Numerical Analysis and Computer Programming

Curve fitting by least squares

Md. Ashraful Islam **Lecturer, Dept. of BECM, RUET**

Least- Square Method

Let the set of data points be (x_i, y_i), i = 1, 2, …, m, and let the curve given by **Y = f(x)** be fitted to this data. At x = x_i, the experimental (observed) value of the ordinate is y_i and the corresponding value on the fitting curve is f(x_i). If e_i is the error of approximation at x = x_i, then we have

$\boldsymbol{e_i}$ **=** $\boldsymbol{y_i}$ **- f(**$\boldsymbol{x_i}$**).**

We can write

S = $[y_1 - f(x_1)]^2 + [y_2 - f(x_2)]^2 + [y_3 - f(x_3)]^2 + \ldots + [y_m - f(x_m)]^2$

= ${e_1}^2 + {e_2}^2 + {e_3}^2 + \ldots + {e_m}^2$ (i.e. the sum of the squares of the errors)

Fitting a Straight Line

Let $\mathbf{Y} = \boldsymbol{a_0} + \boldsymbol{a_1}\mathbf{x}$ be the straight line to be fitted to the given data. Then, the corresponding Eq. we have

S = $[y_1 - (a_0 + a_1x_1)]^2 + [y_2 - (a_0 + a_1x_2)]^2 + \ldots + [y_m - (a_0 + a_1x_m)]^2$

For the minimum value of S,

$\frac{\delta S}{\delta a_0} = 0 = -2[y_1 - (a_0 + a_1x_1)] - 2[y_2 - (a_0 + a_1x_2)] - \ldots - 2[y_m - (a_0 + a_1x_m)]$ and

$\frac{\delta S}{\delta a_1} = 0 = -2x_1[y_1 - (a_0 + a_1x_1)] - 2x_2[y_2 - (a_0 + a_1x_2)] - \ldots - 2x_m[y_m - (a_0 + a_1x_m)]$

Fitting a Straight Line

From $\frac{\delta S}{\delta a_0}$ = 0, we get

$y_1 + y_2 + \ldots + y_m = (a_0 + a_1x_1) + (a_0 + a_1x_2) + \ldots + (a_0 + a_1x_m)$

$\sum_{i=1}^{m} y_i = ma_0 + a_1(x_1 + x_2 + \ldots + x_m)$

$\boldsymbol{\sum_{i=1}^{m} y_i = ma_0 + a_1 \sum_{i=1}^{m} x_i}$ ....(i)

From $\frac{\delta S}{\delta a_1}$ = 0

$a_0(x_1 + x_2 + \ldots + x_m) + a_1({x_1}^2 + {x_2}^2 + \ldots + {x_m}^2) = x_1y_1 + x_2y_2 + \ldots + x_my_m$

$\boldsymbol{a_0 \sum x_i + a_1 \sum {x_i}^2 = \sum x_iy_i}$ ....(ii)

Example-1

Given the following noisy data fit a straight line to this data by using least square method:

i	1	2	3	4	5
x	2.10	6.22	7.17	10.52	13.68
f(x)	2.90	3.83	5.98	5.71	7.74

Sol^n:

Let the equation of the straight line is $\mathbf{y} = a_0 + a_1 \mathbf{x}$

We have,

$\sum y_i = m a_0 + a_1 \sum x_i$ ……………………………………(i)

$\sum x_i y_i = a_0 \sum x_i + a_1 \sum {x_i}^2$ ……………………………(ii)

i	1	2	3	4	5	$m = 5$
x	2.10	6.22	7.17	10.52	13.68	$\sum \boldsymbol{x_i}$ = 39.69
f(x)	2.90	3.83	5.98	5.71	7.74	$\sum \boldsymbol{y_i}$ = 26.16
x^2	4.41	38.68	51.41	110.67	187.14	$\sum \boldsymbol{x_i}^2$ = 392.31
$x_i y_i$	6.09	23.82	42.88	60.07	105.88	$\sum \boldsymbol{x_i y_i}$ = 238.74

From Eq. (i)

5 a_0 + 39.69 a_1 = 26.16 …………..(iii)

From Eq. (ii)

39.69 a_0 + 392.31 a_1 = 238.74 …………..(iv)

From Eq. (iii) & (iv), we get

a_0 = 2.04

a_1 = 0.40

y = 2.04 + 0.40x

Nonlinear Curve Fitting

❖ Power function:

Let **y** = $\boldsymbol{ax^c}$ be the function to be fitted to the given data.

log y = log a + c log x [taking logarithms of both sides]

Which is of the form **Y** = $\boldsymbol{a_0 + a_1}$**X** ,

Where

Y = log y

a_0 = log a

a_1 = c and

X = log x

❖ Exponential function:

Let the curve, **y** = $\boldsymbol{a_0 e^{a_1 x}}$ be fitted to the given data.

log y = log a_0 + a_1x [taking logarithms of both sides]

Which can be written in the form, **Y = A + Bx**

where,

Y = log y

A = log a_0

B = a_1

Nonlinear Curve Fitting

❖ Polynomial of the n^{th} degree:

Let the polynomial of the n^{th} degree

$$Y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_n x^n$$

$$\sum Y = m a_0 + a_1 \sum x + a_2 \sum x^2 + \ldots + a_n \sum x^n$$

$$\sum XY = a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 + \ldots + a_n \sum x^{n+1}$$

$$\sum X^2 Y = a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 + \ldots + a_n \sum x^{2n}$$

……………………………………………………………………………

$$\sum X^n Y = a_0 \sum x^n + a_1 \sum x^{n+1} + \ldots\ldots + a_n \sum x^{2n}$$

Example-2

Fit a polynomial of the second degree to the data points given in the following table

x	y
0.0	1.0
1.0	6.0
2.0	17.0

***Sol^n*:**

Let the polynomial of 2nd degree

$$\mathbf{y = a_0 + a_1 x + a_2 x^2} \quad \text{.................................(i)} \qquad \text{and}$$

$$\sum y = m a_0 + a_1 \sum x + a_2 \sum x^2 \quad \text{.........................(ii)}$$

$$\sum xy = a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3 \quad \text{.....................(iii)}$$

$$\sum x^2 y = a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4 \quad \text{........................(iv)}$$

x	0	1	2	$\sum x$ = 3
y	1	6	17	$\sum y$ = 24
x^2	0	1	4	$\sum x^2$ = 5
x^3	0	1	8	$\sum x^3$ = 9
x^4	0	1	16	$\sum x^4$ = 17
xy	0	6	34	$\sum xy$ = 40
x^2y	0	6	68	$\sum x^2 y$ = 74

From Eq. (ii), Eq. (iii) and Eq. (iv), we obtain the Equations-

$3a_0 + 3a_1 + 5a_2 = 24$

$3a_0 + 5a_1 + 9a_2 = 40$

$5a_0 + 9a_1 + 17a_2 = 74$

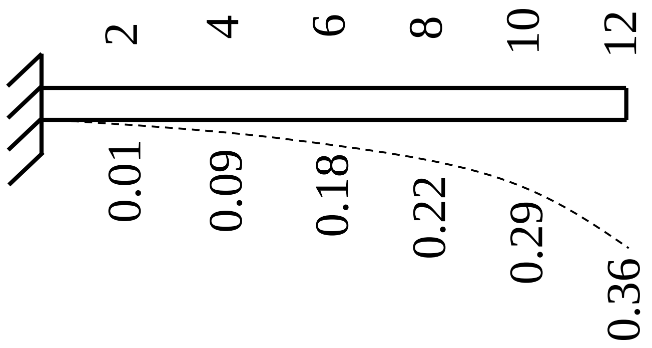
the solution to which is a_0 = 1, a_1 = 2 and a_2 = 3.

Therefore the required polynomial is

$$\mathbf{y = 1 + 2x + 3x^2}$$

Exercises

- In a laboratory the deformation (mm) of a pre-stressed cantilever beam with respect to distance (m) from the fixed support is shown below:



Fit a parabolic curve line from this data to find the deflection for any point from the fixed support.

Exercises

- Determine the constants a and b by the method of least squares such that y = ae^{bx} fits the following data-

x	y
2	4.077
4	11.084
6	30.128
8	81.897
10	222.62

- Given the data in the following table use the least square criteria to fit a function of the form AX^B to this data-

i	1	2	3	4	5	6
x	1.2	2.8	4.3	5.4	6.8	7.9
y	2.1	11.5	28.1	41.9	72.3	91.4

Thank You All