Numerical Analysis

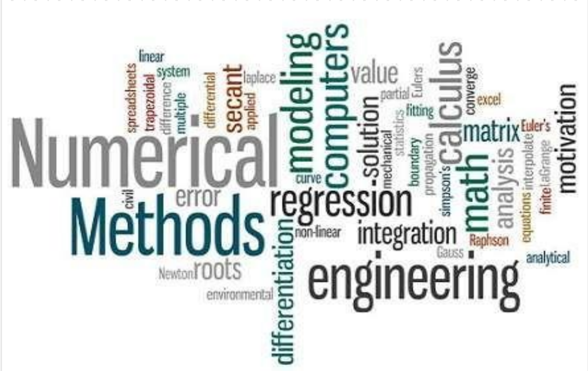
Dr. Farman Ali

Assistant Professor

DEPARTMENT OF SOFTWARE

SEJONG UNIVERSITY

Week 11, Lecture-I-II



farmankanju@sejong.ac.kr

Course Syllabus

- **Introduction:** Numerical analysis and background, definitions of computer number systems, floating point representation, representation of numbers in different bases, and round-off errors.
- **Root Finding:** Bisection method, fixed-point iteration, Newton's method, the secant method and their error analysis and order of convergence (Newton’s method and Secant method).
- **Direct Methods for Solving Linear Systems:** Gaussian elimination, LU decomposition, pivoting strategies, and PA=LU-factorization,…..
- **Polynomial:** Polynomial interpolation, Lagrange interpolation, Piecewise interpolation, divided differences interpolation, and curve fitting in interpolation (Application: Regression)…...
- **Integration:** Numerical differentiation, numerical integration, and composite numerical integration…….
- **Ordinary Differential Equations:** Euler's Method, and Runge-Kutta methods…..

Polynomial Interpolation

- **Interpolation and Curve Fitting**

- **Newton's Divided Difference Interpolating Polynomial**

- **Linear Regression**

Interpolation and Curve Fitting

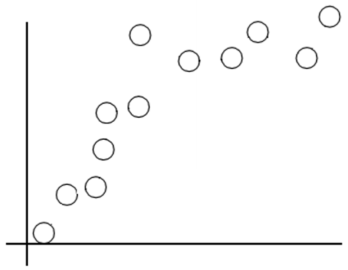
Interpolation

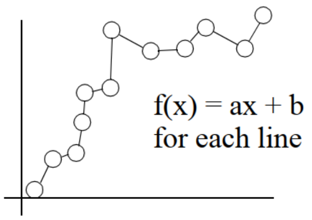
- The method of finding a value between known values of data points.
  - It determines a polynomial that gives the exact value at the data points

- If there is small number of data points
  - A single polynomial is enough

- If there is large number of data points
  - Different polynomials are used in the intervals between the points.
  - If data is reliable, we can plot it and connect the dots. This is piece-wise linear interpolation.

- Since its really a group of small s, connecting one point to the next it doesn't work very well for data that has built in random error (scatter).





Interpolation

Interpolation and Curve Fitting

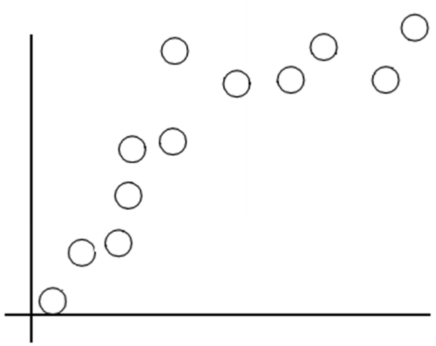
Curve Fitting

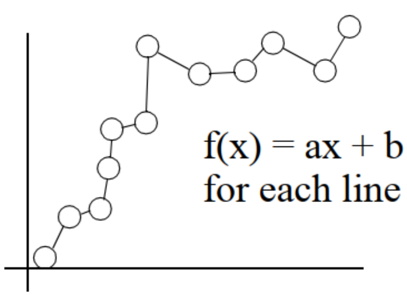
- ❑ Capturing the trend in the data by assigning a single function across the entire range.

- ❑ Objective of curve fitting
  - To identify a function that fits the overall data points.
  - The function does not have to give the exact value at any single point, but fits the data well overall.
  - It can be used typically when the values of data points have some error.

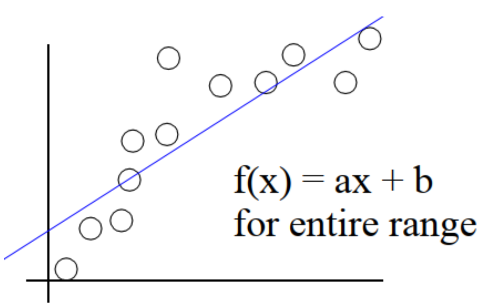
- ❑ Curve fitting can be carried out with many types of functions and with polynomials of various orders.

- ❑ A straight line is described generically by $f(x) = ax + b$.



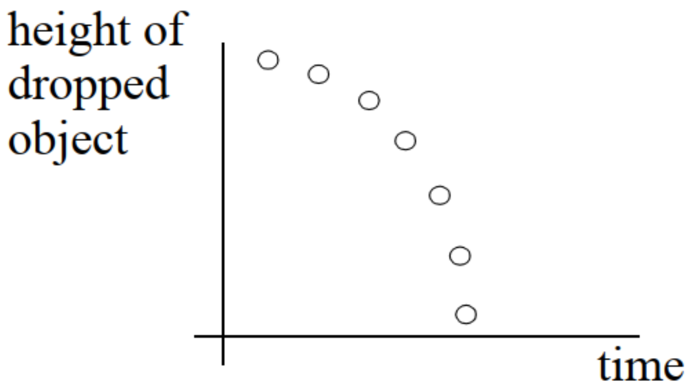


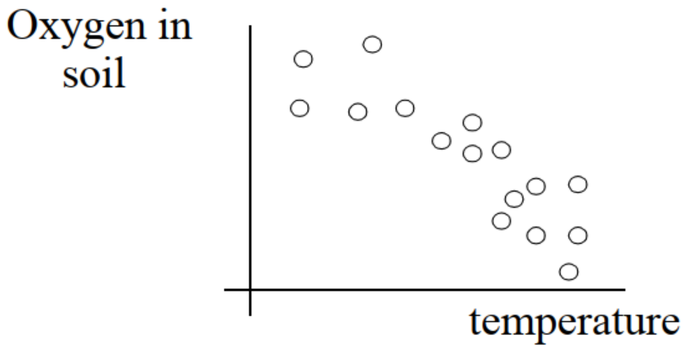
Interpolation

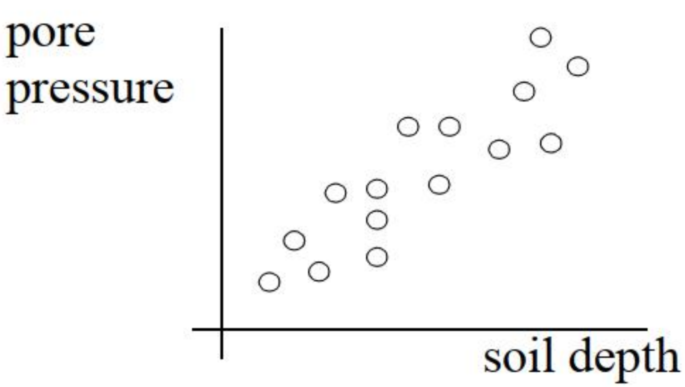


Curve Fitting

Interpolation and Curve Fitting







❑Interpolation

- Lagrange Interpolation Polynomials
- Piecewise Interpolation Polynomials
- Newton’s Divided Difference Interpolating Polynomial

❑Regression

- Linear Regression
- Polynomial Regression

Newton's Divided Difference Interpolating Polynomial

Newton's Divided Difference Interpolating Polynomial

❑ What is divided differences

- Suppose that $P_n(x)$ is the nth Lagrange polynomial that agrees with the function f at the distinct numbers $x_0, x_1, x_2 \dots, x_n$. The divided differences of f with respect to $x_0, x_1, x_2 \dots, x_n$ are used to express $P_n(x)$ in the form.

$$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots + a_n(x - x_0) \dots (x - x_{n-1})$$

For appropriate constants $a_0, a_1, a_2 \dots, a_n$.

- To determine the first of these constants, a_0, note that if $P_n(x)$ is written in the form of the above equation, then evaluating $P_n(x)$ at a_0 leaves only the constant term a_0; that is

$$\mathbf{x_0}: \quad a_0 = P_n(x_0) = f(x_0).$$

- Similarly, when $P(x)$ is evaluated at x_1, the only nonzero terms in the evaluation of $P_n(x_1)$ are the constant and linear terms,

$$\text{at } \mathbf{x_1}: \quad f(x_0) + a_1(x_1 - x_0) = P_n(x_1) = f(x_1)$$

$$\Rightarrow a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Newton's Divided Difference Interpolating Polynomial

Newton's Divided Difference Interpolating Polynomial

❑ Generating the divided differences table

x	$f(x)$	First divided differences	Second divided differences	Third divided differences
x_0	$f[x_0]$			
		$f[x_0, x_1] = \dfrac{f[x_1] - f[x_0]}{x_1 - x_0}$		
x_1	$f[x_1]$		$f[x_0, x_1, x_2] = \dfrac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}$	
		$f[x_1, x_2] = \dfrac{f[x_2] - f[x_1]}{x_2 - x_1}$		$f[x_0, x_1, x_2, x_3] = \dfrac{f[x_1, x_2, x_3] - f[x_0, x_1, x_2]}{x_3 - x_0}$
x_2	$f[x_2]$		$f[x_1, x_2, x_3] = \dfrac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1}$	
		$f[x_2, x_3] = \dfrac{f[x_3] - f[x_2]}{x_3 - x_2}$		$f[x_1, x_2, x_3, x_4] = \dfrac{f[x_2, x_3, x_4] - f[x_1, x_2, x_3]}{x_4 - x_1}$
x_3	$f[x_3]$		$f[x_2, x_3, x_4] = \dfrac{f[x_3, x_4] - f[x_2, x_3]}{x_4 - x_2}$	
		$f[x_3, x_4] = \dfrac{f[x_4] - f[x_3]}{x_4 - x_3}$		$f[x_2, x_3, x_4, x_5] = \dfrac{f[x_3, x_4, x_5] - f[x_2, x_3, x_4]}{x_5 - x_2}$
x_4	$f[x_4]$		$f[x_3, x_4, x_5] = \dfrac{f[x_4, x_5] - f[x_3, x_4]}{x_5 - x_3}$	
		$f[x_4, x_5] = \dfrac{f[x_5] - f[x_4]}{x_5 - x_4}$		
x_5	$f[x_5]$			

Newton's Divided Difference Interpolating Polynomial

❑ Zeroth divided difference

$$f[x_i] = f(x_i).$$

at $\mathbf{x_0}$: $a_0 = P_n(x_0) = f(x_0).$

❑ First divided difference

$$f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i}.$$

at $\mathbf{x_1}$: $f(x_0) + a_1(x_1 - x_0) = P_n(x_1) = f(x_1)$

$$\Rightarrow a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

❑ Second divided difference

$$f[x_i, x_{i+1}, x_{i+2}] = \frac{f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]}{x_{i+2} - x_i}.$$

❑ kth divided difference

$$f[x_i, x_{i+1}, \ldots, x_{i+k}] = \frac{f[x_{i+1}, x_{i+2}, \ldots, x_{i+k}] - f[x_i, x_{i+1}, \ldots, x_{i+k-1}]}{x_{i+k} - x_i}.$$

$$\boldsymbol{f(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1)}$$

$$\boldsymbol{+ f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) \ldots\ldots}$$

Newton's Divided Difference Interpolating Polynomial

❑ What is the notation of divided differences

- Another form of Newton's Forward Difference Formula is constructed by using operator Δ.
- We now introduce the divided-difference notation, which is related to Aitken's Δ^2 notation

❑ First divided difference

$$\Delta f_0 = \frac{f_1 - f_0}{x_1 - x_0} \qquad \Delta f_1 = \frac{f_2 - f_1}{x_2 - x_1} \qquad \Delta f_2 = \frac{f_3 - f_2}{x_3 - x_2}$$

❑ Second divided difference

$$\Delta^2 f_0 = \frac{\Delta f_1 - \Delta f_0}{x_2 - x_0} \qquad \Delta^2 f_1 = \frac{\Delta f_2 - \Delta f_1}{x_3 - x_1} \qquad \Delta^2 f_2 = \frac{\Delta f_3 - \Delta f_2}{x_4 - x_2}$$

❑ Third divided difference

$$\Delta^3 f_0 = \frac{\Delta^2 f_1 - \Delta^2 f_0}{x_3 - x_0} \qquad \Delta^3 f_1 = \frac{\Delta^2 f_2 - \Delta^2 f_1}{x_4 - x_1}$$

$$\boldsymbol{f(x) = f_0 + \Delta f_0(x - x_0) + \Delta^2 f_0(x - x_0)(x - x_1) + \Delta^3 f_0(x - x_0)(x - x_1)(x - x_2) \ldots\ldots}$$

Newton's Divided Difference Interpolating Polynomial

Question: Using newton's divided difference formula to find f(x), given

$$x: \quad 0 \quad 2 \quad 3 \quad 6$$

$$f(x): 648 \quad 704 \quad 729 \quad 792$$

Also find $f(4)$ and $f'(4)$

Solution: The divided difference table for given data

X	$f = f(x$	First divided difference	Second divided difference	Third divided difference
0	648	$\Delta f_0 = \dfrac{704-648}{2-0} = 28$	$\Delta^2 f_0 = \dfrac{25-28}{3-0} = -1$	$\Delta^3 f_0 = 0$
2	704	$\Delta f_1 = \dfrac{729-704}{3-2} = 25$	$\Delta^2 f_1 = \dfrac{21-25}{6-2} = -1$	
3	729	$\Delta f_2 = \dfrac{792-729}{6-3} = 21$		
6	792			

Newton's Divided Difference Interpolating Polynomial

By newton's divided difference interpolation formula

$$\boldsymbol{f(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1)}$$
$$\boldsymbol{+ f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) \ldots\ldots}$$

OR

$$\boldsymbol{f(x) = f_0 + \Delta f_0(x - x_0) + \Delta^2 f_0(x - x_0)(x - x_1) + \Delta^3 f_0(x - x_0)(x - x_1)(x - x_2) \ldots\ldots}$$

$$f(x) = 648 + 28(x - 0) + (-1)(x - 0)(x - 2)$$

$$f(x) = -x^2 + 30x + 648$$

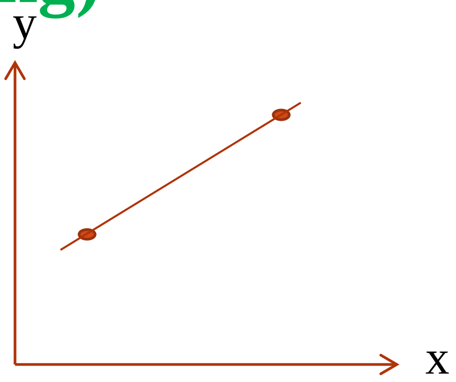
Now we should find the required $f(4)$ and $f'(4)$

$$f(4) = -4^2 + 30(4) + 648$$
$$f(4) = -16 + 120 + 648$$
$$f(4) = 752$$

$$f'(x) = -2x + 30$$
$$f'(4) = -2(4) + 30$$
$$f'(4) = 22$$

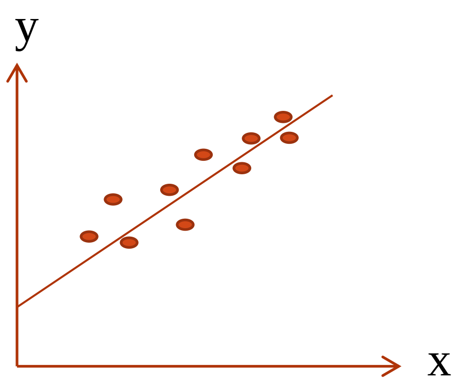
Linear Regression

Linear Regression (Linear Curve Fitting)



- **Curve fitting using a linear equation**

$$f(x) = ax + b$$

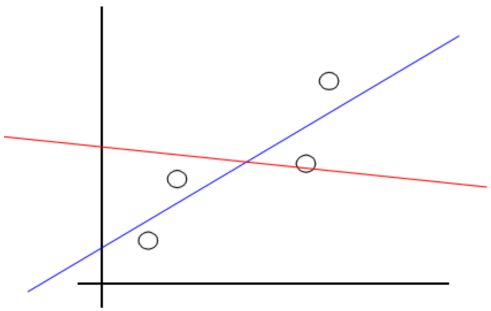


- With two data points
  - The constant can be determined that gives the exact values at the points.
- With more than two data points
  - Constant a and b are determined such that the line has the **best fit overall (?)**.
- What makes a particular straight line a 'good' fit?

Linear Regression

❑ Measuring how good is fit

- To quantify the overall agreement between the points and the function

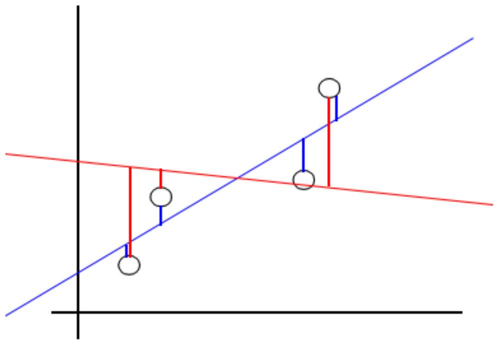


❑ Why does the blue line appear to us to fit the trend better?

❑ Method

- Calculate the error (**residual,** the difference between a data point and the predicted value)
- Calculate a total error using the residuals

$$r_i = y_i - f(x_i)$$

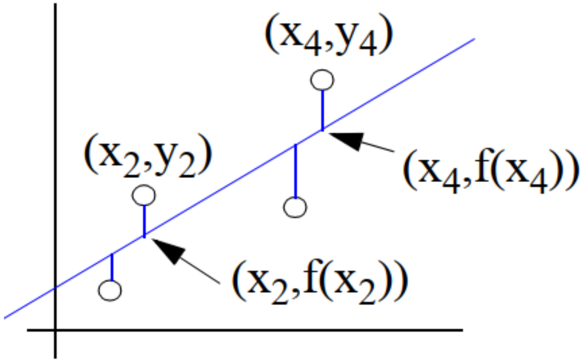


- The one line that provides a minimum error is then the 'best' straight line .

Linear Regression

❑ **Quantifying error in a curve fit**

- Both the positive and negative error have the same value.
- we can do both of these things by squaring the distance.



❑ Linear Regression

- A procedure to determine the coefficients a and b.
- Best fit
  - ○ The smallest possible total error calculated by adding the squares of the residuals.

$$err = \sum (d_i)^2 = (y_1 - f(x_1))^2 + (y_2 - f(x_2))^2 + (y_3 - f(x_3))^2 + (y_4 - f(x_4))^2$$

  - ○ Our fit is a straight line, so now substitute $f(x) = ax + b$

$$err = \sum_{i=1}^{\#\text{ data points}} (y_i - f(x_i))^2 = \sum_{i=1}^{\#\text{ data points}} (y_i - (ax_i + b))^2$$

Linear Regression

❑ Least Square approach

- This is called the **least squares approach**, since we minimize the square of the error.

$$\text{minimize} \quad err = \sum_{i=1}^{\#\text{ data points } = n} (y_i - (ax_i + b))^2$$

$$\frac{\partial err}{\partial a} = -2 \sum_{i=1}^{n} x_i(y_i - ax_i - b) = 0$$

$$\frac{\partial err}{\partial b} = -2 \sum_{i=1}^{n} (y_i - ax_i - b) = 0$$

❑ Solve for the a and b so that the previous two equations both = 0 and re-write these two equations .

$$a\sum x_i^2 + b\sum x_i = \sum (x_i y_i)$$

$$a\sum x_i + b*n = \sum y_i$$

Linear Regression

❑ Matrix form

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum (x_i y_i) \end{bmatrix}$$

- we have the data points (x_i, y_i) for i=1,2,….,n, so we have all the summation terms in the matrix, but we don't have a and b

- Find a and b using Gaussian elimination.

$$A = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}, \quad X = \begin{bmatrix} b \\ a \end{bmatrix}, \quad B = \begin{bmatrix} \sum y_i \\ \sum (x_i y_i) \end{bmatrix}$$

$$AX = B$$

- The coefficients a and b can be solved using matrix inversion.

$$X = A^{-1} * B$$

Linear Regression

Example: Fit a straight line f(x)=ax+b to the following data

i	1	2	3	4	5	6
x	0	0.5	1.0	1.5	2.0	2.5
y	0	1.5	3.0	4.5	6.0	7.5

Solution Solve for the a and b

$$a\sum x_i^2 + b\sum x_i = \sum (x_i y_i) \qquad a\sum x_i + b*n = \sum y_i$$

- First we find values for all the summation terms

$$n = 6$$

$$\sum x_i = 7.5, \quad \sum y_i = 22.5, \quad \sum x_i^2 = 13.75, \quad \sum x_i y_i = 41.25$$

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum (x_i y_i) \end{bmatrix}$$

Linear Regression

- Put these values in the matrix.

$$\begin{bmatrix} 6 & 7.5 \\ 7.5 & 13.75 \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 22.5 \\ 41.25 \end{bmatrix}$$

$$\begin{bmatrix} b \\ a \end{bmatrix} = inv \begin{bmatrix} 6 & 7.5 \\ 7.5 & 13.75 \end{bmatrix} * \begin{bmatrix} 22.5 \\ 41.25 \end{bmatrix}$$

$$X = A^{-1} * B$$

- The solution is

$$\begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

- Put these values in the straight line linear equation

$$f(x) = ax + b$$

$$f(x) = 3x + 0$$

Linear Regression

Example: Fit a straight line f(x)=ax+b to the following data

x = [0 .5 1 1.5 2 2.5], y = [-0.4326 -0.1656 3.1253 4.7877 4.8535 8.6909]

Solution Solve for the a and b

$$a\sum x_i + b*n = \sum y_i \longrightarrow \text{Eq (1)}$$

$$a\sum x_i^2 + b\sum x_i = \sum (x_i y_i) \longrightarrow \text{Eq (2)}$$

x_i	y_i	x_i^2	$x_i y_i$
0	-0.4326	0	0
0.5	-0.1656	0.25	-0.0828
1	3.1253	1	3.1253
1.5	4.7877	2.25	7.18155
2	4.8535	4	9.707
2.5	8.6909	6.25	21.72725
$\sum x_i = 7.5$	$\sum y_i = 20.8593$	$\sum x_i^2 = 13.75$	$\sum x_i y_i = 41.7411$

Linear Regression

Put the last row values in equation 1 and 2

$$a\sum x_i + b*n = \sum y_i$$

$$7.5a + 6b = 20.8593$$

$$a = \frac{20.8593 - 6b}{7.5} \longrightarrow \text{Eq (3)}$$

$$a\sum x_i^2 + b\sum x_i = \sum (x_i y_i)$$

$$13.75a + 7.5b = 41.7411$$

$$b = \frac{41.7411 - 13.75a}{7.5} \longrightarrow \text{Eq (4)}$$

Put the value of b in equation 3

$$a = \frac{20.8593 - 6\,\dfrac{41.7411 - 13.75a}{7.5}}{7.5}$$

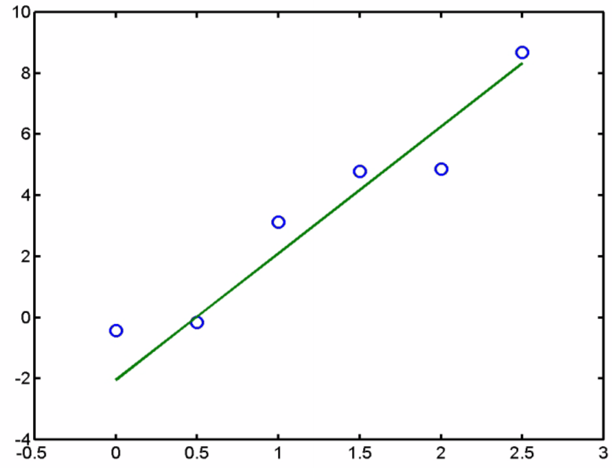
$$a = 3.581$$

Put the value of a in equation 4

$$b = \frac{41.7411 - 13.75(3.581)}{7.5}$$

$$b = -0.999$$

$$f(x) = 3.581x - 0.999$$



Thank You !