2. Scatter Plots of Transform Dependent versus Independent Variables by MW1.

From the scatter plots in the Fig 2, it is seen that the 20 observations are more clustered when MW1 is applied to the original data. The effects of outliers are downweighted by the proposed MW1.

Table 2. Estimates of Regression Coefficients to Coleman Data Set**.**

Estimates	OLS	MW1	AND	HAM	HUB	TUR	WIN
$\hat{\beta}_0^*$	19.9486	**12.9230**	22.9935	20.0533	19.9486	-2.9302	3.6342
$\hat{\beta}_1^*$	-1.7933	**-1.7650**	-1.7640	-1.7928	-1.7933	1.3206	0.6985
$\hat{\beta}_2^*$	0.0436	**-0.0346**	0.0557	0.0435	0.0436	-0.3011	-0.2340
$\hat{\beta}_3^*$	0.5558	**0.3656**	0.5869	0.5557	0.5558	0.5582	0.5595
$\hat{\beta}_4^*$	1.1102	**0.7789**	1.1277	1.1085	1.1102	-0.3568	-0.1616
$\hat{\beta}_5^*$	-1.8109	**1.1868**	-2.4846	-1.8087	-1.8109	8.4603	6.4865
MSE	**4.3027**	**1.0452**	**2.1904**	**7.9839**	**4.3027**	**15.9984**	**8.9065**
R^2	**0.8728**	**0.9028**	**0.7368**	**0.8723**	**0.8728**	**0.7366**	**0.7233**

Table 2 displays the fitted values, the mean squares error and the coefficients of determination. The estimates regression coefficients with outliers will increase estimates of intercept $\left(\hat{\beta}_0^*\right)$ and decrease estimates slope ($\hat{\beta}_1^*$ **,** $\hat{\beta}_2^*$ **,** $\hat{\beta}_3^*$**,** $\hat{\beta}_4^*$ and $\hat{\beta}_5^*$). The result of the estimates of MW are $\hat{\beta}_0^*$ = 12.9230, $\hat{\beta}_1^*$ = -1.7650, $\hat{\beta}_2^*$ = -0.0346, $\hat{\beta}_3^*$ = 0.3656, $\hat{\beta}_4^*$ = 0.7789 and $\hat{\beta}_5^*$ = 1.1868. The property of R^2 in MW1 is the highest value from all methods. It is 0.9028. This example is a good model. The result of MSE of MW1 is the lowest value from all methods. There is 1.0452, which means the proposed modified weights are performed to reduce influential observations.

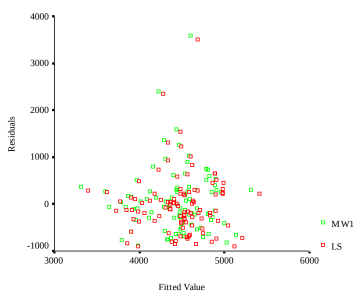


Figure 3. Plots of Residuals versus Fitted Values by Ordinary Least-Squares and MW1.

From the residual plot in Fig 3, the proposed MW1 is good performance to reduce influential observations.

7. Comparison of Estimates

Consider the model $\underline{y} = X\underline{\beta} + \underline{\varepsilon}$ where $\underline{y}$ is an $n \times 1$ vector of responses, $X = (\underline{x}_1, \underline{x}_2, ..., \underline{x}_p)'$,an $n \times p$ matrix of row vector $\underline{x}_i' = (x_{i1}, x_{i2}, ..., x_{ip})$ and x_{ij} the i^{th} value of the j^{th} regressors, $i = 1,2,...,n.$, $j = 1,2,...,p$, $\underline{\beta}$ a $p \times 1$ vector of regression coefficients and $\underline{\varepsilon}$ an $n \times 1$ vector of errors. The proposed estimated are interest in value of three regressors. The sample sizes also vary from small ($n = 10$), medium ($n = 30$) to large ($n = 50$). We can generate the residuals ε_i from a normal $N(0,1)$ distribution. Chosen $x_{i1}, x_{i2}, ..., x_{ip}$, for $i = 1,2,...,n.$ each form generated at random from $N(0,100)$ distribution. Fixed the values of $\beta_0, \beta_1, ..., \beta_p$ is one. Replace the linear model of $y_i = \underline{x}'_i \underline{\beta} + \varepsilon_i$ for $\underline{x}'_i = (1, x_{i1}, x_{i2}, ..., x_{ip})$ and $\underline{\beta} = (\beta_0, \beta_1, ..., \beta_p)'$ as chosen. For samples size of n with percentages $10\%, 20\%$ and 30% of X-outliers, Y-outliers and XY-outliers. From each sample so obtained, find the estimates of regression coefficients, $R^{2,}$ and MSE using MW, AND, HAM, HUB, TUR and WIN. The averages of R^2 and MSE form the 5,000 samples are then computed and compared. The results for three regressors appear in Tables 3-8 and Figures 4-9.

Table 3. Coefficients of Determination for Different Estimates by Sample Sizes and Percentages of X-Outliers.

Sample Sizes	% of X-Outliers	OLS	MW1	AND	HAM	HUB	TUR	WIN
		Coefficients of Determination						
10	10	0.9869	**0.9874**	0.9862	0.8609	0.9864	0.9290	0.9871
	20	0.9985	**0.9986**	0.9980	0.9668	0.9982	0.9899	0.9985
	30	0.9996	**0.9997**	0.9996	0.9982	0.9995	0.9990	0.9996
30	10	0.9986	**0.9988**	0.9985	0.9965	0.9984	0.9980	0.9986
	20	0.9997	**0.9997**	0.9997	0.9997	0.9996	0.9997	0.9997
	30	0.9998	**0.9998**	0.9998	0.9998	0.9998	0.9998	0.9998
50	10	0.9993	**0.9993**	0.9992	0.9990	0.9992	0.9993	0.9993
	20	0.9997	**0.9997**	0.9997	0.9997	0.9997	0.9998	0.9997
	30	0.9998	**0.9998**	0.9998	0.9998	0.9998	0.9999	0.9998

With X-outliers, Table 3 shows that for all samples. MW1 estimates are the best in terms of R^2 values. The performance of MW1 has the highest value of R^2 **(0.9998)** in other sample sizes and percentages of outliers. The next best estimates are HUB and WIN in term of R^2 with high percentages of X-outliers [Fig. 4].

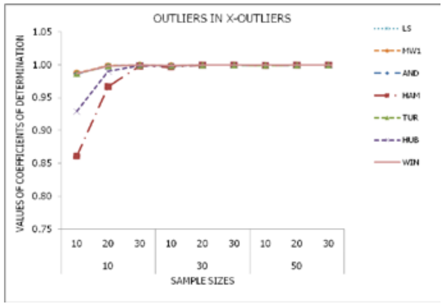


Figure 4. Graph of Coefficients of Determination for Different Estimates by Sample Sizes and Percentages of X-Outliers.

From the graph of coefficients of determination in Fig 4, the proposed MW1 has the highest value of R^2 in other sample sizes and percentages of X-outliers.

Table 4. Coefficients of Determination for Different Estimates by Sample Sizes and Percentages of Y-Outliers.

Sample Sizes	% of Y-Outliers	Coefficients of Determination						
		OLS	MW1	AND	HAM	HUB	TUR	WIN
10	10	0.9800	**0.9789**	0.9809	0.8960	0.9778	0.9430	0.9801
	20	0.9787	**0.9073**	0.9072	0.8865	0.9074	0.8928	0.9069
	30	0.8399	**0.8947**	0.8426	0.8789	0.8547	0.8378	0.8402
30	10	0.8386	**0.8831**	0.8390	0.8340	0.8390	0.8385	0.8385
	20	0.7210	**0.7224**	0.7083	0.7103	0.7054	0.7170	0.7065
	30	0.7093	**0.8248**	0.7154	0.7110	0.7102	0.7283	0.7111
50	10	0.7594	**0.8465**	0.7616	0.7571	0.7593	0.7637	0.7599
	20	0.6913	**0.7060**	0.6918	0.6925	0.6888	0.7010	0.6904
	30	0.6913	**0.8286**	0.6977	0.6970	0.6908	0.7187	0.6927

With Y-outliers, Table 4 shows that for all sample sizes. MW1 estimates also give higher value of R^2 when the percentages of Y-outliers and sample sizes are small [Fig. 5].

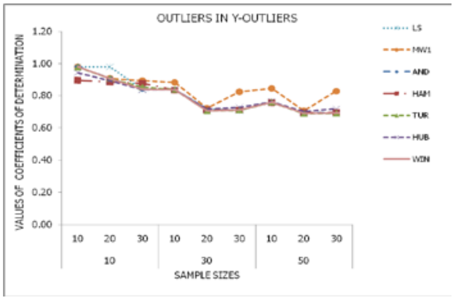


Figure 5. Graph of Coefficients of Determination for Different Estimates by Sample Sizes and Percentages of Y-Outliers.

From the graph of coefficients of determination in Fig 5, the proposed MW1 has the highest value of R^2 in other sample sizes and percentages of Y-outliers.

Table 5. Coefficients of Determination for Different Estimates by Sample Sizes and Percentages of XY-Outliers.

Sample Sizes	% of XY-Outliers	OLS	MW1	AND	HAM	HUB	TUR	WIN
		Coefficients of Determination						
10	10	0.2512	**0.3074**	0.2604	0.2355	0.2514	0.2480	0.2535
	20	0.2246	**0.2797**	0.2320	0.2117	0.2250	0.2196	0.2252
	30	0.2209	**0.2803**	0.2314	0.2191	0.2205	0.2312	0.2237
30	10	0.0746	**0.2229**	0.0771	0.0668	0.0743	0.0777	0.0748
	20	0.0710	**0.1879**	0.0724	0.0648	0.0708	0.0727	0.0708
	30	0.0695	**0.1926**	0.0717	0.0660	0.0690	0.0747	0.0694
50	10	0.0452	**0.2432**	0.0462	0.0402	0.0450	0.0490	0.0450
	20	0.0431	**0.2312**	0.0440	0.0388	0.0429	0.0453	0.0429
	30	0.0428	**0.2166**	0.0439	0.0395	0.0426	0.0440	0.0429

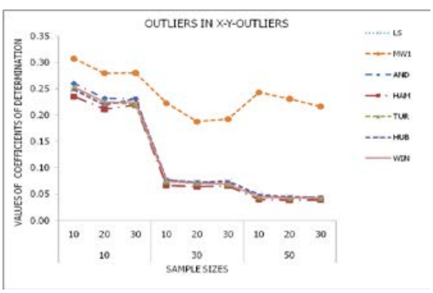


Figure 6. Graph of Coefficients of Determination for Different Estimates by Sample Sizes and Percentages of XY-Outliers.

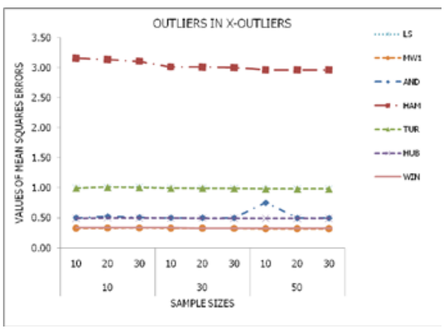
With XY-outliers, Table 5 shows that for all samples. MW1 estimates are the best in terms of R^2 values. The performance of MW1 has the highest value of R^2 (0.9824) for all sample sizes and percentages of outliers with medium and large sizes, [Fig. 6]. From the graph of coefficients of determination in Fig 5, the proposed MW1 has the highest value of R^2 in other sample sizes and percentages of XY-outliers.

Table 6. Mean Squares of Error for Different Estimates by Sample Sizes and Percentages of X-Outliers.

Sample Sizes	% of X-Outliers	Mean Squares Error						
		OLS	MW1	AND	HAM	HUB	TUR	WIN
10	10	1.0001	**0.3279**	0.5032	3.1601	0.9983	0.4933	0.3336
	20	1.0132	**0.3318**	0.5234	3.1390	1.0117	0.5034	0.3379
	30	1.0147	**0.3345**	0.5099	3.1088	1.0074	0.4988	0.3393
30	10	0.9979	**0.3272**	0.5028	3.0162	0.9966	0.4971	0.3330
	20	0.9968	**0.3269**	0.5026	3.0105	0.9957	0.4959	0.3327
	30	0.9932	**0.3257**	0.5006	3.0068	0.9921	0.4941	0.3315
50	10	0.9870	**0.3233**	0.7555	2.9639	0.9865	0.4927	0.3292
	20	0.9876	**0.3235**	0.4975	2.9641	0.9871	0.4930	0.3294
	30	0.9864	**0.3231**	0.4969	2.9641	0.9859	0.4924	0.3290

With X-outliers, Table 6 shows that for all sample sizes. MW1 estimates give the lowest of MSE (0.3231) for all sample sizes and percentages of X-outliers. The next best estimates are AND and WIN in term of MSE with percentages of X-outliers [Fig. 7].

Figure 7. Graph of Mean Squares of Error for Different Estimates by Sample Sizes and Percentages of X-Outliers.



From the graph of mean squares of error in Fig 7, the proposed MW1 has the highest value of MSE in other sample sizes and percentages of X-outliers.

Table 7 Mean Squares of Error for Different Estimates by Sample Sizes and Percentages of Y-Outliers.

Sample Sizes	% of Y-Outliers	Mean Squares Error						
		OLS	MW1	AND	HAM	HUB	TUR	WIN
10	10	1.4787	**0.8475**	1.0483	2.5388	1.4764	1.0376	0.8546
	20	4.0596	**1.9263**	2.4201	5.6291	3.3061	2.3187	1.9311
	30	28.7810	**17.1638**	19.5727	63.0058	24.7333	19.1582	16.5928
30	10	15.2936	**8.8489**	12.6273	31.5189	15.0549	10.6970	8.8595
	20	30.0146	**14.8782**	18.5803	47.0704	25.8850	18.2231	14.9845
	30	48.0682	**27.8164**	36.2041	112.8139	46.6765	33.5606	27.7576
50	10	22.3703	**12.8859**	16.6498	45.3387	22.1525	15.7284	12.9419
	20	34.0150	**16.7428**	21.0546	52.6560	29.1862	20.6266	16.8752
	30	50.4750	**29.0860**	41.8833	111.4511	49.6787	35.4542	29.1553

With Y-outliers, Table 7 shows that for all sample sizes. MW1 estimates give the lowest MSE (0.8475) for all sample sizes and percentages of Y-outliers. The next best estimates are AND and WIN in term of MSE with percentages of Y-outliers [Fig. 8].

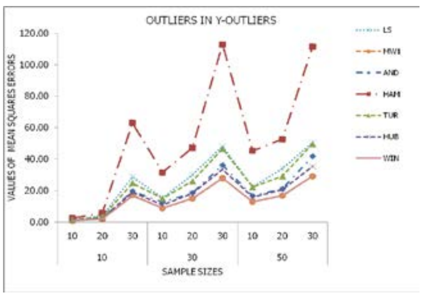


Figure 8. Graph of Mean Squares of Error for Different Estimates by Sample Sizes and Percentages of Y-Outliers.

From the graph of mean squares of error in Fig 8, the proposed MW1 has the highest value of MSE in other sample sizes and percentages of Y-outliers.

With XY-outliers, Table 8 shows MW1 estimates give the lowest MSE (0.8477) for all sample sizes and percentages of XY-outliers. The next best estimates are AND and WIN in term of MSE with percentages of XY-Outliers [Fig. 9].

Table 8. Mean Squares of Error for Different Estimates by Sample Sizes and Percentages of XY-Outliers.

Sample Sizes	% of XY-Outliers	Mean Squares Error						
		OLS	MW1	AND	HAM	HUB	TUR	WIN
10	10	25.2228	**14.6314**	17.8273	48.5600	24.9207	17.4544	14.6389
	20	39.1387	**22.6894**	27.6909	73.4637	38.6834	27.1021	22.7111
	30	49.5984	**28.7021**	35.1211	92.3356	49.0558	34.4156	28.7358
30	10	28.4817	**16.4049**	20.2929	52.7601	28.3542	19.9945	16.4841
	20	42.5516	**24.4894**	30.4059	77.7482	42.3836	29.9101	24.6149
	30	53.0639	**30.5290**	37.8719	96.6856	52.8583	37.3218	30.6869
50	10	29.6836	**17.0488**	21.1494	54.2209	29.6117	20.9249	17.1590
	20	42.8752	**24.6304**	31.0211	78.1280	42.7571	30.2160	24.7782
	30	53.3872	**30.6606**	38.1572	96.4851	53.2451	37.6390	30.8478

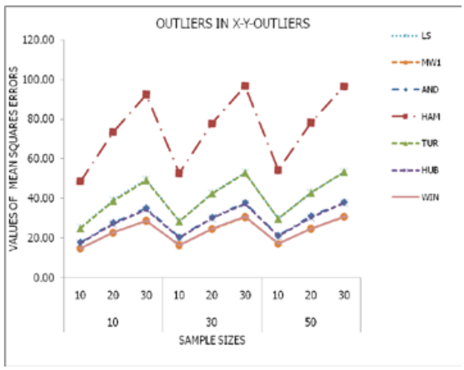


Figure 9. Graph of Mean Squares of Error for Different Estimates by Sample Sizes and Percentages of XY-Outliers.

From the graph of mean squares of error in Fig 9, the proposed MW1 has the highest value of MSE in other sample sizes and percentages of XY-outliers.

8. Conclusions

An influence function or a $\rho - function$ is given to constructs proper weights for residual in multiple linear regression. The new weights are used in the estimation of regression coefficients with outliers. The weights so constructed downweight the outliers observations. Study this explains the process of derivation of weights for the proposed influence function and shows better results by simulation in terms of the coefficient of determination and mean squares error comparing to the ordinary least squares and the other M-estimators. However, MW1 perform better than other, which is show the best criterion in the fit model

Acknowledgements

I sincerely acknowledge the Suvarnaphumi Rajamangala University of Technology who supported my scholarship. I would like to give many thanks to Miss Supaporn Tunpornpituk for kindly provide the SAS program. The remaining errors belong to the first author.

References

[1] Thomas, P. R. Modern Regression Methods. New York: John Wiley & Sons, 1997

[2] Draper, N.R. Applied Regression Analysis. 3rd ed, New York: John Wiley & Sons, 1998.

[3] Montgomery, D.C., Peck E.A. and Vining, G.G. Introduction to Linear Regression Analysis. 3rd ed. New York: John Wiley & Sons, 2003.

[4] Rousseeuw, P.J. and Leroy, A.M. Robust Regression and Outlier Detection. New York: John Wiley & Sons, 1987.

[5] Yohai, V.J. High Breakdown Point and High Efficiency Robust Estimates for Regression. *Annals of Statistics,* 1987; 15: 642-656.

[6] Hampel, F.R. The Influence Curve and Its Role in Robust Estimation. *Journal of the American Statistical Association,* 1974; 69: 383-393.

[7] Krasker, W.S. and Welsch, R.E. Efficientv Bounded-Influence Regression Estimation. *Journal of the American Statistical Association,* 1982; 77: 595-604.

[8] Simpson, D. G., Ruppert D. and Carroll R. J. On one-step GM-estimates and Stability of Inference in Linear Models. *Journal of the American Statistical Association,* 1992; 87: 439-450.

[9] Maronna, R. A., Martin R.D. and Yohai, V.J. Robust Statistics: Theory and Methods. England: John Wiley & Sons, 2006.

[10] Huber, P. J. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics,* 1964; 35: 73-101.

[11] Tukey, J. W. Explortary Data Analysis, Addison-Wesley, Reading, MA. *The Annals of Mathematical Statistics,* 1977; 31: 385-391

[12] Ampanthong, P., and Suwattee, P. A Comparative Study of Outlier Detection Procedures in Multiple Linear Regression. Proceedings of Conference by IAENG

International Conference on Data Mining and Applications (ICDMA'09), Kowloon, HongKong, 18-20 March, 2009: 704-709.

[13] Ampanthong P. and Suwattee P. A Monte Carlo Simulation Comparison of Outlier Detection Procedures in Regression. Proceedings of Conference 14th Annual Meeting in Mathematics (AMM'09), Suratthani, Thailand. 5-6 March, 2009: 131-149.

[14] Ampanthong P. and Suwattee P. Estimation of Regression Coefficients with Outliers. Proceedings of Conference by NIDA International (NIDA'09), Chonburi, Thailand, 21-22 May, 2009: 561-576.

[15] Ampanthong P. and Suwattee P. Construction of Weights for the Estimation of Regression Coefficients with Outliers. Proceedings of Conference 5th IMT-GT 2009 by International Conference on Mathematics, Statistics, and their Applications (ICMSA'09), Bukittinggi, Indonesia, 9-11 June, 2009: 198-203.

[16] Andrews, D.F. A Robust Method for Multiple Linear Regression. Technometrics, 1974; 16: 523-531.

[17] Coakley, C.W. and Hettamansperger, T.P. A Bounded Influence, High Breakdown, Efficient Regression Estimator. *Journal of the American Statistical Association,* 1993; 88: 872-880.

[18] Dixon, W.J. Simplified Estimation from Censored Normal Samples. The Annals of Mathematical Statistics, 1960; 31: 385-391.

[19] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel W.A. Robust Statistics: The Approach Based on Influence Functions. New York: John Wiley & Sons, 1986.

[20] Huber, P. J. The Wald Lecture Robust Statistics: A Review, *The Annals of Mathematical Statistics.* 1972; 43: 1041-1067.

[21] Huber, P.J. Robust Regression: Asymptotics, Conjectures and Monte Carlo. *Annals of Statistics,* 1973; 1: 799-821.

[22] Maronna, R.A. and Yohai, V.J. The Breakdown Point of Simultaneous General M Estimates of Regression and Scale. *Journal of the American Statistical Association,* 1991; 86: 699-703.

[23] Rousseeuw, P.J. Least Median of Squares Regression. *Journal of the American Statistical Association,* 1984; 79: 871-880.

[24] Rousseeuw, P.J. and Zomeren, B.C.V. Unmasking Multivariate Outliers and Leverage Points. Journal of the American Statistical Association, 1990; 85: 633-639