

CURVE FITTING: STEP-WISE LEAST SQUARES METHOD

Dhritikesh Chakrabarty

Department of Statistics, Handique Girls' College, Guwahati – 781001, Assam, India.
e-mail: dhritikesh.c@rediffmail.com, dhritikeshchakrabarty@gmail.com

ABSTRACT:

A method has been developed for fitting of a mathematical curve to numerical data based on the application of the least squares principle separately for each of the parameters associated to the curve. The method has been termed as step-wise least squares method. In this paper, the method has been presented in the case of fitting of a polynomial curve to observed data.

Key Words : *Step-wise least squares, polynomial curve, parameter estimation.*

1. INTRODUCTION:

The method of least squares is indispensable and is widely used method for curve fitting to numerical data. The method of least squares was first discovered by the French mathematician *Legendre* in 1805 [*Mansfield* (1877), *Paris* (1805)]. The first proof of this method was given by the renowned statistician *Adrian* (1808) followed by its second proof given by the German Astronomer *Gauss* [*Hamburg* (1809)]. Apart from these two proofs as many as eleven proofs were developed at different times by a number of mathematicians viz. *Laplace* (1810), *Ivory* (1825), *Hagen* (1837), *Bassel* (1838), *Donkim* (1844), *Herscel* (1850), *Crofton* (1870) etc.. Though none of the thirteen proofs is perfectly satisfactory but yet it has given new dimension in setting the subject in a new light.

In fitting of a curve by the method of least squares, the parameters of the curve are estimated by solving the normal equations which are obtained by applying the principle of least squares with respect to all the parameters associated to the curve jointly (simultaneously). However, for a curve of higher degree polynomial and / or for a curve having many parameters, the calculation involved in the solution of the normal equations becomes more complicated as the number of normal equations then becomes larger. Moreover, In many situations, it is not possible to obtain normal equations by applying the principle of least squares with respect to all the parameters jointly. These lead to think of searching for some other method of estimating the parameters. For this reason, a new method of fitting of a curve has been framed of which is based on the application of the principle of least squares separately for each of the parameters associated to the curve. This paper is based on this method with respect to the fitting of a polynomial curve.

The problem in this study can be summarized as follows:

Let the polynomial curve which is to be fitted to the n pairs of observations

$$\{(x_i, y_i) : i = 1, 2, \dots, n\}$$

on (X, Y) where Y is a variable dependent on the variable X , be

satisfy the model

$$y_i = \mu + e_i \quad , \quad (i = 1, 2, \dots, n)$$

where μ is the true part of y_i and e_i is the error associated to y_i , then the least squares estimator of the parameter μ

is $\frac{1}{n} \sum_{i=1}^n y_i$.

3. METHOD OF ESTIMATION OF PRAMETERS:

Suppose, the polynomial in X of degree n namely

$$Y = a_0 + a_1 X + a_2 X^2 + \dots + a_k X^k \quad , \quad a_k \neq 0 \tag{3.1}$$

(where $a_0, a_1, a_2, \dots, a_k$ are parameters),

is to be fitted on the n pairs of observations

$$\{(x_i, y_i) : i = 1, 2, \dots, n\}$$

on (X, Y) .

The problem here is to determine the parameters on the basis of the observations.

Here, the notations / operator Δ is used to define $\Delta f(x_i)$ as

$$\Delta f(x_i) = f(x_{i+1}) - f(x_i) .$$

In particular,

$$\Delta x_i = x_{i+1} - x_i \quad \& \quad \Delta y_i = y_{i+1} - y_i$$

i.e. Δ operates over the suffix variable.

Also, the notation $R\{g(.) : h(.)\}$ is used to describe the ratio of $g(.)$ and $h(.)$. For example,

$$R\{\Delta y_i : \Delta x_i\} = \frac{\Delta y_i}{\Delta x_i}$$

Now, let

$$v_0 = v_0(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n),$$

$$v_1 = v_1(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n),$$

.....

$$v_k = v_k(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

be the estimators of the parameters

$$a_0, a_1, \dots, a_k$$

respectively obtained by the application of the principle of least squares separately for the parameters. Then the curve, fitted by this method, will be

$$Y = v_0 + v_1 X + v_2 X^2 + \dots + v_k X^k \tag{3.2}$$

Since the n pairs of values of (X, Y) may not lie on the fitted curve, the observed values satisfy the equations

$$y_i = v_0 + v_1 x_i + v_2 x_i^2 + \dots + v_k x_i^k + e_i \tag{3.3}$$

($i = 1, 2, 3, \dots, n$)

where e_i is the deviation of the observed value Y_i from the corresponding value

$$v_0 + v_1 X_i + v_2 X_i^2 + v_3 X_i^3 + \dots + v_k X_i^k$$

of Y yielded by the fitted curve.

This yield

$$\Delta Y_i = v_1 \Delta X_i + v_2 \Delta X_i^2 + v_3 \Delta X_i^3 + \dots + v_k \Delta X_i^k + \Delta e_i$$

from which one obtains that

$$Y_i(1) = v_1 + v_2 X_i^2(1) + v_3 X_i^3(1) + \dots + v_k X_i^k(1) + e_i(1) \tag{3.4}$$

where

$$Y_i(1) = R\{\Delta Y_i : \Delta X_i\},$$

$$X_i^r(1) = R\{\Delta X_i^r : \Delta X_i\}$$

$$(r = 2, 3, \dots, k)$$

$$\& e_i(1) = R\{\Delta e_i : \Delta X_i\}$$

$$(i = 1, 2, 3, \dots, n - 1).$$

This again yield

$$\Delta Y_i(1) = v_2 \Delta X_i^2(1) + v_3 \Delta X_i^3(1) + \dots + v_k \Delta X_i^k(1) + \Delta e_i(1)$$

$$\text{i.e. } Y_i(2) = v_2 + v_3 X_i^3(2) + \dots + v_k X_i^k(2) + e_i(2) \tag{3.5}$$

where

$$Y_i(2) = R\{\Delta Y_i(1) : \Delta X_i^2(1)\},$$

$$X_i^r(2) = R\{\Delta X_i^r(1) : \Delta X_i^2(1)\}$$

$$(r = 3, 4, \dots, k)$$

$$\& e_i(2) = R\{\Delta e_i(1) : \Delta X_i^2(1)\}$$

$$(i = 1, 2, 3, \dots, n - 2).$$

This further yield

$$\Delta Y_i(2) = v_3 \Delta X_i^3(2) + \dots + v_k \Delta X_i^k(2) + \Delta e_i(2)$$

$$\text{i.e. } Y_i(3) = v_3 + v_4 X_i^4(3) + \dots + v_k X_i^k(3) + e_i(3) \tag{3.6}$$

where

$$Y_i(3) = R\{\Delta Y_i(2) : \Delta X_i^3(2)\},$$

$$X_i^r(3) = R\{\Delta X_i^r(2) : \Delta X_i^3(2)\}$$

$$(r = 4, 5, \dots, k)$$

$$\& e_i(3) = R\{\Delta e_i(2) : \Delta X_i^3(2)\}$$

$$(i = 1, 2, 3, \dots, n - 3).$$

At the s^{th} step, one obtains that

$$\Delta Y_i(s - 1) = v_s \Delta X_i^s(s - 1) + \dots + v_k \Delta X_i^k(s - 1) + \Delta e_i(s - 1)$$

$$\text{i.e. } Y_i(s) = v_s + v_{s+1} X_i^{s+1}(s) + \dots + v_k X_i^k(s) + e_i(s) \tag{3.7}$$

where

$$Y_i(s) = R\{\Delta Y_i(s - 1) : \Delta X_i^s(s - 1)\},$$

$$X_i^r(s) = R\{\Delta X_i^r(s - 1) : \Delta X_i^s(s - 1)\}$$

$$\begin{aligned} & (r = s + 1, s + 2, \dots, k) \\ & \& e_i(s) = R\{\Delta e_i(s-1) : \Delta \mathbf{x}_i^3(s-1)\} \\ & (i = 1, 2, 3, \dots, n-s). \end{aligned}$$

At the $(k-1)^{\text{th}}$ step, one obtains that

$$\begin{aligned} \Delta \mathbf{y}_i(k-2) &= v_{k-1} \Delta \mathbf{x}_i^{k-1}(k-2) + v_k \Delta \mathbf{x}_i^k(k-2) + \Delta e_i(k-2) \\ \text{i.e. } \mathbf{y}_i(k-1) &= v_{k-1} + v_k \mathbf{x}_i^k(k-1) + e_i(k-1) \end{aligned} \quad (3.8)$$

$$\text{where } \mathbf{y}_i(k-1) = R\{\Delta \mathbf{y}_i(k-2) : \Delta \mathbf{x}_i^{k-1}(k-2)\}$$

$$\& e_i(k-1) = R\{\Delta e_i(k-2) : \Delta \mathbf{x}_i^{k-1}(k-2)\}$$

$$(i = 1, 2, 3, \dots, n-k+1).$$

Finally, at the k^{th} step one obtains that

$$\begin{aligned} \Delta \mathbf{y}_i(k-1) &= v_k \Delta \mathbf{x}_i^k(k-1) + \Delta e_i(k-1) \\ \text{i.e. } \mathbf{y}_i(k) &= v_k + e_i(k) \end{aligned} \quad (3.9)$$

$$\text{where } \mathbf{y}_i(k) = R\{\Delta \mathbf{y}_i(k-1) : \Delta \mathbf{x}_i^k(k-1)\}$$

$$\& e_i(k) = R\{\Delta e_i(k-1) : \Delta \mathbf{x}_i^k(k-1)\}$$

$$(i = 1, 2, 3, \dots, n-k).$$

This is of the form (2.1).

Therefore by Theorem (2.1),

$$\begin{aligned} v_k &= v_k(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n) \\ &= \frac{1}{n-k} \sum_{i=1}^{n-k} \mathbf{y}_i(k) \end{aligned} \quad (3.10)$$

Also from (3.8), one obtains that

$$\mathbf{y}_i(k-1) - v_k \mathbf{x}_i^k(k-1) = v_{k-1} + e_i(k-1) \quad (3.11)$$

This is also of the form (2.1).

Therefore by the same theorem,

$$\begin{aligned} v_{k-1} &= v_{k-1}(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n) \\ &= \frac{1}{n-k+1} \sum_{i=1}^{n-k+1} \{ \mathbf{y}_i(k-1) - v_k \mathbf{x}_i^k(k-1) \} \end{aligned} \quad (3.12)$$

Similarly from the equation obtained at the $(k-2)^{\text{th}}$ step, one obtains that

$$\mathbf{y}_i(k-2) - v_{k-1} \mathbf{x}_i^{k-1}(k-2) - v_k \mathbf{x}_i^k(k-2) = v_{k-2} + e_i(k-2) \quad (3.13)$$

This is also of the form (2.1).

Therefore by the same theorem,

$$\begin{aligned} v_{k-2} &= v_{k-2}(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n) \\ &= \frac{1}{n-k+2} \sum_{i=1}^{n-k+2} \{ \mathbf{y}_i(k-2) - v_{k-1} \mathbf{x}_i^{k-1}(k-2) - v_k \mathbf{x}_i^k(k-2) \} \end{aligned} \quad (3.14)$$

Applying the same technique at the earlier steps successively, one can obtain the estimators of the parameters as follows:

$$v_s = v_s(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

$$= \frac{1}{n-s} \sum_{i=1}^{n-s} \{y_i(s) - v_{s+1} x_i^{s+1}(s) - v_{s+2} x_i^{s+2}(s) - \dots - v_k x_i^k(s)\} \quad (3.15)$$

$$v_{s-1} = v_{s-1}(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

$$= \frac{1}{n-s+1} \sum_{i=1}^{n-s+1} \{y_i(s-1) - v_s x_i^s(s-1) - v_{s+1} x_i^{s+1}(s-1) - \dots - v_k x_i^k(s-1)\} \quad (3.16)$$

$$v_1 = v_1(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

$$= \frac{1}{n-1} \sum_{i=1}^{n-1} \{y_i(1) - v_2 x_i^2(1) - v_3 x_i^3(1) - \dots - v_k x_i^k(1)\} \quad (3.17)$$

$$v_0 = v_0(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

$$= \frac{1}{n} \sum_{i=1}^n \{y_i - v_1 x_i - v_2 x_i^2 - \dots - v_k x_i^k\} \quad (3.18)$$

Note:

In this method, the estimates of the parameters are to be computed in the following order:

First, estimate v_k of the parameter a_k ,

Then, estimate v_{k-1} of the parameter a_{k-1} ,

.....

next, estimate v_s of the parameter a_s ,

next, estimate v_{s-1} of the parameter a_{s-1} ,

.....

next, estimate v_1 of the parameter a_1 ,

finally, estimate v_0 of the parameter a_0 .

Special Cases

Case -1 (Fitting of a linear curve):

If the curve $Y = f(X)$ is linear of the form

$$Y = a_0 + a_1 X \quad (3.19)$$

where a_0 & a_1 are the parameters, the estimators

$$v_0 = v_0(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

$$\& v_1 = v_1(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

of the parameters a_0 & a_1 respectively will be as follows:

$$v_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} y_i(1) \quad (3.20)$$

$$\& v_0 = \frac{1}{n} \sum_{i=1}^n (y_i - v_1 x_i) \quad (3.21)$$

where $y_i(1) = R\{\Delta y_i : \Delta x_i\}$, $(i = 1, 2, 3, \dots, n-k+1)$.

Case - 2 (Fitting of a quadratic curve):

If the curve $Y = f(X)$ is linear of the form

$$Y = a_0 + a_1 X + a_2 X^2 \quad (3.22)$$

where a_0 , a_1 & a_2 are the parameters, the estimators

$$v_0 = v_0(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n),$$

$$v_1 = v_1(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

$$\& v_2 = v_2(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

of the parameters a_0 , a_1 & a_2 respectively will be as follows:

$$v_2 = \frac{1}{n-2} \sum_{i=1}^{n-2} \mathbf{y}_i(2) \quad (3.23)$$

$$v_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} \{\mathbf{y}_i(1) - v_2 \mathbf{x}_i^2(1)\} \quad (3.24)$$

$$v_0 = \frac{1}{n} \sum_{i=1}^n \{\mathbf{y}_i - v_1 \mathbf{x}_i - v_2 \mathbf{x}_i^2\} \quad (3.25)$$

where $\mathbf{y}_i(1) = R\{\Delta \mathbf{y}_i : \Delta \mathbf{x}_i\}$, ($i = 1, 2, 3, \dots, n-1$),

$\mathbf{x}_i^2(1) = R\{\Delta \mathbf{x}_i^2 : \Delta \mathbf{x}_i\}$, ($i = 1, 2, 3, \dots, n-1$)

& $\mathbf{y}_i(2) = R\{\Delta \mathbf{y}_i(1) : \Delta \mathbf{x}_i^2(1)\}$, ($i = 1, 2, 3, \dots, n-2$).

Case - 5 (Fitting of a cubic curve):

If the curve $Y = f(X)$ is linear of the form

$$Y = a_0 + a_1 X + a_2 X^2 + a_3 X^3 \quad (3.26)$$

where a_0 , a_1 , a_2 & a_3 are the parameters, the estimators

$$v_0 = v_0(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n),$$

$$v_1 = v_1(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n),$$

$$v_2 = v_2(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

$$\& v_3 = v_3(x_1, x_1, \dots, x_n : y_1, y_1, \dots, y_n)$$

of the parameters a_0 , a_1 , a_2 & a_3 respectively will be as follows:

$$v_3 = \frac{1}{n-3} \sum_{i=1}^{n-3} \mathbf{y}_i(3) \quad (3.27)$$

$$v_2 = \frac{1}{n-2} \sum_{i=1}^{n-2} \{\mathbf{y}_i(2) - v_3 \mathbf{x}_i^3(2)\} \quad (3.28)$$

$$v_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} \{\mathbf{y}_i(1) - v_2 \mathbf{x}_i^2(1) - v_3 \mathbf{x}_i^3(1)\} \quad (3.29)$$

$$v_0 = \frac{1}{n} \sum_{i=1}^n (y_i - a_0 - a_1 x_i - a_2 x_i^2 - a_3 x_i^3) \tag{3.30}$$

where $y_i(1) = R\{\Delta y_i : \Delta x_i\}$, ($i = 1, 2, 3, \dots, n-1$),
 $x_i^2(1) = R\{\Delta x_i^2 : \Delta x_i\}$, ($i = 1, 2, 3, \dots, n-1$),
 $x_i^3(1) = R\{\Delta x_i^3 : \Delta x_i\}$, ($i = 1, 2, 3, \dots, n-1$),
 $x_i^3(2) = R\{\Delta x_i^3(1) : \Delta x_i^2(1)\}$, ($i = 1, 2, 3, \dots, n-2$),
 $y_i(2) = R\{\Delta y_i(1) : \Delta x_i^2(1)\}$, ($i = 1, 2, 3, \dots, n-2$)
 & $y_i(3) = R\{\Delta y_i(2) : \Delta x_i^3(2)\}$, ($i = 1, 2, 3, \dots, n-3$).

4. NUMERICAL EXAMPLES:

Example 4.1: Consider the problem of fitting of the linear curve

$$Y = a_0 + a_1 X$$

(where a_0 & a_1 are the parameters)

to the following observations on X and Y :

x_i :	0	5	7	10	16	20	21	25	31	40
y_i :	2	17	24	33	51	63	64	76	94	123

In order to determine the estimates of a_0 & a_1 , the following table (**Table - 4.1**) is to be prepared:

Table-4.1

(1)	(2)	(3)	(4)	(5)	(6)
x_i	y_i	Δx_i	Δy_i	$y_i(1)$	$y_i - v_1 x_i$ with $v_1 = 3.001234568$
0	2	5	17	3.4	-1.001234568
5	17	2	7	3.5	1.993827160
7	24	3	9	3	2.991358024
10	33	6	18	3	2.987654320
16	51	4	12	3	2.980246912
20	63	1	2	2	2.975308640
21	65	4	12	3	1.974074072
25	77	6	18	3	1.969135800
31	95	9	28	3.1111111	1.961728392
40	123				2.950617280
Total = 175	Total = 550			Total =27.0111111	Total = 24.78305064

Now, $v_1 = \frac{1}{9} \sum_{i=1}^9 y_i(1) = 3.001234568$

$$\& v_0 = \frac{1}{10} \sum_{i=1}^{10} (Y_i - v_1 X_i) = 2.478305064$$

Thus the linear curve fitted to the observations becomes

$$Y = 2.478305064 + 3.001234568 X$$

Example 4.2: Consider the problem of fitting of the quadratic curve

$$Y = a_0 + a_1 X + a_2 X^2$$

(where a_0 , a_1 & a_2 are parameters)

to the following observations on X and Y :

x_i :	0	1	3	4	7	8	10	12
y_i :	5	9	19	30	69	84	124	175

In order to determine the estimates of a_0 , a_1 & a_2 , the following table (**Table - 4.2**) is to be prepared:

Table - 4.2

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
x_i	y_i	Δx_i	Δx_i^2	Δy_i	$x_i^2(1)$	$y_i(1)$	$\Delta x_i^2(1)$
0	5	1	1	4	1	4	3
1	9	2	8	10	4	5	3
3	19	1	7	11	7	11	4
4	30	3	33	39	11	13	4
7	69	1	15	15	15	15	3
8	84	2	36	40	18	20	4
10	124	2	44	51	22	25.5	
12	175						
Total = 45							

Table - 4.2 continued

(9)	(10)	(11)	(12)
$\Delta y_i(1)$	$y_i(2)$	$y_i(1) - v_2 x_i^2(1)$ with $v_2 = 1.063$	$y_i - v_1 x_i - v_2 x_i^2$ with $v_1 = 1.512$ & $v_2 = 1.063$
1	0.333333333	2.937	5
6	2	0.748	6.425
2	0.5	3.559	4.897
2	0.5	1.307	6.944

5	1.666666666	- 0.945	6.329
5.5	1.375	0.866	3.872
		2.114	2.588
			3.784
	Total = 6.375	Total = 10.586	Total = 39.839

Now, $v_2 = \frac{1}{6} \sum_{i=1}^6 y_i(2) = 1.063$,

$$v_1 = \frac{1}{7} \sum_{i=1}^7 \{y_i(1) - v_2 x_i^2(1)\} = 1.512$$

$$\& v_0 = \frac{1}{8} \sum_{i=1}^8 (y_i - a_0 - a_1 x_i - a_2 x_i^2 - a_3 x_i^3) = 4.98$$

Thus the quadratic curve fitted to the observations becomes

$$Y = 4.98 + 1.512 X + 1.063 X^2$$

5. CONCLUSION:

- (1) The method developed in this paper is based on the application of the principle of least squares with respect to the parameters separately while the usual method of least squares is based on the application of the principle of least squares with respect to the parameters jointly.
- (2) The method is suitable for fitting of a polynomial curve of any finite order.
- (3) It is yet to be investigated whether this method can be applicable to a curve other than polynomial curve to observed values.
- (4) It is yet to be searched for whether the estimates of parameter obtained by this method and those obtained by solving the normal equations of the curve under study are identical.

6. REFERENCE:

1. *Adrian A.* (1808): “ The Analyst ”, No.- IV, 93-109.
2. *Bessel* (1838): “ On utersuchungen ueber die Wahrscheinlichkeit der Beobachtungsfehler ”.
3. *Crofton* (1870): “ On the Proof of the Law of Error of Observations”, Phil. Trans, London, 175-188.
4. *Donkin* (1844): “ An Essay on the Theory of the Combination of Observations ”, Joua. Math., XV, 297-322.
4. *Gauss C. F.* (1809): “ Theory of Motion of Heavenly Bodies ”, Humburg , 205-224.
6. *Hagen* (1837): “ Grandzuge der Wahrscheinlichkeitsrechnung ”, Berlin.
7. *Herschel J.* (1850): “ Quetelet On Probabilities ”, Edinburgh Review, XCII, 1-57.
8. *Ivory* (1825): “ On the Method of Least Squares ”, Phil. Mag., LX, 3-10.
9. *Koppelman R.* (1971): “ The Calculus of Operation ”, History of Exact Sc., 8, 152-242.
10. *Laplace* (1810): “ Theoretic Analytique des prob. ”, Ch-IV.

11. *Merriman M.* (1877): “The Analyst”, Vol. -IV, NO-2.
12. D. Chakrabarty [2010] “Probability as the Maximum Occurrence of Relative Frequency” *Aryabhatta J. of Maths & Info.*, Vol. 2 [2], pp 339-344.
13. D. Chakrabarty [2013] “One Table of Three Digit Random Numbers” *Aryabhatta J. of Maths & Info.*, Vol. 5 [2], pp 279-284.