A Least Deviation Estimation Approach for Time Series Models

Silvey Shamsi[1], Mian Arif Shams Adnan[2]
[1]Ball State University, Muncie, Indiana, 47304
[2] Bowling Green State University, Bowling Green, Ohio, 43402

Abstract
Estimates of the Time Series Models suffer from the exaggeration of the contribution of the extreme observation(s) as well as the outlier(s). Unlike the traditional fitting procedure of Time Series Models based on the least squares estimates, it uses least deviation method for minimizing the total sum of errors. It also suggests some new measures of fit.

Key Words: Absolute Deviation, Outlier, Relative Coefficient of Determination.

1. Introduction

A time series is a series of data points indexed (or listed or graphed) in time order (1977, 1986, 1987, and 2008). Most commonly, a time series is a sequence taken at successive equally spaced points in time. Thus it is a sequence of discrete-time data. Examples of time series are heights of ocean tides, counts of sunspots, and the daily closing value of the Dow Jones Industrial Average.

Existing regression including Simple Linear Regression, Multiple Regression, etc use Least Square Method for estimating regression parameters. Unfortunately, the least square estimates of regression parameters leave the presence of extreme observations and/or outliers exaggerated that mislead a researcher or analyst with significant (or insignificant) value of the parameters with insignificant (or significant) effects. So, one dimensional distances should be used instead of squared distances for estimating regression parameters. But, the one-dimensional distances of the data from the fitted regression line makes the total sum of errors zero which does not help the mathematicians to differentiate with respect to the parameters to calculate least deviation estimate of regression parameters. This is due to the fact that sum of positive deviations (positive errors) of the dependent variable apart from the fitted regression line nullifies the negative deviations (negative errors). As a result, statisticians used least square deviations not only to make the deviations apart from the fitted regression line positive but also to make the sum of squares of errors differentiable with respect to parameters so that a class of normal equations are accessible that result least square estimates. So, there was no way of using the one-dimensional naïve difference between observed values of the dependent variable and its expected or fitted values.

Fortunately, one dimensional transformed differences of the aforesaid values might be used for the sake of having the regression estimates free from exaggeration by the presence extreme observation and/or outlier(s). For estimating the regression parameters, if we retransform the normal equations for fitting the regression line, we should get a fitted regression line along with least deviation regression estimates that overcome the problem for the presence or extreme as well as outlier(s). Shamsi and Adnan (2016, 2017) proposed Least Deviation Approach for Simple and Multiple Linear Regression. We have observed that Least Deviation Estimators performs better than those by Ordinary Least Square Estimators for simple linear regression.

Attempt has been made here to find a proper transformation of the one-dimensional concern difference so that we can smoothly estimate the multiple regression parameters. Using the proper transformation of the one-dimensional distance from the fitted regression line we have estimated regression parameters and checked whether the estimators follow the BLUE properties. The performance and the limiting behavior of the parameters have also been observed under simulations.

2. Estimation Methods

Shamsi *et al* (2016) suggested that for the simple linear regression model

$$y = \beta_0 + \beta_1 x + \varepsilon$$

where the intercept β_0 and the slope β_1 are unknown constant known as regression coefficients and ε is a random error component, the Least Deviation Estimators are

$$\hat{\beta}_1 = \frac{\overline{\left(\frac{y}{x}\right)} - \bar{y}\,\overline{\left(\frac{1}{x}\right)}}{\left[1 - \bar{x}\,\overline{\left(\frac{1}{x}\right)}\right]}$$

and

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$

It was also observed that the relative fit to the error as

$$FBR = \frac{Fit}{Error} = \frac{\sum|\hat{y}_{Method} - \bar{y}|}{\sum|y - \hat{y}_{Method}|}$$

is less for Ordinary Least Deviation Method compared to Ordinary Least Square Method for the above case. That is

$$FBR_{OLD} = \frac{Fit_{OLD}}{Error_{OLD}} = \frac{\sum|\hat{y}_{OLD} - \bar{y}|}{\sum|y - \hat{y}_{OLD}|} > FBR_{OLS} = \frac{Fit_{OLS}}{Error_{OLS}} = \frac{\sum|\hat{y}_{OLS} - \bar{y}|}{\sum|y - \hat{y}_{OLS}|}.$$

We want to study the Least Deviation regression estimators and their performance for multiple regression.

Let the simple linear regression model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \tag{2.1}$$

where the intercept β_0 and the slopes $\beta_1, \beta_2$are unknown constant known as regression coefficients and ε is a random error component. The errors are assumed to have mean zero and unknown variance σ^2. Here the errors are uncorrelated. There is a Normal probability distribution for y at each possible value for x such that

$$E(y|x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

and
$$V(y|x_i) = V(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon) = \sigma^2.$$

Although the mean of y is a linear function of x that is the conditional mean of y depends on all x, but the conditional variance of y does not depend on any x. Responses y are uncorrelated since the errors ε are uncorrelated. Moreover, the independent variables are mutually independent.

Since the parameters β_i are unknown, they should be estimated using sample data. Suppose that we have $n$tuples of data, say $(y_1, x_{11}, x_{21}), (y_2, x_{12}, x_{22}), \dots, (y_n, x_{1n}, x_{2n})$ obtained from a controlled experimental design or from an observational study or from existing historical records. Least Square method estimates β_i so that the sum of squares of differences between the observations y_i and the straight line is minimum. From equation 2.1 we can write
$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i ; i = 1, 2, \dots, n \qquad (2.2)$$

Equation 2.1 presents the Population Multiple Regression Model and equation 2.2 expresses the Sample Multiple Regression Model. Now the sum of squares of deviations from the true line is
$$S = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i})^2. \qquad (2.3)$$

Now the least square estimates of β_i must satisfy

$$\frac{\partial S}{\partial \beta_0} = -2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}) = 0, \qquad (2.4)$$

$$\frac{\partial S}{\partial \beta_1} = -2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}) x_{1i} = 0, \qquad (2.5)$$

$$\frac{\partial S}{\partial \beta_2} = -2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}) x_{2i} = 0. \qquad (2.6)$$

After simplification the normal equations are generally found such that

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{1i} + \hat{\beta}_2 \sum_{i=1}^{n} x_{2i} = \sum_{i=1}^{n} y_i \qquad (2.7)$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_{1i} + \hat{\beta}_1 \sum_{i=1}^{n} x_{1i}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{1i} x_{2i} = \sum_{i=1}^{n} x_{1i} y_i \qquad (2.8)$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_{2i} + \hat{\beta}_1 \sum_{i=1}^{n} x_{2i} x_{1i} + \hat{\beta}_2 \sum_{i=1}^{n} x_{2i}^2 = \sum_{i=1}^{n} x_{2i} y_i \qquad (2.9)$$

The solution to the normal equations is

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2, \qquad (2.10)$$

$$\hat{\beta}_1 = \frac{\left(\sum_{i=1}^{n} y_i x_{1i} - n\bar{y}\bar{x}_1\right)\left(\sum_{i=1}^{n} {x_2}^2 - n{\bar{x}_2}^2\right) - \left(\sum_{i=1}^{n} y_i x_{2i} - n\bar{y}\bar{x}_2\right)\left(\sum_{i=1}^{n} x_1 x_2 - n\bar{x}_1\bar{x}_2\right)}{\left(\sum_{i=1}^{n} x_i^2 - n{\bar{x}_1}^2\right)\left(\sum_{i=1}^{n} x_2^2 - n{\bar{x}_2}^2\right) - \left(\sum_{i=1}^{n} x_1 x_2 - n\bar{x}_1\bar{x}_2\right)^2}, \qquad (2.11)$$

and
$$\hat{\beta}_2 = \frac{\left(\sum_{i=1}^{n} y_i x_{2i} - n\bar{y}\bar{x}_2\right)\left(\sum_{i=1}^{n} {x_1}^2 - n{\bar{x}_1}^2\right) - \left(\sum_{i=1}^{n} y_i x_{1i} - n\bar{y}\bar{x}_1\right)\left(\sum_{i=1}^{n} x_1 x_2 - n\bar{x}_1\bar{x}_2\right)}{\left(\sum_{i=1}^{n} x_i^2 - n{\bar{x}_1}^2\right)\left(\sum_{i=1}^{n} x_2^2 - n{\bar{x}_2}^2\right) - \left(\sum_{i=1}^{n} x_1 x_2 - n\bar{x}_1\bar{x}_2\right)^2} \qquad (2.12)$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \bar{x}_1 = \frac{1}{n}\sum_{i=1}^{n} x_{1i}$ and $\bar{x}_2 = \frac{1}{n}\sum_{i=1}^{n} x_{2i}$

Therefore, $\hat{\beta}_0$ and $\hat{\beta}_1$, $\hat{\beta}_2$ are the Least Square estimates of the intercept and slopes respectively. The fitted Simple Linear Regression Model is

$$\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_1 x_2. \tag{2.13}$$

For the aforesaid regression model the sum of absolute deviations from the true line is

$$\sum_{i=1}^{n} |\varepsilon_i| = \sum_{i=1}^{n} |y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}|. \tag{2.14}$$

Now the least deviation estimates of β_0, β_1 and β_2 will satisfy

$$\frac{\partial \sum_{i=1}^{n} |\varepsilon_i|}{\partial \beta_0} = 0, \tag{2.15}$$

$$\frac{\partial \sum_{i=1}^{n} |\varepsilon_i|}{\partial \beta_1} = 0, \tag{2.16}$$

and

$$\frac{\partial \sum_{i=1}^{n} |\varepsilon_i|}{\partial \beta_2} = 0 \tag{2.17}$$

which are equivalent to

$$\frac{\partial \sum_{i=1}^{n} ln|\varepsilon_i|}{\partial \beta_0} = -\infty \tag{2.18}$$

$$\frac{\partial \sum_{i=1}^{n} ln|\varepsilon_i|}{\partial \beta_1} = -\infty, \tag{2.19}$$

and

$$\frac{\partial \sum_{i=1}^{n} ln|\varepsilon_i|}{\partial \beta_2} = -\infty. \tag{2.20}$$

Therefore,

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2 \tag{2.21}$$

Moreover,

$$\hat{\beta}_1 = \overline{\left(\frac{y}{x_1}\right)} - \hat{\beta}_0 \overline{\left(\frac{1}{x_1}\right)} - \hat{\beta}_2 \overline{\left(\frac{x_2}{x_1}\right)} \tag{2.22}$$

$$\therefore \hat{\beta}_1 \left[1 - \bar{x}_1 \overline{\left(\frac{1}{x_1}\right)}\right] - \left[\bar{x}_2 \overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right] \hat{\beta}_2 - \left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y} \overline{\left(\frac{1}{x_1}\right)}\right] = 0 \tag{2.23}$$

$$\therefore \hat{\beta}_1 = \frac{\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right] + \left[\bar{x}_2 \overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\hat{\beta}_2}{\left[1 - \bar{x}_1 \overline{\left(\frac{1}{x_1}\right)}\right]} \tag{2.24}$$

Again,

$$\frac{\partial \sum_{i=1}^{n} ln|\varepsilon_i|}{\partial \beta_2} = \frac{\partial \sum_{i=1}^{n} ln|y_i - \beta_0 - \beta_1 x_{1i} - \beta_2 x_{2i}|}{\partial \beta_2} = -\infty$$

$$\therefore \hat{\beta}_2 = \overline{\left(\frac{y}{x_2}\right)} - \hat{\beta}_0 \overline{\left(\frac{1}{x_2}\right)} - \hat{\beta}_1 \overline{\left(\frac{x_1}{x_2}\right)} \tag{2.25}$$

$$\therefore \hat{\beta}_1 \left[\bar{x}_1 \overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right] - \hat{\beta}_2 \left[1 - \bar{x}_2 \overline{\left(\frac{1}{x_2}\right)}\right] + \left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y} \overline{\left(\frac{1}{x_2}\right)}\right] = 0, \tag{2.26}$$

$$\therefore\ \hat{\beta}_2 = \frac{\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right] + \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\hat{\beta}_1}{\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right]}, \tag{2.27}$$

Therefore, from (2.22) and (2.25) we get the following equations of $\hat{\beta}_1$, $\hat{\beta}_2$ such that

$$\hat{\beta}_1\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right] - \hat{\beta}_2\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right] - \left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right] = 0$$

$$\hat{\beta}_1\left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right] - \hat{\beta}_2\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] + \left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right] = 0.$$

After the cross multiplication we get the following equations

$$\frac{\hat{\beta}_1}{-\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right] - \left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right]}$$

$$= \frac{\hat{\beta}_2}{-\left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right] - \left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right]}$$

$$= \frac{1}{-\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] + \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]}$$

Therefore,

$$\therefore\ \hat{\beta}_1 = \frac{\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right] + \left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right]}{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]} \tag{2.28}$$

$$\therefore\ \hat{\beta}_2 = \frac{\left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right] + \left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right]}{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]} \tag{2.29}$$

Unlike the least square estimates, least deviations estimates may follow other properties, which might not be BLUE.

Theorem 2.1: *If for the multiple linear regression model* $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$, β_0 *and* $\beta_{1,}\ \beta_2$ *are the unknown the intercept and the slope constants known as regression coefficients and* ε *is a random error component, and if we have* n *pairs of data, say* $(y_1, x_{11}, x_{21}), (y_2, x_{12}, x_{22}), \ldots, (y_n, x_{1n}, x_{2n})$ *, then R be the percentage of observations that explains the extent fit of the model such that*

$$0 < \frac{\sum_{i=1}^{n}|\hat{y} - \bar{y}|}{\sum_{i=1}^{n}|y_i - \bar{y}|} = R < 1. \tag{2.30}$$

Here R presents the percentage of observations being explained by the fitted model.

Theorem 2.2: *If for the simple linear regression model* $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$, β_0 *and* β_i *are the unknown the intercept and the slope constants known as regression coefficients and* ε *is a random error component, and if we have* n *pairs of data, say* $(y_1, x_{11}, x_{21}), (y_2, x_{12}, x_{22}), \ldots, (y_n, x_{1n}, x_{2n})$, *obtained from a controlled experimental* *design or from an observational study or from existing historical records, then the estimators of β_1 and β_2 will be*

$$\hat{\beta}_1 = \frac{\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right] + \left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right]}{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]}$$

$$\hat{\beta}_2 = \frac{\left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right] + \left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right]}{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]}$$

such that $\hat{\beta}_1$, $\hat{\beta}_2$ are the unbiased estimators i.e.

$$E(\hat{\beta}_1) = \beta_1, E(\hat{\beta}_2) = \beta_2, \tag{2.31}$$

Theorem 2.3: *If for the simple linear regression model $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i$, β_0 and β_i are the unknown the intercept and the slope constants known as regression coefficients and ε is a random error component, and if we have n pairs of data, say $(y_1, x_{11}, x_{21}), (y_2, x_{12}, x_{22}), \ldots, (y_n, x_{1n}, x_{2n})$, then the estimator of β_1* and *β_2 will be*

$$\hat{\beta}_1 = \frac{\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right] + \left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right]}{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]}$$

$$\hat{\beta}_2 = \frac{\left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\overline{\left(\frac{y}{x_1}\right)} - \bar{y}\overline{\left(\frac{1}{x_1}\right)}\right] + \left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[\overline{\left(\frac{y}{x_2}\right)} - \bar{y}\overline{\left(\frac{1}{x_2}\right)}\right]}{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]}$$

such that the variance of the estimator $\hat{\beta}_1$ will be

$$V(\hat{\beta}_1) = \frac{\sigma^2}{n}\frac{\left\{\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\right\}^2\left[\overline{\left(\frac{1}{x_2^2}\right)} + \overline{\left(\frac{1}{x_2}\right)}^2\right] + \left\{\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right]\right\}^2\left[\overline{\left(\frac{1}{x_1^2}\right)} + \overline{\left(\frac{1}{x_1}\right)}^2\right]}{\left\{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\right\}^2}$$

$$V(\hat{\beta}_2) = \frac{\sigma^2}{n}\frac{\left\{\left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\right\}^2\left[\overline{\left(\frac{1}{x_1^2}\right)} + \overline{\left(\frac{1}{x_1}\right)}^2\right] + \left\{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\right\}^2\left[\overline{\left(\frac{1}{x_2^2}\right)} + \overline{\left(\frac{1}{x_2}\right)}^2\right]}{\left\{\left[1 - \bar{x}_1\overline{\left(\frac{1}{x_1}\right)}\right]\left[1 - \bar{x}_2\overline{\left(\frac{1}{x_2}\right)}\right] - \left[\bar{x}_1\overline{\left(\frac{1}{x_2}\right)} - \overline{\left(\frac{x_1}{x_2}\right)}\right]\left[\bar{x}_2\overline{\left(\frac{1}{x_1}\right)} - \overline{\left(\frac{x_2}{x_1}\right)}\right]\right\}^2}$$

Some real-life example along with data will be cited to explain the credibility of the Least Deviation Estimates over Least Square Estimates. Multiple simulations will be backed to simplify and amplify the limiting behaviors of the Least Deviation Estimators and Estimates.

2.1 Delivery time Data: Two independent variables
A set of 25 paired observations of average number of delivery time, y, and the number of cases, x_1, and the distances, x_1, are available in page 74 of the same text. From the scatter plot we observe that the 3-tupled observations form a 3D form.

Two linear regressions lines having following equations have been fitted according to the Least Square Method and Least Deviation Method respectively as

$$\hat{y}_{OLS} = 2.34 - 1.62x_1 + 0.01x_2,$$

$$\hat{y}_{OLD} = 3.94 + 1.68x_1 + 0.01x_2.$$

We also observe the following dispersion measures.

$$\sum(y-\bar{y})^2 = \sum(y-\hat{y}_{OLS})^2 + \sum(\hat{y}_{OLS}-\bar{y})^2 < \sum(y-\hat{y}_{OLD})^2 + \sum(\hat{y}_{OLD}-\bar{y})^2$$

$$5785 = 234 + 5551 > 280 + 4774$$

$$\sum|y-\bar{y}| < \sum|y-\hat{y}_{OLS}| + \sum|\hat{y}_{OLS}-\bar{y}| < \sum|y-\hat{y}_{OLD}| + \sum|\hat{y}_{OLD}-\bar{y}|$$

$$251 < 57 + 257 > 58 + 235$$

$$\sum(y-\hat{y}_{OLS})^2 = 234 < \sum(y-\hat{y}_{OLD})^2 = 280$$
and
$$\sum|y-\hat{y}_{OLS}| = 150 < \sum|y-\hat{y}_{OLD}| = 223$$

$$R^2{}_{OLS} = \frac{\sum(\hat{y}_{OLS}-\bar{y})^2}{\sum(y-\bar{y})^2} = \frac{5551}{5785} = 0.96 > R^2{}_{OLD} = \frac{\sum(\hat{y}_{OLD}-\bar{y})^2}{\sum(y-\bar{y})^2} = \frac{4774}{5785} = 0.83$$

$$R_{OLS} = \frac{\sum|\hat{y}_{OLS}-\bar{y}|}{\sum|y-\bar{y}|} = \frac{257}{251} = 1.02 > R_{OLD} = \frac{\sum|\hat{y}_{OLD}-\bar{y}|}{\sum|y-\bar{y}|} = \frac{235}{251} = 0.94$$

$\sum e^2$ is less (234<280) for OLS but $\sum|e|$ is not much more (150<223) for OLD method. Moreover, total variation of the unidimensional data (y), being explained by two dimensional co-efficient of determination, is greater (0.96>0.83) for OLS fit, and also being explained by one dimensional co-efficient of determination, is also greater (1.02>0.94) for OLS fit. One dimensional scaled dispersion conducts little bit more error and commit less good quality of estimation for OLD method compared to OLS method. So, the Ordinary Least Dispersion Method does not worth better estimation compared to Ordinary Least Square Method for multiple regression.

2.2 Comparison of the Least Deviation Estimates with the Least Square Estimates
Performances and potentialities will be better for the Least Deviations Estimates compared to those of Least Square Method due to the reason is that the estimators of the regression coefficients are maintaining the similar dimension of the true data since they are expressed in terms of the mathematical operations with unidimensional scaling. Besides, for the presence of extreme values or outliers, the two-dimensional scaling of the dispersions of the fitted line from the data are being exaggerated. So, the malicious influences of the

extreme value(s) and/or outlier(s) drastically affects the overall estimation of parameters in regression based on Least Square Method. On the other hand, the absolute deviations for the dispersion of the observations apart from fitted regression line in regression estimators are apathetic to the extreme observation(s) and/or outlier(s) and equi-sensitive to all observations.

Moreover, error and quality estimation are two complementary factors. If error increases and quality estimation decreases or vice versa, then it is better to assess the relative performance like estimation divided by error. If for one-dimensional data, one-dimensional relative variation due to regression with respect to total error encountered for fitting the model by one method is greater than that of the other method, then it is better to use the first method for fitting regression line. Interestingly enough it is observed that the relative fit to the error as

$$FBR = \frac{Fit}{Error} = \frac{\sum|\hat{y}_{Method} - \bar{y}|}{\sum|y - \hat{y}_{Method}|} \qquad (2.32)$$

is less for Ordinary Least Deviation Method compared to Ordinary Least Square Method for the above case. That is

$$FBR_{OLD} = \frac{Fit_{OLD}}{Error_{OLD}} = \frac{\sum|\hat{y}_{OLD} - \bar{y}|}{\sum|y - \hat{y}_{OLD}|} > FBR_{OLS} = \frac{Fit_{OLS}}{Error_{OLS}} = \frac{\sum|\hat{y}_{OLS} - \bar{y}|}{\sum|y - \hat{y}_{OLS}|}. \qquad (2.33)$$

Here, one-dimensional FBR for OLS is greater than one-dimensional FBR for OLD (4.50>4.04). The two dimensional co-efficient of determination can be described as below

$$FBR = \frac{Fit}{Error} = \frac{\sum(\hat{y}_{Method} - \bar{y})^2}{\sum(y - \hat{y}_{Method})^2} \quad . \qquad (2.34)$$

Two-dimensional FBR for OLS is higher than two-dimensional FBR for OLD (23.74>17.08). In example, two dimensional relative co-efficient of determination for OLS is greater than that for OLD, because in the current multiple regression there are two independent variables and the regression model is a plane rather that a line in 3D space. Here we also observe the following inequalities.

$$FBR_{OLS} = \frac{Fit_{OLS}}{Error_{OLS}} = \frac{\sum(\hat{y}_{OLS} - \bar{y})^2}{\sum(y - \hat{y}_{OLS})^2} > FBR_{OLD} = \frac{Fit_{OLD}}{Error_{OLD}} = \frac{\sum(\hat{y}_{OLD} - \bar{y})^2}{\sum(y - \hat{y}_{OLD})^2} \qquad (2.35)$$

As a result, the distance between any point/data and the fitted plane is a perpendicular plane rather than a perpendicular line.

3. Least Absolute Deviation Method for Time Series Models

Let the simplest AR(1) model is

$$\begin{aligned} y_i^* &= y_i - \rho y_{i-1} \\ &= \beta_0 + \beta_1 x_i + \varepsilon_i - \rho(\beta_0 + \beta_1 x_{i-1} + \varepsilon_{i-1}) \\ &= \beta_0(1-\rho) + \beta_1(x_i - \rho x_{i-1}) + \varepsilon_i - \rho\varepsilon_{i-1} \end{aligned}$$

$$y_i^* = \beta_0^* + \beta_1^* x_i^* + \varepsilon_i^*$$

Similarly, from MA(1) model, for the white noise ε_i, $\varepsilon_i - \rho\varepsilon_{i-1}$ is ε_i^*, and we can find the ols estimate of ρ.

Conclusion

The extended version of this paper could be to find the estimators and their properties of the parameters of the several time series models.

References

Adnan, M. A. S, Shamsi, S., Imon, R. 2017. An Alternative Approach for Multiple Regression Model. In JSM *Proceedings*, Statistical Computing Section. Alexandria, VA: American Statistical Association. 760-771.

Adnan, M. A. S. Shamsi, S, Imon, R. 2016. A Least Deviation Approach of Fitting Regression Line: *An Alternative Approach to Least Square Estimates.* In JSM *Proceedings*, Business and Economic Statistic Section. Alexandria, VA: American Statistical Association. 304 - 317.

Shamsi, S., Adnan, M. A. S., 2018. A Moving Variance 2D Time Series Model. Unpublished paper presented in SDSS 2018.

Montogomery D. C.et al. 2008. Introduction to Time Series Analysis and Forecasting Analysis. Wiley Publisher.

Myers R. H. 1986. Classical and Modern Regression with Applications. PWS Publishers.
Robinson E. A. 1981. Least Square Regression Analysis in terms of Linear Algebra. Goose Pond Press.

Rousseeuw P. J. and Leroy A. M. 1987. Robust Regression and Outlier Detection. Wiley Publisher.

Seber, G. A. F. 1977. Linear Regression Analysis. Wiley Publishers.

https://en.wikipedia.org/wiki/Time_series