

5. Curve Fitting

In **Curve Fitting** we are given n points (pairs of numbers) (x_1, y_1) , $(x_2, y_2), \dots, (x_n, y_n)$ and we want to determine a function $f(x)$ approximately. The type of function (for example, polynomials, exponential functions, sine and cosine functions) may be suggested by the nature of the problem (the underlying physical law, for instance), and in many cases a polynomial of a certain degree will be appropriate.

Errors: (error = exact value - approx. value)

1. Max error

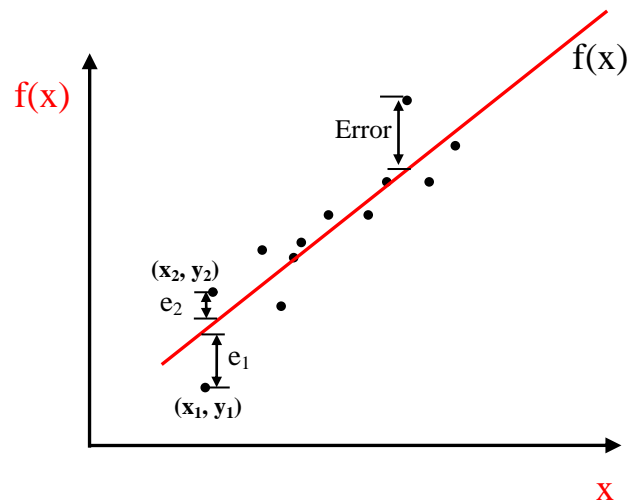
$$E_{\infty}(f) = \max_{1 < i < n} |y_i - f(x_i)|$$

2. Average error

$$E_a(f) = \frac{1}{n} \sum_{i=1}^n |y_i - f(x_i)|$$

3. **Root Mean Square (RMS) error**

$$E_{rms}(f) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2}$$



Using **Least Square Method** to find approximate function to minimize error

$$E_{rms}(f) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2} \xrightarrow{\min.} E = \sum_{i=1}^n (y_i - f(x_i))^2$$

5.1 Linear Fitting

$f(x) = ax + b$ a, b : constants

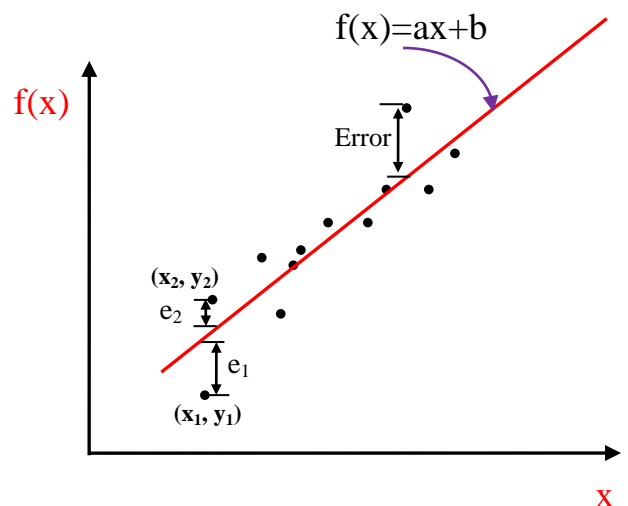
$$e_1 = [y_1 - f(x_1)]^2 = [y_1 - (ax_1 + b)]^2$$

$$e_2 = [y_2 - y(x_2)]^2 = [y_2 - (ax_2 + b)]^2$$

⋮

$$e_i = [y_i - f(x_i)]^2 = [y_i - (ax_i + b)]^2$$

$$E = \sum_{i=1}^n (e_i) \quad n: \text{No. of points}$$



$$\therefore E = \sum_{i=1}^n [y_i - (ax_i + b)]^2 \quad \text{Should be minimum.}$$

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

$$\therefore \frac{\partial E}{\partial a} = \sum_{i=1}^n 2 \cdot [y_i - (ax_i + b)] \cdot (-x_i) = 0$$

$$\sum_{i=1}^n [y_i - (ax_i + b)] x_i = 0$$

$$\sum_{i=1}^n y_i x_i = \sum_{i=1}^n (ax_i^2 + bx_i) \quad \text{divided both sides by n}$$

$$\overline{yx} = a\overline{x^2} + b\overline{x} \quad \dots(1)$$

$$\frac{\partial E}{\partial b} = \sum_{i=1}^n 2 \cdot [y_i - (ax_i + b)] \cdot (-1) = 0$$

$$\sum_{i=1}^n y_i = \sum_{i=1}^n (ax_i + b) \quad \text{divided both sides by n}$$

$$\overline{y} = a\overline{x} + b \quad (2)$$

$$\overline{yx} = a\overline{x^2} + b\overline{x} \quad (1)$$

Solve the set Eqs. using any method

Multiply eq.(2) by \overline{x} and subtract from eq.(1);

$$\therefore a = \frac{\overline{yx} - \overline{y} \overline{x}}{\overline{x^2} - (\overline{x})^2}, \quad \text{and} \quad b = \overline{y} - a\overline{x}$$

Ex. Find the best line representing the following points (0, 1), (1, 2), (2, 4), (3, 8), (4, 10) and calculate the approximate errors.

Sol. $n = 5$, using equation of line $f(x)=ax+b$

i	x_i	y_i	$y_i x_i$	x_i^2	$f(x_i)=ax_i+b$	e_i	e_i^2
1	0	1	0	0	0.2	0.8	0.64
2	1	2	2	1	2.6	-0.6	0.36
3	2	4	8	4	5.0	-1	1
4	3	8	24	9	7.4	0.6	0.36
5	4	10	40	16	9.8	0.2	0.04
Σ	10	25	74	30	-	-	2.4
Avg.	2	5	14.8	6			

$$a = \frac{\overline{yx} - \overline{y}\overline{x}}{\overline{x^2} - (\overline{x})^2} = \frac{14.8 - (5)(2)}{6 - (2)^2} = 2.4$$

$$b = \overline{y} - a\overline{x} = 5 - (2.4)(2) = 0.2$$

$$\therefore y = 2.4x + 0.2$$

To calculate the approximate errors. **Root Mean Square (RMS)**

$$E_{rms}(f) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2} = \sqrt{\frac{1}{5}(2.4)} = 0.693$$

5.2 Polynomial fitting

In General

$$y(x) = a_m x^m + a_{m-1} x^{m-1} + a_{m-2} x^{m-2} + \dots + a_0$$

where **m** degree of polynomial

- **For example m = 2**

$$y(x) = a_2 x^2 + a_1 x + a_0$$

$$\frac{\partial E}{\partial a_2} = 0 \quad \Rightarrow \quad \overline{yx^2} = a_2 \overline{x^4} + a_1 \overline{x^3} + a_0 \overline{x^2} \quad \dots(1)$$

$$\frac{\partial E}{\partial a_1} = 0 \quad \Rightarrow \quad \overline{yx} = a_2 \overline{x^3} + a_1 \overline{x^2} + a_0 \overline{x} \quad \dots(2)$$

$$\frac{\partial E}{\partial a_0} = 0 \quad \Rightarrow \quad \overline{y} = a_2 \overline{x^2} + a_1 \overline{x} + a_0 \quad \dots(3)$$

Ex. Find the best second degree curve representing the following points (0, 1), (1, 3), (2, 5), (3, 11), (4, 15).

Sol.

$$y(x) = a_2x^2 + a_1x + a_0 \quad ; \quad n = 5$$

i	x	y	x ²	x ³	x ⁴	yx	yx ²
1	0	1	0	0	0	0	0
2	1	3	1	1	1	3	3
3	2	5	4	8	16	10	20
4	3	11	9	27	81	33	99
5	4	15	16	64	256	60	240
Σ	10	35	30	100	354	106	362
Avg.	2	7	6	20	70.8	21.2	72.4

$$\overline{yx^2} = a_2\overline{x^4} + a_1\overline{x^3} + a_0\overline{x^2} \quad \Rightarrow \quad 72.4 = 70.8a_2 + 20a_1 + 6a_0 \quad \dots(1)$$

$$\overline{yx} = a_2\overline{x^3} + a_1\overline{x^2} + a_0\overline{x} \quad \Rightarrow \quad 21.2 = 20a_2 + 6a_1 + 2a_0 \quad \dots(2)$$

$$\overline{y} = a_2\overline{x^2} + a_1\overline{x} + a_0 \quad \Rightarrow \quad 7 = 6a_2 + 2a_1 + a_0 \quad \dots(3)$$

Solve the above linear simultaneous equation using iteration or Gauss method to find a_2 , a_1 , and a_0

5.3 Exponential fitting, $y = ae^{bx}$

Step (1) Take ln for both sides (to convert Exp. Function to linear function);

$$\ln y = \ln a + bx$$

$$Y = A + bx$$

Step (2) Perform linear fitting to evaluate b, A

$$b = \frac{\overline{Yx} - \overline{Y}\overline{x}}{\overline{x^2} - (\overline{x})^2}, \quad \text{and} \quad A = \overline{Y} - b\overline{x}$$

$$\therefore a = e^A$$

Or solve equations:

$$\overline{Y} = A + b\overline{x} \quad (1)$$

$$\overline{Yx} = A\overline{x} + b\overline{x^2} \quad (2)$$

Ex. Find the best exponential function representing the following points (1, 4.7), (2, 9), (3, 25), (4, 80) and calculate the approximate errors.

Sol. $n = 4$

$$y = ae^{bx} \Rightarrow \ln y = \ln a + bx \Rightarrow Y = A + bx$$

where $Y = \ln y$, $A = \ln a$

i	x	y	Y=lny	Yx	x ²	y=ae ^{bx}	e _i	e _i ²
1	1	4.7	1.548	1.548	1	4.437	0.263	0.07
2	2	9	2.197	4.394	4			
3	3	25	3.219	9.387	9			
4	4	80	4.382	17.53	15			
∑	10	118.4	11.346	32.859	30			
Avg.	2.5	29.67	2.837	8.215	7.5			

$$b = \frac{\overline{Yx} - \overline{Y} \overline{x}}{\overline{x^2} - (\overline{x})^2} = \frac{8.215 - 2.837 * 2.5}{7.5 - (2.5)^2} = 0.8978$$

$$A = \overline{Y} - b\overline{x} = 2.837 - 0.8978 * 2.5 = 0.5925$$

$$\therefore a = e^A = 1.808$$

$$\therefore y = 1.808e^{0.8978x}$$

5.4 Power fitting $y = ax^b$

Step (1) Take ln for both sides;

$$\ln y = \ln a + b \ln x$$

$$Y = A + bX$$

Step (2) Perform linear fitting to evaluate b, A

$$b = \frac{\overline{YX} - \overline{Y} \overline{X}}{\overline{X^2} - (\overline{X})^2}, \text{ and } A = \overline{Y} - b\overline{X}$$

$$\therefore a = e^A$$

or Solve equation $\overline{Y} = A + b\overline{X}$ (1)

$$\overline{YX} = A\overline{X} + b\overline{X^2} \quad (2)$$

H.W. For the same points in the last example, find the best power function? and calculate the approximation error. Ans. $y = 3.614 x^{1.954}$

Ex. An investigator has reported the data tabulated below. It is known that such data can be modeled by the following equation

$$x = e^{(y-b)/a}$$

where a & b are parameters. Use a transformation to linearize this equation and then employ linear regression to determine a and b.

Based on your analysis predict y at x = 2.6

x	1	2	3	4	5
y	0.5	2	2.9	3.5	4

Answer: $x = e^{(y-0.49)/2.177}$