A Least-Squares Approach To Prediction The Future Sales of Pharmacy

Mohammad Abazid, Duaa Alkoud

Abstract: ***The least square methods (LSM) are widely utilized in data fitting, with the best fit minimizing the residual squared sum. LSM can be divided into two categories, the ordinary or linear LSM and the nonlinear LSM, this depends on the residuals. To statistically analyzed regression, the linear LSM most have a solution that is closed form. Conversely, the nonlinear LSM is analyzed using the iterative refinement. In this paper, the best fit of the data that correspond to pharmacy sales for the year 2019 and 2020 is evaluated using the LS method. The results revealed that the trend value for 2014, 2015, 2016, 2017, 2018, 2019, 2020 and 2021 is found to be 10200, 27300, 44400, 61500, 78600, 95700, 112800 and 129900 respectively.***

Keywords: The Least Square Methods, LSM, Pharmacy, Sales.

I. INTRODUCTION

Least square methods (LSM) have since been used as an estimation method that was independently developed in 1796 by Gauss [3], Legendre and Adrain in 1805 and 1808 respectively. It was first published in the early nineteenth century. Thus, LS is known to be one of the most widely used methods of data analysis in geophysics. LSM is a statistical way of evaluating regression analysis in order to approximate solution of over determined systems. i.e there are more unknown in a set of equations. LSM connotes that the complete solution reduces the sum of square of the residuals made in each equation results. However, LSM is widely utilized in data fitting, with the best fit minimizing the residual squared sum. LSM can be divided into two categories, the ordinary or linear LSM and the nonlinear LSM, this depends on the residuals. To statistically analyzed regression, the linear LSM most have a solution that is closed form. Conversely, the nonlinear LSM is analyzed using the iterative refinement [6]. The aim of this project is to estimate the future sales of pharmacy for three consecutive years (2019, 2020 and 2021) using the least square method.

II. PROBLEM STATEMENT

In daily problems, one is faced with solving a LSM problem and believes to have a linear relationship between variables. The LSM is widely utilized in data fitting, with the best fit minimizing the residual squared sum. LSM can be divided into two categories, the ordinary or linear LSM and the nonlinear LSM, this depends on the residuals. Several real-life problems that need future estimation or relationship between two or more variables are solved using LSM.

Revised Manuscript Received on 19 June 2018.

Mohammad Abazid, Ph.D. Candidate, Department of Civil Engineering, Near East University, Nicosia, Turkey. E-mail: mohamad.rezq@gmail.com

Duaa Alkoud, Master Student, Faculty of Pharmacy and Medical Sciences, Al-Ahliyya Amman University, Amman, Jordan. E-mail: doaa.alkoud@gmail.com

III. AIMS AND OBJECTIVES

The main aim of this study is to determine the future sales of pharmacy for the year 2019, 2020 and 2021 using the LS method.

IV. MOTIVATION

Our motivation for the topic is based on the fact that some companies and organization need to forecast their future sales using the previously obtained data so as to have a proper plan that can lead them in accomplishing their set goals and target. Also, lack of forecast data has led to the collapse of several businesses in the past.

V. REVIEW ON STATISTICS AND PROBABILITY

To understand LSM, it is important to review some statistical and probability measures. For a set of data say $x_i \dots\dots\dots\dots\dots . x_N$, the mean is given as $(x_i + \cdots \dots\dots\dots . + x_N)$/N. it can be written as [11]:

$$\bar{x} = \frac{1}{N}\sum_{n=1}^{N} x_n$$

However, the mean of the data is the average value.

The variance of $\{x_i \dots\dots\dots\dots\dots . x_N\}$ is given as σ_x^2

$$\sigma_x^2 = \frac{1}{N}\sum_{n=1}^{N}(x_i - \bar{x})^2$$

σ_x Connote the deviation, and it is the square root of variance. This measures the deviations of x's around the mean.

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x_i - \bar{x})^2}$$

It should be noted that the unit of SD is the same with the unit of x.

VI. METHODOLOGY

In this section, the LSM, optimality of LS estimate, weighted LS will be discussed.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

A. Least Square Method

One of the oldest and frequent used of ordinary least square (OLS) method is the linear regression, this relates to the problem of finding a curve or a line that best fits a set of data points. In standard form, a set of N pairs of observations

$\{x_i, y_i\}$ is used in finding a function that relate to the value of the dependent variable (Y) to the values of the independent variable (X). With one variable and a linear function, the prediction is given by the following equation [11]:

$$\hat{Y} = a + bx$$

From the data (X_i, Y_i) …….(Y_N, Y_N) the error associated to

$\hat{Y} = a + bx$ is given by:

$$E(a,b) = \sum_{n=1}^{N} (y_n - (ax_n + b))^2 .$$

It is N times the variance of the data set $\{y_1 - (ax_1 + b), \ldots, y_n - (ax_N + b)\}$ It makes no alteration whether or not they study the variance or N times the variance as the error, and note that the error is a function of two variables. The aim is to get value of a and b that reduce the error. In multivariable calculus the values of (a, b) is such that.

$$\frac{\partial E}{\partial a} = 0, \quad \frac{\partial E}{\partial b} = 0$$

It should be noted that boundary points are not considered, thus, a and b become large, the fit will visibly get worse. Differentiating E (a, b) we get [11]:

$$\begin{aligned} \frac{\partial E}{\partial a} &= \sum_{n=1}^{N} 2\left(y_n - (ax_n + b)\right) \cdot (-x_n) \\ \frac{\partial E}{\partial b} &= \sum_{n=1}^{N} 2\left(y_n - (ax_n + b)\right) \cdot 1. \end{aligned}$$

Letting $\frac{\partial E}{\partial a} = \frac{\partial E}{\partial b} = 0$, we have:

$$\begin{aligned} \sum_{n=1}^{N} (y_n - (ax_n + b)) \cdot x_n &= 0 \\ \sum_{n=1}^{N} (y_n - (ax_n + b)) &= 0. \end{aligned}$$

These equations can be written as:

$$\begin{aligned} \left(\sum_{n=1}^{N} x_n^2\right) a + \left(\sum_{n=1}^{N} x_n\right) b &= \sum_{n=1}^{N} x_n y_n \\ \left(\sum_{n=1}^{N} x_n\right) a + \left(\sum_{n=1}^{N} 1\right) b &= \sum_{n=1}^{N} y_n. \end{aligned}$$

The values of a and b that reduces the error satisfies the following equation:

$$\begin{pmatrix} \sum_{n=1}^{N} x_n^2 & \sum_{n=1}^{N} x_n \\ \sum_{n=1}^{N} x_n & \sum_{n=1}^{N} 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{n=1}^{N} x_n y_n \\ \sum_{n=1}^{N} y_n \end{pmatrix}.$$

The matrix is invertible as shown below:

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{n=1}^{N} x_n^2 & \sum_{n=1}^{N} x_n \\ \sum_{n=1}^{N} x_n & \sum_{n=1}^{N} 1 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{n=1}^{N} x_n y_n \\ \sum_{n=1}^{N} y_n \end{pmatrix}.$$

Also, the determinant of the matrix can be denoted as:

$$\det M = \sum_{n=1}^{N} x_n^2 \cdot \sum_{n=1}^{N} 1 - \sum_{n=1}^{N} x_n \cdot \sum_{n=1}^{N} x_n.$$

Also,

$$\overline{x} = \frac{1}{N} \sum_{n=1}^{N} x_n$$

It was found that:

$$\begin{aligned} \det M &= N \sum_{n=1}^{N} x_n^2 - (N\overline{x})^2 \\ &= N^2 \left(\frac{1}{N} \sum_{n=1}^{N} x_n^2 - \overline{x}^2 \right) \\ &= N^2 \cdot \frac{1}{N} \sum_{n=1}^{N} (x_n - \overline{x})^2, \end{aligned}$$

There are two main parameters in the equation above which identify the slope and intersection of the regression line. These parameters can be estimated using the LSM as the values that reduce the sum of squares. This leads to minimizing the expression:

$$\mathscr{E} = \sum_i (Y_i - \hat{Y}_i)^2 = \sum_i [Y_i - (a + bX_i)]^2$$

Here, E connotes “error" either the minimized quantity. The estimation of the parameters is found by basic results from calculus and, precisely, utilizes the property that a quadratic expression reaches its minimum value when its derivatives disappear. Taking the derivative of E with respect to a and b and letting them to zero gives the following set of equations (called the normal equations):

$$\frac{\partial \mathscr{E}}{\partial a} = 2Na + 2b \sum X_i - 2 \sum Y_i = 0$$

And

$$\frac{\partial \mathscr{E}}{\partial b} = 2b \sum X_i^2 + 2a \sum X_i - 2 \sum Y_i X_i = 0$$

Solving the normal equations generate the following LS estimates of *a* and *b* as:

$$a = M_Y - bM_X$$

(With *MY* and *MX* denoting the means of *X* and *Y*) and

$$b = \frac{\sum (Y_i - M_Y)(X_i - M_X)}{\sum (X_i - M_X)^2}$$

To extend the OLS to more than one independent variable, the matrix algebra is used.

VII. GEOMETRY OF LEAST SQUARES

The interception of geometrical structure as a projection in orthogonal is defined by those variables that are independent. There is no correlation between the actual and the predicted values, hence it is known as orthogonal projection. As presented in figure 1, X_1 and X_2 are the independent variables while y is a vector that shows the error vector y-y and it is orthogonal to LS y that lies in the subspace [3].

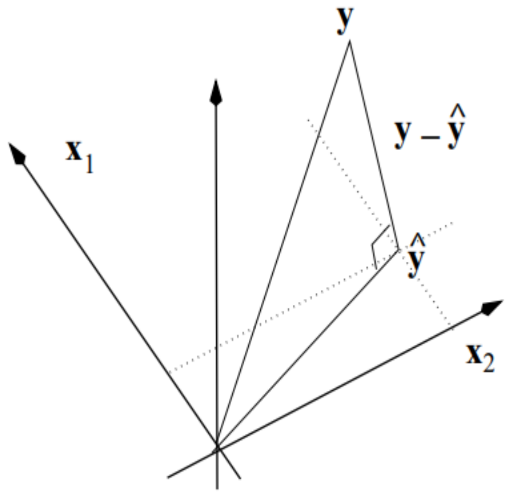

Figure 1: The LSM of The Data [13]

VIII. LS OPTIMALITY ESTIMATES

There is a solid statistical property in OLS estimates. Specifically, for data that are randomly sampled from a population that is well defined. Then there is a linear population model with an expected error with zero value, variables that are independent and linearly independent with normally distributed error that are uncorrelated. However, the unbiased linear estimate is usually written BLUE. Thus, above are known as the Gauss-Markov theorem [11].

A. Weighted Least Squares

OLS optimally depends greatly on the assumption that is homoscedastically. For data that come from different sub-populations with an available error estimate that are independent, weighted LS or generalized LS can be used in order to assign a weight to each observation which reflect the measurement uncertainty. However, for weight wi , allocated to the ith observation, will therefore be a function of the variance of this observation, represented σ 2 i . A candid weighting schema is to describe wi = σ −1. For regression that is linear, example, the values of a and b minimizing can be found using WLS:

$$\mathscr{E}_w = \sum_i w_i (Y_i - \hat{Y}_i)^2 = \sum_i w_i [Y_i - (a + bX_i)]^2 \; .$$

IX. PROBLEMS WITH LS, AND ALTERNATIVES

Regardless of it flexibility and popularity, LSM has its complications. One of the most demerits of LSM is that the outlier is highly sensitive. This is as a result of utilizing squares since squaring amplifies the amount of differences and therefore provides much tougher importance to observations that are extreme. However, to address this problem, a robust technique is used since it has less sensitivity to the outlier effect.

X. RESULTS AND DISCUSSION

Using least square methods, find the sales estimate in 2019, 2020 and 2021.

Table 1: Pharmaceutical Sales

Year	Sales ($)
2014	10000
2015	21000
2016	50000
2017	70000
2018	71000
2019	?
2020	?
2021	?

XI. MANUAL CALCULATION OF LSM

The above question is solved using the least square method as presented below:

Table 2: Pharmaceutical Sales Solution using Least Square Method

Years	Sales	X	X^2	XY	Trend value
2014	10000	-2	4	-20000	**10200**
2015	21000	-1	1	-21000	**27300**
2016	50000	0	0	0	**44400**
2017	70000	1	1	70000	**61500**
2018	71000	2	4	142000	**78600**
2019	?	3	9		**95700**
2020	?	4	16		**112800**
2021	?	5	25		**129900**

N= 5

$$\sum X = 0$$

$$\sum Y = 222000 \;\; i.e\ (sum\ of\ sales)$$

$\sum X^2$ =10

$\sum XY$ = 171000

For y = a + bx

$$\sum Y = Na + b \times \sum x$$

When 222000 = 5a+ b $\times 0$ we found a = 44400

Also, $\sum XY = a\sum X + b\sum X^2$

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

When $171000 = (222000 \times 0) + (b \times 10)$

171000= 10b we found b = 17100
Therefore, for trend value, Y = a + bx
Y = 44400 + 17100 X
Trends values can be calculated for each year.
For Y 2014, where X = -2. We have:
Y 2014 = 44400 + 17100 (-2)
Y 2014 = 10200
For Y 2015, where X = -1. We have:
Y 2014 = 44400 + 17100 (-1)
Y 2014 = 27300
For Y 2016, where X = 0. We have:
Y 2016 = 44400 + 17100 (0)
Y 2016 = 44400
For Y 2017, where X = 1. We have:
Y 2017 = 44400 + 17100 (1)
Y 2017 = 61500
For Y 2018, where X = 2. We have:
Y 2018 = 44400 + 17100 (2)
Y 2018 = 78600
For Y 2019, where X = 3. We have:
Y 2019 = 44400 + 17100 (3)
Y 2019 = 95700
For Y 2020, where X = 4. We have:
Y 2020 = 44400 + 17100 (4)
Y 2020 = 112800
For Y 2021, where X = 5. We have:
Y 2021 = 44400 + 17100 (5)
Y 2021 = 129900

XII. LSM CALCULATION USING EXCEL

Apart from manual calculations, the LSM can estimate a forecasted sales using Microsoft excel. Table 3 below shows the excel calculations.

Table 3: LSM Excel Calculation

Years	Sales
2014	10000
2015	21000
2016	50000
2017	70000
2018	71000
2019	**95700**
2020	**112660**
2021	**125178**

XIII. DISCUSSION

To estimate or forecast sales using LSM, both manual and computer can be used in achieving the set objectives. In this study, the forecast sales for 2019, 2020 and 2021 were computed using manual calculation and computer. However, both methods were found to yield the same results. Thus, LSM is seen as one of the effective and widely utilized in data fitting, with the best fit reducing the residual squared sum.

XIV. CONCLUSION

The LSM are generally used in data fitting, with the best fit minimizing the residual squared sum. LSM can be divided into two categories, the ordinary or linear LSM and the nonlinear LSM, this depends on the residuals. Also, LSM is a statistical way of evaluating regression analysis in order to approximate solution of over determined systems. i.e there are more unknown in a set of equations. In this study, LSM is used in finding the trend values for unknown years and the results revealed that the trend value for 2014, 2015, 2016, 2017, 2018, 2019, 2020 and 2021 is found to be 10200, 27300, 44400, 61500, 78600, 95700, 112800 and 129900 respectively. Thus, it can be concluded that LSM is an effect way of estimating or forecasting the future sales of product or services in a pharmacy.

REFERENCES

1. Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. Journal of the society for Industrial and Applied Mathematics, 11(2), 431-441.
2. Reid, F. (2000). The mathematician and the banknote: Carl Friedrich Gauss. Parabola, 36(2), 2-9.
3. Menke, W. (2015). Review of the generalized least squares method. Surveys in Geophysics, 36(1), 1-25.
4. Bates, D. M., & Watts, D. G. (1988). Nonlinear regression analysis and its applications Wiley. New York.
5. Harper, H. L. (1974). The method of least squares and some alternatives. Part I, II, II, IV, V, VI. International Satistical Review, 42, 147-174.
6. Seal, H. L. (1967). Studies in the History of Probability and Statistics. XV The historical development of the Gauss linear model. Biometrika, 54(1-2), 1-24.
7. Nie, N. H., Bent, D. H., & Hull, C. H. (1970). SPSS: Statistical package for the social sciences (No. HA29 S6). New York: McGraw-Hill.
8. Larson, R., & Farber, E. (2000). Elementary statistics: picturing the World. Prentice Hall.
9. Boyd, S., & Vandenberghe, L. (2004). Convex optimization. Cambridge university press.
10. Deakin, R. E. (1991). A review of Least Squares theory applied to traverse adjustment. Australian surveyor, 36(3), 245-253.
11. Miller, S. J. (2006). The method of least squares. Mathematics Department Brown University, 1-7.
12. Abdi, H. (2007). The method of least squares. Encyclopedia of Measurement and Statistics. CA, USA: Thousand Oaks.
13. Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural processing letters, 9(3), 293-300.
14. Mosteller, F. (2006). Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. In Selected Papers of Frederick Mosteller (pp. 157-162). Springer, New York, NY.
15. Teunissen, P. J. (1995). The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. Journal of geodesy, 70(1-2), 65-82.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.