Unit 2

Interpolation and Approximation

2.1 Interpolation

Suppose we are given a set of $n+1$ data points in the tabular form as below:

x	x_0	x_1	$\cdots$	x_n
f	f_0	f_1	$\cdots$	f_n

Then **interpolation** is the method of finding a function $f(x)$, called an **interpolation function**, such that $f(x_i) = f_i$, $0 \le i \le n$ and estimating the value of f by $f(x)$ for some x lying in between $x_0, x_1, \ldots, x_n$. If the estimation is done for some x lying outside the range of $x_0, \cdots, x_n$, then we call it **extrapolation** .

For example, suppose we are given the following table of data points:

x	0	1	3
f	1	3	2

Then
$$f(x) = \frac{-5}{6}x^2 + \frac{17}{6}x + 1$$
is an interpolation function for this set of data points because $f(0) = 1$, $f(1) = 3$ and $f(3) = 2$. If $x = 2$, then the estimated value of corresponding f using above interpolation function is
$$f(2) = \frac{-5}{6} \times 2^2 + \frac{17}{6} \times 2 + 1 = \frac{10}{3}.$$

Given a set of data points, we can use different types of functions to interpolate those points. We can use polynomial functions to interpolate the given set of data points as in the example above. In some cases, we may use exponential function or trigonometric function as well.

Polynomial Interpolation:

Given a set of data points (x_i, f_i), $i = 0, 1, \ldots, n$, the process of finding a polynomial function that interpolates these data points is called **polynomial interpolation** .

Polynomials are a very popular choice for interpolating any set of data points. These can easily be differentiated and integrated which again results in polynomials. Moreover, there is a famous theorem by Weierstrass which says that a continuous function defined on a closed interval can be approximated as accurately as we like by a polynomial of sufficiently high degree. More precisely

Weierstrass Approximation Theorem: Suppose f is a continuous function defined on the interval $[a, b]$. Then for any $\varepsilon > 0$, there exists a polynomial $p(x)$ such that

$$|f(x) - p(x)| < \varepsilon$$

for all $x \in [a, b]$.

There are various methods of polynomial interpolation:

1. Lagrange's interpolation
2. Newton's interpolation
3. Cubic spline interpolation

2.2 Lagrange's Interpolation

Suppose we are given a set of $n + 1$ data points (x_i, f_i), $i = 0, 1, \ldots, n$. Using Lagrange's interpolation method, we obtain a polynomial of degree n that interpolates these data points. To derive this polynomial, we suppose for simplicity that we are given only three data points (x_i, f_i), $i = 0, 1, 2$. We want to interpolate these points by a polynomial function of degree 2. Let that polynomial be

$$P_2(x) = b_1(x - x_0)(x - x_1) + b_2(x - x_1)(x - x_2) + b_3(x - x_0)(x - x_2) \cdots\cdots (A)$$

We need to determine the values of b_1, b_2 and b_3. Since $P_2(x)$ interpolates (x_i, f_i), $i = 0, 1, 2$, so we have,

$$P_2(x_0) = f_0 \text{ which implies } b_2(x_0 - x_1)(x_0 - x_2) = f_0 \cdots\cdots (1)$$

$$P_2(x_1) = f_1 \text{ which implies } b_3(x_1 - x_0)(x_1 - x_2) = f_1 \cdots\cdots (2)$$

$$P_2(x_2) = f_2 \text{ which implies } b_1(x_2 - x_1)(x_2 - x_0) = f_2 \cdots\cdots (3)$$

From (1), (2) and (3), we obtain

$$b_2 = \frac{f_0}{(x_0 - x_1)(x_0 - x_2)},$$

$$b_3 = \frac{f_1}{(x_1 - x_0)(x_1 - x_2)}$$

and

$$b_1 = \frac{f_2}{(x_2 - x_0)(x_2 - x_1)}.$$

Substituting these in equation (A), we get

$$P_2(x) = f_2\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} + f_0\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + f_1\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} \cdots\cdots (B)$$

Equation (B) may be represented as,

$$P_2(x) = f_0 l_0(x) + f_1 l_1(x) + f_2 l_2(x) = \sum_{i=0}^{2} f_i l_i(x)$$

where $l_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{2} \frac{(x-x_j)}{(x_i-x_j)}$.

Generalizing this formula for $n+1$ data points (x_i, f_i), $i = 0, 1, \ldots, n$, we obtain an n degree polynomial as

$$P_n(x) = \sum_{i=0}^{n} f_i l_i(x) \cdots\cdots (C)$$

where

$$l_i(x) = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{(x-x_j)}{(x_i-x_j)} \cdots\cdots (D)$$

The polynomial $P_n(x)$ given by (C) is the required interpolation polynomial and is called the Lagrange's interpolation polynomial . The polynomials $l_i(x)$, $i = 0, 1, \ldots, n$ represented by (D) are called Lagrange's basis polynomials .

Algorithm (Lagrange's interpolation polynomial):

INPUT: A set of $n+1$ data points (x_i, f_i), $i = 0, 1, \ldots, n$.

PROCESS:

SET $lpoly = 0$

FOR $i = 0$ TO n {

SET $w = 1$

FOR $j = 0$ TO n {

IF $j \neq i$ SET $w = w \times \frac{(x-x_j)}{(x_i-x_j)}$

}

SET $lpoly = lpoly + w \times f_i$

}

OUTPUT: Lagrange's interpolation polynomial $lpoly$.

Rolle's Theorem: Let f be a function that is continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b) = 0$, then $f'(c) = 0$ for some point $c \in (a, b)$.

Theorem (Error in polynomial interpolation): If $P_n(x)$ is the polynomial of degree at most n that interpolates f at the $n+1$ distinct nodes $x_0, x_1, \cdots, x_n \in [a, b]$ and if $f^{(n+1)}$ exists and is continuous, then for each $x \in [a, b]$, there exists $r \in (a, b)$ such that

$$f(x) - P_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(r)\omega(x)$$

where $\omega(x) = (x - x_0)(x - x_1) \cdots (x - x_n)$.

Proof: If $x = x_i$ for any $i = 0, 1, \ldots, n$ then $f(x_i) = f_i = P_n(x_i)$ and $\omega(x_i) = 0$ and so the given equation obviously holds. Suppose now that $x \neq x_i$ for any $i = 0, 1, \ldots, n$. We define a function $\phi(t)$ by

$$\phi(t) = f(t) - P_n(t) - c\omega(t)$$

where c is a constant chosen in such a way that $\phi(x) = 0$. Then $\phi(t)$ has at least $n+2$ zeros namely $x_0, x_1, \cdots, x_n, x$ in the interval $[a, b]$, i.e.,

$$\phi(x_0) = \phi(x_1) = \cdots = \phi(x_n) = \phi(x) = 0.$$

By Rolle's theorem, between any two roots of ϕ, there must exist a root of ϕ'. So ϕ' has at least $n+1$ roots. By similar reasoning, ϕ'' has at least n roots, ϕ''' has at least $n-1$ roots and finally, $\phi^{(n+1)}$ must have at least one root, say r, in (a, b). So

$$\phi^{(n+1)}(r) = 0$$

$$\Rightarrow f^{(n+1)}(r) - P_n^{(n+1)}(r) - c\omega^{(n+1)}(r) = 0.$$

Since $P_n(x)$ is a polynomial of degree at most n, so $P_n^{(n+1)}(r) = 0$. Also, $\omega^{(n+1)}(r) = (n+1)!$. So we have,

$$f^{(n+1)}(r) - c(n+1)! = 0$$

$$\Rightarrow c = \frac{1}{(n+1)!} f^{(n+1)}(r).$$

Therefore,

$$\phi(x) = 0$$

$$\Rightarrow f(x) - P_n(x) - c\omega(x) = 0$$

$$\Rightarrow f(x) - P_n(x) = c\omega(x)$$

$$\Rightarrow f(x) - P_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(r)\omega(x).$$

■

Disadvantages of using Lagrange's method of interpolation:

1. This method requires $2(n+1)$ multiplications and divisions and $2n+1$ additions and subtractions. So it involves more arithmetic operations.

2. If we want to add or remove one or more data points, then we have to compute the polynomial from the very beginning, i.e., it cannot use the polynomial $P_n(x)$ to compute the polynomial $P_{n+1}(x)$.

2.2.1 Exercise

1. Find Lagrange's interpolating polynomial and the value of f at $x = \frac{1}{2}$ for the following data points:

x	$\frac{1}{3}$	$\frac{1}{4}$	1
f	2	-1	7

Solution: We have,

$$l_0(x) = \prod_{j=0, j\neq 0}^{2} \frac{(x - x_j)}{(x_0 - x_j)} = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - \frac{1}{4})(x - 1)}{(\frac{1}{3} - \frac{1}{4})(\frac{1}{3} - 1)} = -18\left(x - \frac{1}{4}\right)(x-1)$$

$$l_1(x) = \prod_{j=0, j\neq 1}^{2} \frac{(x - x_j)}{(x_1 - x_j)} = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = 16\left(x - \frac{1}{3}\right)(x - 1)$$

$$l_2(x) = \prod_{j=0, j\neq 2}^{2} \frac{(x - x_j)}{(x_2 - x_j)} = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = 2\left(x - \frac{1}{3}\right)\left(x - \frac{1}{4}\right)$$

Therefore, the Lagrange's interpolating polynomial is

$$\begin{aligned} P_2(x) &= f_0 l_0(x) + f_1 l_1(x) + f_2 l_2(x) \\ &= -36\left(x - \frac{1}{4}\right)(x - 1) - 16\left(x - \frac{1}{3}\right)(x - 1) + 14\left(x - \frac{1}{3}\right)\left(x - \frac{1}{4}\right) \\ &= -38x^2 + \frac{349}{6}x - \frac{79}{6} \end{aligned}$$

The value of f at $x = \frac{1}{2}$ is $P_2\left(\frac{1}{2}\right) = \frac{77}{12}$.

2. Derive the equation for Lagrange's interpolating polynomial and find the value of $f(x)$ at $x = 0$ for following function:

x	-1	-2	2	4
$f(x)$	-1	-9	11	69

Solution: We have,

$$\begin{aligned} l_0(x) &= \prod_{j=0, j\neq 0}^{3} \frac{(x - x_j)}{(x_0 - x_j)} = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} \\ &= \frac{(x + 2)(x - 2)(x - 4)}{(-1 + 2)(-1 - 2)(-1 - 4)} = \frac{1}{15}(x - 4)(x^2 - 4) \end{aligned}$$

$$\begin{aligned} l_1(x) &= \prod_{j=0, j\neq 1}^{3} \frac{(x - x_j)}{(x_1 - x_j)} = \frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} \\ &= \frac{(x + 1)(x - 2)(x - 4)}{(-2 + 1)(-2 - 2)(-2 - 4)} = \frac{-1}{24}(x + 1)(x - 2)(x - 4) \end{aligned}$$

$$l_2(x) = \prod_{j=0,j\neq 2}^{3} \frac{(x-x_j)}{(x_2-x_j)} = \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}$$
$$= \frac{(x+1)(x+2)(x-4)}{(2+1)(2+2)(2-4)} = \frac{-1}{24}(x+1)(x+2)(x-4)$$
$$l_3(x) = \prod_{j=0,j\neq 3}^{3} \frac{(x-x_j)}{(x_0-x_j)} = \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}$$
$$= \frac{(x+1)(x+2)(x-2)}{(4+2)(4+2)(4-2)} = \frac{1}{60}(x+1)(x^2-4)$$

So the Lagrange's interpolating polynomial is,

$$\begin{aligned} P_3(x) &= f_0l_0(x) + f_1l_1(x) + f_2l_2(x) + f_3l_3(x) \\ &= \frac{-1}{15}(x-4)(x^2-4) + \frac{9}{24}(x+1)(x-2)(x-4) - \frac{11}{24}(x+1)(x+2)(x-4) \\ &\quad + \frac{69}{60}(x+1)(x^2-4) \\ &= x^3 + x + 1. \end{aligned}$$

The value of f at $x = 0$ is $P_3(0) = 1$.

3. Use Lagrange's interpolation method to find the interpolating polynomial for the following data:

x	0	1	3
f	-12	0	12

What would be the interpolating polynomial if $(4, 24)$ is added to the table?

4. Use Lagrange's interpolation method to find the interpolating polynomial for the following data:

x	1	3	4	6
f	-3	9	30	132

5. Use Lagrange's interpolation method to find the interpolating polynomial for the following data:

x	-2	-1	2	3
f	-12	-8	3	5

6. Use the numbers $x_0 = 2$, $x_1 = 2.75$ and $x_2 = 4$ to find the Lagrange interpolating polynomial for $f(x) = \dfrac{1}{x}$. Use this polynomial to approximate $f(3) = \dfrac{1}{3}$.

2.3 Newton's Interpolation

Divided differences:

Suppose we are given a set of $n+1$ data points $(x_0, f_0), \ldots, (x_n, f_n)$. We define

$$\begin{aligned} f[x_i] &= f_i \\ f[x_i, x_j] &= \frac{f[x_j] - f[x_i]}{x_j - x_i} \\ f[x_i, x_j, x_k] &= \frac{f[x_j, x_k] - f[x_i, x_j]}{x_k - x_i} \\ &\vdots \\ f[x_i, x_j, \ldots, x_l, x_m] &= \frac{f[x_j, \ldots, x_l, x_m] - f[x_i, x_j, \ldots, x_l]}{x_m - x_i} \end{aligned}$$

and so on. These quantities are called the divided differences .

Newton's Interpolation Polynomial:

Suppose we are given a set of $n+1$ data points (x_i, f_i), $i = 0, 1, \ldots, n$. Then Newton's interpolation polynomial for interpolating these data points is of the form

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

To compute this polynomial, we need to determine the values of the coefficients $a_0, a_1, \cdots, a_n$ in (A). For this, we note that since (A) is an interpolating polynomial, we must have $P_n(x_i) = f_i$ for $i = 0, 1, 2, \cdots, n$. So for $i = 0$ we have

$$\begin{aligned} P_n(x_0) &= f_0 \\ \Rightarrow a_0 &= f_0 = f[x_0]. \end{aligned}$$

For $i = 1$ we have

$$\begin{aligned} P_n(x_1) = f_1 \Rightarrow a_0 + a_1(x_1 - x_0) = f_1 \Rightarrow f_0 + a_1(x_1 - x_0) = f_1 \\ \Rightarrow a_1 = \frac{f_1 - f_0}{x_1 - x_0} = \frac{f[x_1] - f[x_0]}{x_1 - x_0} = f[x_0, x_1]. \end{aligned}$$

For $i = 2$ we have

$$\begin{aligned} P_n(x_2) = f_2 \Rightarrow a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1) = f_2 \\ \Rightarrow a_2 = \frac{\dfrac{f_2 - f_1}{x_2 - x_1} - \dfrac{f_1 - f_0}{x_1 - x_0}}{x_2 - x_0} = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = f[x_0, x_1, x_2]. \end{aligned}$$

Continuing similarly, we get $a_i = f[x_0, x_1, \cdots, x_i]$ for $i = 0, 1, \ldots, n$. Substituting the value of a_i's in (A), we get,

$$\begin{aligned} P_n(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) + \cdots \\ \cdots + f[x_0, x_1, \cdots, x_n](x - x_0)(x - x_1)\cdots(x - x_{n-1}) \cdots\cdots (B) \end{aligned}$$

Equation (B) is known as the Newton's (divided differences) interpolation polynomial .

2.3.1 Exercise

1. Find Newton's interpolating polynomial for the following data points:

x	$\frac{1}{3}$	$\frac{1}{4}$	1
f	2	-1	7

Solution: We have,

$$f[x_0] = f_0 = 2$$

$$f[x_0, x_1] = \frac{f_1 - f_0}{x_1 - x_0} = \frac{-1-2}{\frac{1}{4} - \frac{1}{3}} = -3 \times -12 = 36$$

$$f[x_1, x_2] = \frac{f_2 - f_1}{x_2 - x_1} = \frac{7-(-1)}{1 - \frac{1}{4}} = 8 \times \frac{4}{3} = \frac{32}{3}$$

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{\frac{32}{3} - 36}{1 - \frac{1}{3}} = \frac{32 - 108}{3} \times \frac{3}{2} = -38$$

Therefore, the Newton's interpolating polynomial is

$$\begin{aligned} P_2(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\ &= 2 + 36\left(x - \frac{1}{3}\right) - 38\left(x - \frac{1}{3}\right)\left(x - \frac{1}{4}\right) \\ &= -38x^2 + \frac{349}{6}x - \frac{79}{6}. \end{aligned}$$

2. Derive the Newton's interpolating polynomial for the following data points:

x	-1	-2	2
$f(x)$	-1	-9	11

What would be the Newton's interpolating polynomial if we add an additional data point $(4, 69)$?

Solution: We have,

$$f[x_0] = f_0 = -1$$

$$f[x_0, x_1] = \frac{f_1 - f_0}{x_1 - x_0} = \frac{-9-(-1)}{-2-(-1)} = 8$$

$$f[x_1, x_2] = \frac{f_2 - f_1}{x_2 - x_1} = \frac{11+9}{2-(-2)} = 5$$

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{5-8}{2-(-1)} = -1$$

Therefore, Newton's interpolating polynomial is

$$\begin{aligned}
P_2(x) &= f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\
&= -1 + 8(x + 1) - (x + 1)(x + 2) \\
&= -x^2 + 5x + 5
\end{aligned}$$

If we now add data point $(4, 69)$ as well, then the Newton's interpolating polynomial is

$$\begin{aligned}
P_3(x) = f[x_0] + f[x_0, x_1](x - x_0) &+ f[x_0, x_1, x_2](x - x_0)(x - x_1) \\
&+ f[x_0, x_1, x_2, x_3](x - x_0)(x - x_1)(x - x_2) \\
= P_2(x) + f[x_0, x_1, x_2, x_3]&(x - x_0)(x - x_1)(x - x_2)
\end{aligned}$$

Now,

$$\begin{aligned}
f[x_2, x_3] &= \frac{f_3 - f_2}{x_3 - x_2} = \frac{69 - 11}{4 - 2} = 29 \\
f[x_1, x_2, x_3] &= \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} = \frac{29 - 5}{4 - (-2)} = 4 \\
f[x_0, x_1, x_2, x_3] &= \frac{f[x_1, x_2, x_3] - f[x_0, x_1, x_2]}{x_3 - x_0} = \frac{4 - (-1)}{4 - (-1)} = 1
\end{aligned}$$

Therefore,

$$P_3(x) = -x^2 + 5x + 5 + (x + 1)(x + 2)(x - 2) = x^3 + x + 1.$$

3. Use Newton's method to find the interpolating polynomial for the following data:

x	0	1	3	4
$f(x)$	-12	0	12	24

What would be the interpolating polynomial if $(2, 9)$ is added to the table?

4. Use Newton's interpolation method to find the interpolating polynomial for the following data:

x	1	3	4	6
f	-3	9	30	132

5. Use Newton's interpolation method to find the interpolating polynomial for the following data:

x	-2	-1	2	3
f	-12	-8	3	5

2.3.2 Divided Difference Table

Given a set of $n+1$ data points (x_i, f_i), $i = 0, 1, \ldots, n$, its divided differences table lists all the possible divided differences of all orders as follows:

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$\cdots\cdots$	$f[x_i, \cdots, x_{i+n-1}]$	$f[x_0, \cdots, x_n]$
x_0	$f[x_0]$				
		$f[x_0, x_1]$			
x_1	$f[x_1]$				
		$f[x_1, x_2]$			
x_2	$f[x_2]$			$f[x_0, \cdots, x_{n-1}]$	
$\vdots$	$\vdots$	$\vdots$			$f[x_0, \cdots, x_n]$
$\vdots$	$\vdots$	$\vdots$		$f[x_1, \cdots, x_n]$	
x_{n-1}	$f[x_{n-1}]$				
		$f[x_{n-1}, x_n]$			
x_n	$f[x_n]$				

Note that to calculate Newton's interpolation polynomial, we only need the divided differences forming the top diagonal elements of the divided differences table.

Algorithm (Divided differences table):

INPUT: A set of $n+1$ data points (x_i, f_i), $i = 0, 1, \ldots, n$.

PROCESS:

FOR $i = 1$ TO n {

FOR $j = 0$ TO $n - i$ {

Compute

$$f[x_j, x_{j+1}, \cdots, x_{j+i}] = \frac{f[x_{j+1}, \cdots, x_{j+i}] - f[x_j, \cdots, x_{j+i-1}]}{x_{j+i} - x_j}$$

and enter into column i of the table

}

}

OUTPUT: Divided difference table

Exercise

1. Construct the divided differences table and Newton's interpolation polynomial for the following data points:

x	0	1	3
f	1	3	2

Solution: The divided differences table for the given data is as follows:

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$
0	1		
		2	
1	3		$-\frac{5}{6}$
		$-\frac{1}{2}$	
3	2		

Therefore

$$P_2(x) = 1 + 2(x-0) - \frac{5}{6}(x-0)(x-1) = \frac{-5}{6}x^2 + \frac{17}{6}x + 1.$$

2. Construct the divided differences table from the following data:

x	3.2	2.7	1.0	4.8	5.6
y	22.0	17.8	14.2	38.3	51.7

Solution:

x_i	$f[x_i]$	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$	$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}, x_{i+4}]$
3.2	22				
		8.4			
2.7	17.8		2.85		
		2.12		-0.52	
1.0	14.2		2.01		0.25
		6.34		0.09	
4.8	38.3		2.26		
		16.75			
5.6	51.7				

3. Construct the divided differences table and Newton's interpolation polynomial for the following data points:

x	-1	0	3	6	7
f	3	-6	39	822	1611

4. Construct the divided differences table and Newton's interpolation polynomial for the following data points:

x	1	3	4	6
f	-3	9	30	132

5. Construct the divided differences table and Newton's interpolation polynomial for the following data points:

x	-2	-1	2	3
f	-12	-8	3	5

Ans: $-\dfrac{1}{15}x^3 - \dfrac{3}{20}x^2 + \dfrac{241}{60}x - 3.9$

6. Construct the divided differences table and Newton's interpolation polynomial for the following data points:

x	0	1	3	4	2
$f(x)$	-12	0	12	24	9

7. Construct the divided differences table and Newton's interpolation polynomial for the following data points:

x	-2	-1	0	1	2	3
$f(x)$	-5	1	1	1	7	25

2.3.3 Interpolating evenly-spaced data

Newton's Forward Difference Interpolation Formula:

Suppose we are given $n+1$ pairs of data points (x_i, f_i), $i = 0, 1, \ldots, n$ but this time we assume that $x_0, x_1, \cdots, x_n$ are equal distance h apart, i.e., $x_k = x_0 + kh$, $k = 1, 2, \ldots, n$.
Now for each $i = 0, 1, \ldots, n-1$, let us define

$$\begin{aligned}
\Delta f_i &= f_{i+1} - f_i \\
\Delta^2 f_i &= \Delta f_{i+1} - \Delta f_i \\
\Delta^3 f_i &= \Delta^2 f_{i+1} - \Delta^2 f_i
\end{aligned}$$

We call these quantities the first forward difference, second forward difference and third forward difference respectively. In general,

$$\Delta^j f_i = \Delta^{j-1} f_{i+1} - \Delta^{j-1} f_i$$

is called the j^{th}-forward difference.

We now express the divided differences in terms of forward differences. We have,

$$f[x_0, x_1] = \frac{f_1 - f_0}{x_1 - x_0} = \frac{\Delta f_0}{h}$$

Also,

$$f[x_1, x_2] = \frac{f_2 - f_1}{x_2 - x_1} = \frac{\Delta f_1}{h}$$

Now,

$$f[x_0,x_1,x_2] = \frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0} = \frac{\frac{\Delta f_1}{h}-\frac{\Delta f_0}{h}}{2h}$$
$$= \frac{\Delta f_1-\Delta f_0}{2h^2} = \frac{\Delta^2 f_0}{2h^2}$$

Continuing this way, we get in general

$$f[x_0,x_1,\cdots,x_j] = \frac{\Delta^j f_0}{j!h^j}$$

Substituting the value of these divided differences in Newton's interpolation polynomial, we get,

$$\begin{aligned} P_n(x) &= f_0 + \frac{\Delta f_0}{h}(x-x_0) + \frac{\Delta^2 f_0}{2h^2}(x-x_0)(x-x_1)+\cdots \\ &\qquad \cdots + \frac{\Delta^n f_0}{n!h^n}(x-x_0)(x-x_1)\cdots(x-x_{n-1}) \\ &= f_0 + \Delta f_0\left(\frac{x-x_0}{h}\right) + \frac{\Delta^2 f_0}{2}\left(\frac{x-x_0}{h}\right)\left(\frac{x-x_1}{h}\right)+\cdots \\ &\qquad \cdots + \frac{\Delta^n f_0}{n!}\left(\frac{x-x_0}{h}\right)\left(\frac{x-x_1}{h}\right)\cdots\left(\frac{x-x_{n-1}}{h}\right). \end{aligned}$$

Let $s=\frac{x-x_0}{h}$. Then for $j=1,2,\dots,n-1$,

$$\frac{x-x_j}{h} = \frac{x-(x_0+jh)}{h} = \frac{x-x_0-jh}{h} = s-j.$$

Thus,

$$P_n(x) = f_0 + s\Delta f_0 + s(s-1)\frac{\Delta^2 f_0}{2!} + \cdots + s(s-1)\cdots(s-n+1)\frac{\Delta^n f_0}{n!}$$

where $s=\frac{x-x_0}{h}$.

This is known as Gregory-Newton forward difference interpolation formula which is used for interpolating evenly-spaced data.

Exercise

1. Construct the forward difference table for following data:

x	0.1	0.3	0.5	0.7	0.9	1.1	1.3
$f(x)$	0.003	0.067	0.148	0.248	0.370	0.518	0.697

Find $f(0.6)$ using Newton-Gregory forward difference interpolation formula.

Solution: The forward difference table for the given data is

x	f	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$	$\Delta^5 f$	$\Delta^6 f$
0.1	0.003						
		0.064					
0.3	0.067		0.017				
		0.081		0.002			
0.5	0.148		0.019		0.001		
		0.100		0.003		0	
0.7	0.248		0.022		0.001		0
		0.122		0.004		0	
0.9	0.370		0.026		0.001		
		0.148		0.005			
1.1	0.518		0.031				
		0.179					
1.3	0.697						

To find $f(0.6)$, we have $x_0 = 0.1$, $h = 0.2$, $x = 0.6$, so $s = \dfrac{x - x_0}{h} = \dfrac{0.6 - 0.1}{0.2} = 2.5$.
Also $f_0 = 0.003$, $\Delta f_0 = 0.064$, $\Delta^2 f_0 = 0.017$, $\Delta^3 f_0 = 0.002$, $\Delta^4 f_0 = 0.001$, $\Delta^5 f_0 = 0$ and $\Delta^6 f_0 = 0$. Therefore the estimated value of $f(0.6)$ using Newton-Gregory forward difference interpolation polynomial is

$$\begin{aligned}
P_6(0.6) &= f_0 + s\Delta f_0 + s(s-1)\frac{\Delta^2 f_0}{2!} + s(s-1)(s-2)\frac{\Delta^3 f_0}{3!} + s(s-1)(s-2)(s-3)\frac{\Delta^4 f_0}{4!} \\
&\quad + s(s-1)(s-2)(s-3)(s-4)\frac{\Delta^5 f_0}{5!} + s(s-1)(s-2)(s-3)(s-4)(s-5)\frac{\Delta^6 f_0}{6!} \\
&= 0.003 + 2.5 \times 0.064 + 2.5(2.5-1)\frac{0.017}{2} + 2.5(2.5-1)(2.5-2)\frac{0.002}{6} \\
&\quad + 2.5(2.5-1)(2.5-2)(2.5-3)\frac{0.001}{24} + 0 + 0 \\
&= 0.19546
\end{aligned}$$

2. Construct a forward difference table and find the Newton-Gregory forward difference interpolation formula for the following table of data points. Then estimate the value of $\sin\theta$ at $\theta = 25°$.

θ	10	20	30	40	50
$\sin\theta$	0.1736	0.3420	0.5	0.6428	0.7660

Solution: The forward difference table for the given data is as follows:

θ	$\sin\theta$	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$
10	0.1736				
		0.1684			
20	0.3420		−0.0104		
		0.1580		−0.0048	
30	0.5		−0.0152		0.0004
		0.1428		−0.0044	
40	0.6428		−0.0196		
		0.1232			
50	0.7660				

Now, from the given data, we have $\theta_0 = 10$ and $h = 10$. So $s = \dfrac{\theta - \theta_0}{h} = \dfrac{\theta - 10}{10}$.
Also, from the above forward difference table, $f_0 = 0.1736$, $\Delta f_0 = 0.1684$, $\Delta^2 f_0 = -0.0104$, $\Delta^3 f_0 = 0.0048$ and $\Delta^4 f_0 = -0.0004$. So the Newton-Gregory forward difference interpolation polynomial is

$$P_4(\theta) = 0.1736 + 0.1684\left(\frac{\theta-10}{10}\right) + \frac{-0.0104}{2!}\left(\frac{\theta-10}{10}\right)\left(\frac{\theta-10}{10}-1\right) + \frac{-0.0048}{3!}\left(\frac{\theta-10}{10}\right)\left(\frac{\theta-10}{10}-1\right)\left(\frac{\theta-10}{10}-2\right) + \frac{0.0004}{4!}\left(\frac{\theta-10}{10}\right)\left(\frac{\theta-10}{10}-1\right)\left(\frac{\theta-10}{10}-2\right)\left(\frac{\theta-10}{10}-3\right)$$

Using this polynomial, the value of $\sin\theta$ at $\theta = 25^\circ$ is estimated to be

$$\begin{aligned} P_4(25) &= 0.1736 + 0.1684 \times \frac{3}{2} - \frac{0.0104}{2} \times \frac{3}{2} \times \frac{1}{2} + \frac{-0.0048}{6} \times \frac{3}{2} \times \frac{1}{2} \times \frac{-1}{2} \\ &+ \frac{0.0004}{24} \times \frac{3}{2} \times \frac{1}{2} \times \frac{-1}{2} \times \frac{-3}{2} \\ &= 0.1736 + 0.2526 - 0.0039 + 0.0003 + 0.000009375 = 0.4226 \end{aligned}$$

3. Construct the forward difference table for following data:

x	0.0	0.2	0.4	0.6	0.8	1.0	1.2
$f(x)$	0.000	0.203	0.423	0.684	1.030	1.557	2.572

Find $f(0.73)$ using Newton-Gregory forward difference interpolation formula.

4. Construct the forward difference table for following data:

x	0.61	0.62	0.63	0.64	0.65	0.66	0.67
$f(x)$	1.840431	1.858928	1.877610	1.896481	1.915541	1.934792	1.954237

Find $f(0.644)$ using Newton-Gregory forward difference interpolation formula.

5. Construct the forward difference table for following data:

x	1.72	1.73	1.74	1.75	1.76	1.77	1.78
$f(x)$	17.9066	17.7284	17.5520	17.3773	17.2044	17.0332	16.8638

Find $f(1.746)$ using Newton-Gregory forward difference interpolation formula.

Newton's Backward Difference Interpolation Formula:

Given a set of $n+1$ data points (x_i, f_i), $i = 0, 1, \ldots, n$, we take the interpolation polynomial $P_n(x)$ of degree n to be of the form

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

To determine the values of the coefficients $a_0, a_1, \cdots, a_n$ in (A), we note that since (A) is an interpolating polynomial, we must have $P_n(x_i) = f_i$ for $i = 0, 1, 2, \cdots, n$. Using these conditions, we get,

$$a_0 = f[x_n], a_1 = f[x_n, x_{n-1}], \ldots, a_n = f[x_n, x_{n-1}, \ldots, x_0]$$

Substituting the value of a_i's in (A), we get,

$$\begin{aligned} P_n(x) = f[x_n] + f[x_n, x_{n-1}](x-x_n) + f[x_n, x_{n-1}, x_{n-2}](x-x_n)(x-x_{n-1}) + \cdots \\ \cdots + f[x_n, x_{n-1}, \cdots, x_0](x-x_n)(x-x_{n-1})\cdots(x-x_1) \cdots\cdots (B) \end{aligned}$$

We now assume that $x_n, x_{n-1}, \cdots, x_0$ are equal distance h apart, i.e., $x_k = x_n - (n-k)h$, $k = 0, 1, \ldots, n-1$.
Now for each $i = 1, \ldots, n$, let us define

$$\begin{aligned} \nabla f_i &= f_i - f_{i-1} \\ \nabla^2 f_i &= \nabla f_i - \nabla f_{i-1} \\ \nabla^3 f_i &= \nabla^2 f_i - \nabla^2 f_{i-1} \end{aligned}$$

We call these quantities the first backward difference, second backward difference and third backward difference respectively. In general,

$$\nabla^j f_i = \nabla^{j-1} f_i - \nabla^{j-1} f_{i-1}$$

is called the j^{th}-backward difference.

We now express the divided differences in terms of backward differences. We have,

$$f[x_n, x_{n-1}] = \frac{f_{n-1} - f_n}{x_{n-1} - x_n} = \frac{f_n - f_{n-1}}{x_n - x_{n-1}} = \frac{\nabla f_n}{h}$$

Also,

$$f[x_{n-1}, x_{n-2}] = \frac{f_{n-2} - f_{n-1}}{x_{n-2} - x_{n-1}} = \frac{f_{n-1} - f_{n-2}}{x_{n-1} - x_{n-2}} = \frac{\nabla f_{n-1}}{h}$$

Now,

$$f[x_n, x_{n-1}, x_{n-2}] = \frac{f[x_{n-1}, x_{n-2}] - f[x_n, x_{n-1}]}{x_{n-2} - x_n} = \frac{f[x_n, x_{n-1}] - f[x_{n-1}, x_{n-2}]}{x_n - x_{n-2}}$$

$$= \frac{\dfrac{\nabla f_n}{h} - \dfrac{\nabla f_{n-1}}{h}}{2h} = \frac{\nabla f_n - \nabla f_{n-1}}{2h^2} = \frac{\nabla^2 f_n}{2h^2}$$

Continuing this way, we get in general

$$f[x_n, x_{n-1}, \cdots, x_{n-j}] = \frac{\nabla^j f_n}{j!h^j}$$

Substituting the value of these divided differences in (B), we get,

$$\begin{aligned}P_n(x) &= f_n + \frac{\nabla f_n}{h}(x - x_n) + \frac{\nabla^2 f_n}{2h^2}(x - x_n)(x - x_{n-1}) + \cdots \\ &\qquad \cdots + \frac{\nabla^n f_n}{n!h^n}(x - x_n)(x - x_{n-1}) \cdots (x - x_1) \\ &= f_n + \nabla f_n \left(\frac{x - x_n}{h}\right) + \frac{\nabla^2 f_n}{2}\left(\frac{x - x_n}{h}\right)\left(\frac{x - x_{n-1}}{h}\right) + \cdots \\ &\qquad \cdots + \frac{\nabla^n f_n}{n!}\left(\frac{x - x_n}{h}\right)\left(\frac{x - x_{n-1}}{h}\right) \cdots \left(\frac{x - x_1}{h}\right).\end{aligned}$$

Let $s = \dfrac{x - x_n}{h}$. Then for $j = n-1, n-2, \ldots, 1$,

$$\frac{x - x_j}{h} = \frac{x - (x_n - (n-j)h)}{h} = \frac{x - x_n + (n-j)h}{h} = s + n - j.$$

Thus,

$$P_n(x) = f_n + s\nabla f_n + s(s+1)\frac{\nabla^2 f_n}{2!} + \cdots + s(s+1)\cdots(s+n-1)\frac{\nabla^n f_n}{n!}$$

where $s = \dfrac{x - x_n}{h}$.

This is known as Gregory-Newton backward difference interpolation formula which is used for interpolating evenly-spaced data.

Note: If we are estimating f for values of x which lies close to the start of the tabulated values, then we would like to make the earliest use of the data points closest to x and so we use the Newton's forward difference interpolation formula in this case. For same reasons, we use the backward difference formula when estimating f for values of x which lies close to the end of the tabulated values.

Exercise

1. Construct the backward difference table for following data:

x	0.1	0.3	0.5	0.7	0.9	1.1	1.3
$f(x)$	0.003	0.067	0.148	0.248	0.370	0.518	0.697

Find $f(0.6)$ using Newton-Gregory backward difference interpolation formula.

Solution: The backward difference table for the given data is

x	f	∇f	$\nabla^2 f$	$\nabla^3 f$	$\nabla^4 f$	$\nabla^5 f$	$\nabla^6 f$
0.1	0.003						
		0.064					
0.3	0.067		0.017				
		0.081		0.002			
0.5	0.148		0.019		0.001		
		0.100		0.003		0	
0.7	0.248		0.022		0.001		0
		0.122		0.004		0	
0.9	0.370		0.026		0.001		
		0.148		0.005			
1.1	0.518		0.031				
		0.179					
1.3	0.697						

To find $f(0.6)$, we have $x_6 = 1.3$, $h = 0.2$, $x = 0.6$, so $s = \dfrac{x - x_6}{h} = \dfrac{0.6 - 1.3}{0.2} = -3.5$.
Also $f_6 = 0.697$, $\nabla f_6 = 0.179$, $\nabla^2 f_6 = 0.031$, $\nabla^3 f_6 = 0.005$, $\nabla^4 f_6 = 0.001$, $\nabla^5 f_6 = 0$ and $\nabla^6 f_6 = 0$. Therefore the estimated value of $f(0.6)$ using Newton-Gregory backward difference interpolation polynomial is

$$
\begin{aligned}
P_6(0.6) &= f_6 + s\nabla f_6 + s(s+1)\frac{\nabla^2 f_6}{2!} + s(s+1)(s+2)\frac{\nabla^3 f_6}{3!} + s(s+1)(s+2)(s+3)\frac{\nabla^4 f_6}{4!} \\
&+ s(s+1)(s+2)(s+3)(s+4)\frac{\nabla^5 f_6}{5!} + s(s+1)(s+2)(s+3)(s+4)(s+5)\frac{\nabla^6 f_6}{6!} \\
&= 0.697 + (-3.5) \times 0.179 + (-3.5)(-3.5+1)\frac{0.031}{2} + (-3.5)(-3.5+1)(-3.5+2)\frac{0.005}{6} \\
&+ (-3.5)(-3.5+1)(-3.5+2)(-3.5+3)\frac{0.001}{24} + 0 + 0 \\
&= 0.19546
\end{aligned}
$$

2. Do the above problems $2, 3, 4$ and 5 using backward difference interpolation.

2.4 Cubic Spline Interpolation

Given a set of $n+1$ data pairs (x_i, f_i), $i = 0, 1, \cdots, n$, interpolating these data using cubic splines is known as cubic spline interpolation.

A function $S(x)$ is said to be a cubic spline if

(i) $S(x)$ is a cubic polynomial $S_i(x)$ in each subinterval $[x_{i-1}, x_i]$, $i = 1, 2, \cdots, n$ i.e., $S(x) = S_i(x)$ for $x_{i-1} \leq x \leq x_i$, $i = 1, 2, \cdots, n$.

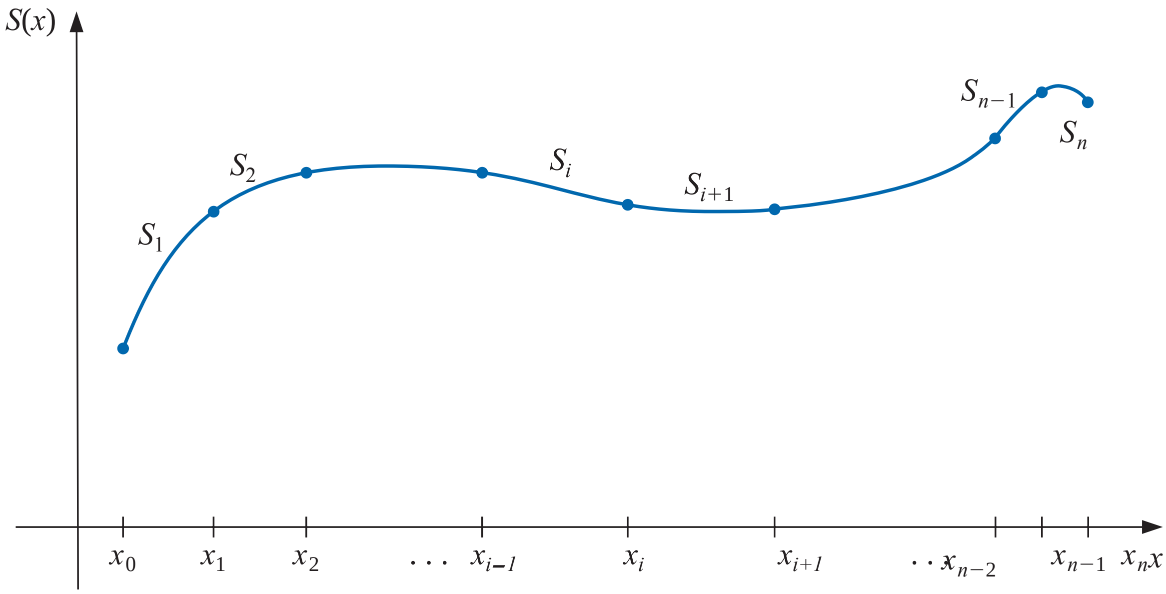
(ii) $S(x_i) = f_i, i = 0, 1, \cdots n$

(iii) $S_i(x_i) = S_{i+1}(x_i), i = 1, 2, \cdots, n-1$

(iv) $S_i'(x_i) = S_{i+1}'(x_i), i = 1, 2, \cdots, n-1$

(v) $S_i''(x_i) = S_{i+1}''(x_i), i = 1, 2, \cdots, n-1$

(vi) $S''(x_0) = S''(x_n) = 0$



Cubic spline with $S''(x_0) = S''(x_n) = 0$ as in (vi) above is a called natural cubic spline. If instead, we fix the value of $S'(x_0)$ and $S'(x_n)$ to some values a and b then it is called a clamped cubic spline.

Derivation of cubic spline interpolation polynomial:

Suppose we are given $n+1$ pair of data points $(x_i, f_i), i = 0, 1, \cdots, n$ and assume that

$$x_0 < x_1 < \cdots < x_n.$$

Let $S(x)$ be the cubic spline function that interpolates these data points. Then for each $i = 1, 2, \cdots, n$ and $x \in [x_{i-1}, x_i]$, $S(x) = S_i(x)$ where $S_i(x)$ is a cubic polynomial. So its second derivative $S_i''(x)$ is a straight line. Since $S_i(x)$ is defined on the interval $[x_{i-1}, x_i]$, so the points $(x_{i-1}, S_i''(x_{i-1}))$ and $(x_i, S_i''(x_i))$ lies on this straight line. By Lagrange's interpolation formula, the polynomial that interpolates these two points and thus approximates $S_i''(x)$ is a straight line given by

$$P_1(x) = S_i''(x_{i-1})\frac{x - x_i}{x_{i-1} - x_i} + S_i''(x_i)\frac{x - x_{i-1}}{x_i - x_{i-1}}$$

Since $S_i''(x)$ is itself a straight line, we have

$$S_i''(x) = S_i''(x_{i-1})\frac{x - x_i}{x_{i-1} - x_i} + S_i''(x_i)\frac{x - x_{i-1}}{x_i - x_{i-1}}$$

Let $S_i''(x_{i-1}) = a_{i-1}$, $S_i''(x_i) = a_i$ and $x_i - x_{i-1} = h_i$. Then

$$S_i''(x) = \frac{-a_{i-1}}{h_i}(x - x_i) + \frac{a_i}{h_i}(x - x_{i-1}) \cdots\cdots (1)$$

Integrating (1) twice w.r.t. x, we get,

$$S_i(x) = \frac{-a_{i-1}}{h_i}\frac{(x - x_i)^3}{6} + \frac{a_i}{h_i}\frac{(x - x_{i-1})^3}{6} + C_1x + C_2$$

$$\text{or, } S_i(x) = \frac{a_i(x - x_{i-1})^3 - a_{i-1}(x - x_i)^3}{6h_i} + C_1x + C_2 \cdots\cdots (2)$$

where C_1 and C_2 are constants of integration. Taking suitable values of b_1 and b_2, we can write

$$C_1x + C_2 = b_1(x - x_{i-1}) + b_2(x - x_i).$$

Let $x - x_i = u_i$, then $C_1x + C_2 = b_1u_{i-1} + b_2u_i$. So (2) becomes

$$S_i(x) = \frac{a_iu_{i-1}^3 - a_{i-1}u_i^3}{6h_i} + b_1u_{i-1} + b_2u_i \cdots\cdots (3)$$

Since $S_i(x_i) = f_i$ and $u_i = 0$, $u_{i-1} = h_i$ for $x = x_i$, so

$$f_i = \frac{a_ih_i^2}{6} + b_1h_i \Rightarrow b_1 = \frac{f_i}{h_i} - \frac{a_ih_i}{6}.$$

Since $S_i(x_{i-1}) = f_{i-1}$ and $u_{i-1} = 0$, $u_i = -h_i$ for $x = x_{i-1}$, so

$$f_{i-1} = \frac{a_{i-1}h_i^2}{6} - b_2h_i \Rightarrow b_2 = \frac{-f_{i-1}}{h_i} + \frac{a_{i-1}h_i}{6}.$$

Substituting the values of b_1 and b_2 in (3), we get,

$$S_i(x) = \frac{a_{i-1}}{6h_i}(h_i^2u_i - u_i^3) + \frac{a_i}{6h_i}(u_{i-1}^3 - h_i^2u_{i-1}) + \frac{1}{h_i}(f_iu_{i-1} - f_{i-1}u_i) \cdots\cdots (4)$$

We now evaluate the two unknowns a_{i-1} and a_i in equation (4). We have

$$\frac{du_i}{dx} = \frac{d(x - x_i)}{dx} = 1.$$

So differentiating (4) w.r.t. x, we get,

$$S_i'(x) = \frac{a_{i-1}}{6h_i}(h_i^2 - 3u_i^2) + \frac{a_i}{6h_i}(3u_{i-1}^2 - h_i^2) + \frac{1}{h_i}(f_i - f_{i-1}) \cdots\cdots (5)$$

Similarly,

$$S'_{i+1}(x) = \frac{a_i}{6h_{i+1}}(h_{i+1}^2 - 3u_{i+1}^2) + \frac{a_{i+1}}{6h_{i+1}}(3u_i^2 - h_{i+1}^2) + \frac{1}{h_{i+1}}(f_{i+1} - f_i)$$

Since $S'_i(x_i) = S'_{i+1}(x_i)$, we get,

$$\frac{a_{i-1}h_i}{6} + \frac{a_i h_i}{3} + \frac{f_i - f_{i-1}}{h_i} = \frac{-a_i h_{i+1}}{3} - \frac{a_{i+1}h_{i+1}}{6} + \frac{f_{i+1} - f_i}{h_{i+1}}$$

$$\text{or, } h_i a_{i-1} + 2(h_i + h_{i+1})a_i + h_{i+1}a_{i+1} = 6\left(\frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i}\right) \cdots\cdots (6)$$

Letting $i = 1, 2, \cdots, n-1$, we get $n-1$ simultaneous linear equations from (6) involving $n+1$ unknowns $a_0, a_1, \cdots, a_n$. But since $S''(x_0) = S''(x_n) = 0$, after differentiating (5) w.r.t. x and letting $x = x_0$ and $x = x_n$, we get

$$a_0 = a_n = 0.$$

Thus the $n-1$ unknowns $a_1, a_2, \cdots, a_{n-1}$ can be found by solving the above $n-1$ linear equations given by (6). This gives us the required cubic spline interpolation polynomial $S(x)$.

2.4.1 Exercise

1. Find the cubic spline interpolating polynomial and the value of f at $x = 7$ for the following set of data points:

x	4	9	16
f	2	3	4

Solution: Here, $f_0 = 2$, $f_1 = 3$, $f_2 = 4$, $h_1 = x_1 - x_0 = 9 - 4 = 5$, $h_2 = x_2 - x_1 = 16 - 9 = 7$. Since $n = 2$, there are two cubic splines, namely,

$$S_1(x),\ x_0 \le x \le x_1$$
$$S_2(x),\ x_1 \le x \le x_2$$

where

$$S_1(x) = \frac{a_0}{6h_1}(h_1^2 u_1 - u_1^3) + \frac{a_1}{6h_1}(u_0^3 - h_1^2 u_0) + \frac{1}{h_1}(f_1 u_0 - f_0 u_1)$$

and

$$S_2(x) = \frac{a_1}{6h_2}(h_2^2 u_2 - u_2^3) + \frac{a_2}{6h_2}(u_1^3 - h_2^2 u_1) + \frac{1}{h_2}(f_2 u_1 - f_1 u_2).$$

Since $a_0 = a_2 = 0$, we have

$$S_1(x) = \frac{a_1}{6h_1}(u_0^3 - h_1^2 u_0) + \frac{1}{h_1}(f_1 u_0 - f_0 u_1)$$

$$S_2(x) = \frac{a_1}{6h_2}(h_2^2 u_2 - u_2^3) + \frac{1}{h_2}(f_2 u_1 - f_1 u_2).$$

To determine a_1, we know that

$$h_1a_0 + 2(h_1+h_2)a_1 + h_2a_2 = 6\left(\frac{f_2-f_1}{h_2} - \frac{f_1-f_0}{h_1}\right)$$

$$\text{or, } 2(h_1+h_2)a_1 = 6\left(\frac{f_2-f_1}{h_2} - \frac{f_1-f_0}{h_1}\right) \ \{\text{since } a_0 = a_2 = 0\}$$

$$\text{or, } a_1 = -0.0143$$

Thus

$$S_1(x) = \frac{-0.0143}{6\times 5}[(x-4)^3 - 25(x-4)] + \frac{1}{5}[3(x-4) - 2(x-9)] \text{ for } 4 \le x \le 9$$

$$S_2(x) = \frac{-0.0143}{6\times 7}[49(x-16) - (x-16)^3] + \frac{1}{7}[4(x-4) - 3(x-9)] \text{ for } 9 \le x \le 16$$

Since 7 lies on the interval $[4, 9]$, so

$$S_1(7) = \frac{-0.0143}{6\times 5}[(7-4)^3 - 25(7-4)] + \frac{1}{5}[3(7-4) - 2(7-9)] = 2.6196.$$

2. Develop cubic splines for the data given below and predict $f(1.5)$.

x	0	1	2	3
f	1	-1	-1	0

Solution: Here, $f_0 = 1$, $f_1 = -1$, $f_2 = -1$, $f_3 = 0$, $h_1 = x_1 - x_0 = 1$, $h_2 = x_2 - x_1 = 1$, $h_3 = x_3 - x_2 = 1$. Since $n = 3$, there are three cubic splines, namely,

$$S_1(x),\ x_0 \le x \le x_1$$
$$S_2(x),\ x_1 \le x \le x_2$$
$$S_3(x),\ x_2 \le x \le x_3$$

where

$$S_1(x) = \frac{a_0}{6h_1}(h_1^2u_1 - u_1^3) + \frac{a_1}{6h_1}(u_0^3 - h_1^2u_0) + \frac{1}{h_1}(f_1u_0 - f_0u_1)$$

$$S_2(x) = \frac{a_1}{6h_2}(h_2^2u_2 - u_2^3) + \frac{a_2}{6h_2}(u_1^3 - h_2^2u_1) + \frac{1}{h_2}(f_2u_1 - f_1u_2)$$

and

$$S_3(x) = \frac{a_2}{6h_3}(h_3^2u_3 - u_3^3) + \frac{a_3}{6h_3}(u_2^3 - h_3^2u_2) + \frac{1}{h_3}(f_3u_2 - f_2u_3).$$

Since $a_0 = a_3 = 0$, we have

$$S_1(x) = \frac{a_1}{6h_1}(u_0^3 - h_1^2u_0) + \frac{1}{h_1}(f_1u_0 - f_0u_1)$$

$$S_2(x) = \frac{a_1}{6h_2}(h_2^2u_2 - u_2^3) + \frac{a_2}{6h_2}(u_1^3 - h_2^2u_1) + \frac{1}{h_2}(f_2u_1 - f_1u_2)$$

and

$$S_3(x) = \frac{a_2}{6h_3}(h_3^2 u_3 - u_3^3) + \frac{1}{h_3}(f_3 u_2 - f_2 u_3).$$

To determine a_1 and a_2, we know that

$$h_1 a_0 + 2(h_1 + h_2)a_1 + h_2 a_2 = 6\left(\frac{f_2 - f_1}{h_2} - \frac{f_1 - f_0}{h_1}\right)$$

$$\text{or, } 2(h_1 + h_2)a_1 + h_2 a_2 = 6\left(\frac{f_2 - f_1}{h_2} - \frac{f_1 - f_0}{h_1}\right) \quad \{\text{since } a_0 = 0\}$$

$$\text{or, } 2(1+1)a_1 + 1.a_2 = 6\left(\frac{-1-(-1)}{1} - \frac{-2}{1}\right)$$

$$\text{or, } 4a_1 + a_2 = 12 \cdots\cdots (A)$$

Also

$$h_2 a_1 + 2(h_2 + h_3)a_2 + h_3 a_3 = 6\left(\frac{f_3 - f_2}{h_3} - \frac{f_2 - f_1}{h_2}\right)$$

$$\text{or, } h_2 a_1 + 2(h_2 + h_3)a_2 = 6\left(\frac{f_3 - f_2}{h_3} - \frac{f_2 - f_1}{h_2}\right) \quad \{\text{since } a_3 = 0\}$$

$$\text{or, } 1.a_1 + 2(1+1)a_2 = 6\left(\frac{0-(-1)}{1} - \frac{(-1)-(-1)}{1}\right)$$

$$\text{or, } a_1 + 4a_2 = 6 \cdots\cdots (B)$$

Solving (A) and (B), we get $a_1 = 2.8$ and $a_2 = 0.8$. Thus

$$S_1(x) = \frac{2.8}{6}[x^3 - x] + [-x - (x-1)] = \frac{2.8}{6}x^3 - \frac{14.8}{6}x + 1 \text{ for } 0 \le x \le 1$$

$$S_2(x) = \frac{2.8}{6}[(x-2)-(x-2)^3] + \frac{0.8}{6}[(x-1)^3 - (x-1)] + [-(x-1)+(x-2)] \text{ for } 1 \le x \le 2$$

$$S_3(x) = \frac{0.8}{6}[(x-3) - (x-3)^3] + [0 - (-1)(x-3)] \text{ for } 2 \le x \le 3$$

Since 1.5 lies on the interval $[1, 2]$, so the estimate of $f(1.5)$ is given by

$$S_2(1.5) = \frac{2.8}{6}[(-0.5) - (-0.5)^3] + \frac{0.8}{6}[0.5^3 - 0.5] + [-0.5 - 0.5]$$

$$= \frac{2.8}{6} \times (-0.375) + \frac{0.8}{6} \times (-0.375) - 1 = -1.225$$

3. Develop cubic splines for data given below and estimate the function value at $x = 1.3$.

x	1	2	3
f	-1	3	15

4. Develop cubic splines for data given below and estimate the function value at $x = 2.1$.

x	1	2	3
f	1	1	5

5. Develop cubic splines for data given below and estimate the function value at $x = 2.7$.

x	1	2	3
f	-8	-1	18

6. Develop cubic splines for data given below and estimate the function value at $x = 1.5$.

x	1.0	3.0	4.0	7.0
f	1.5	4.5	9.0	25.5

2.5 Approximation

Let (x_i, y_i), $i = 1, 2, \cdots, n$ be a given set of n pairs of data points where x is an independent variable and y is a variable dependent upon x. Then the problem of finding an analytic expression of the form $y = f(x)$ that best fits the functional relationship suggested by the given data is called approximation.

Least Squares Approximation:

Let (x_i, y_i), $i = 1, 2, \cdots, n$ be a given set of n data pairs and let $y = f(x)$ be the curve that is fitted to this data. At $x = x_i$, the given value of the ordinate is y_i and the corresponding value on the fitting curve is $f(x_i)$. Then the error of approximation at $x = x_i$ is $e_i = y_i - f(x_i)$. Let

$$\begin{aligned} S &= [y_1 - f(x_1)]^2 + [y_2 - f(x_2)]^2 + \cdots + [y_n - f(x_n)]^2 \\ &= e_1^2 + e_2^2 + \cdots + e_n^2 \\ &= \sum_{i=1}^{n} e_i^2. \end{aligned}$$

Then the method of least squares approximation consists of finding an expression $y = f(x)$ such that the sum of the squares of the errors is minimized, i.e., S is minimized.

Types of least squares approximation:

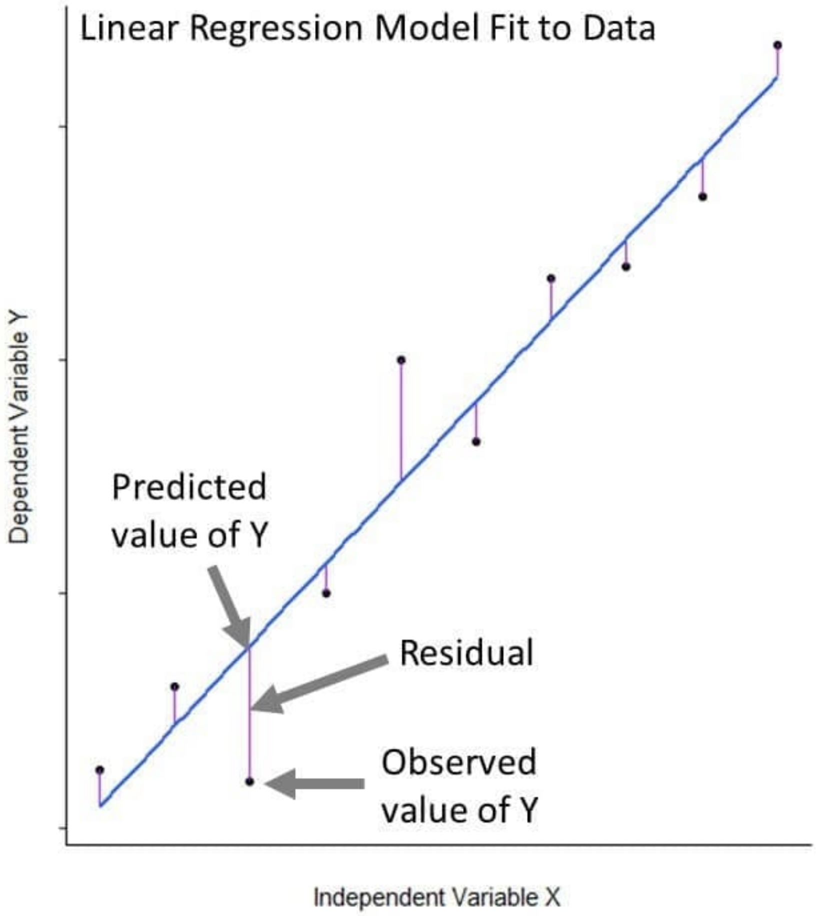
1. Method of linear least squares (Fitting a straight line):
Let (x_i, y_i), $i = 1, 2, \cdots, n$ be a given set of n data pairs. In the linear least squares method, we fit a straight line

$$y = f(x) = a + bx$$

in the given data. To determine this line, we need to determine the values of a and b. For this we proceed as follows:

For the data point (x_i, y_i), the vertical distance of this point from the line $f(x) = a + bx$ is the error e_i. Then

$$e_i = y_i - f(x_i) = y_i - (a + bx_i) = y_i - a - bx_i.$$



The sum of the squares of these individual errors is

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \cdots\cdots (1)$$

In the method of least squares, we choose a and b such that S is minimized. We know that the necessary condition for S to be minimum is

$$\frac{\partial S}{\partial a} = 0 \text{ and } \frac{\partial S}{\partial b} = 0$$

$$\text{i.e. } \frac{\partial}{\partial a}\left[\sum_{i=1}^{n} (y_i - a - bx_i)^2\right] = 0 \text{ and } \frac{\partial}{\partial b}\left[\sum_{i=1}^{n} (y_i - a - bx_i)^2\right] = 0$$

$$\text{i.e. } \sum_{i=1}^{n}\left[\frac{\partial}{\partial a}(y_i - a - bx_i)^2\right] = 0 \text{ and } \sum_{i=1}^{n}\left[\frac{\partial}{\partial b}(y_i - a - bx_i)^2\right] = 0$$

$$\text{i.e. } -2\sum_{i=1}^{n} (y_i - a - bx_i) = 0 \text{ and } -2\sum_{i=1}^{n} x_i(y_i - a - bx_i) = 0$$

$$\text{i.e. } \sum_{i=1}^{n} y_i - na - b\sum_{i=1}^{n} x_i = 0 \text{ and } \sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i - b\sum_{i=1}^{n} x_i^2 = 0$$

These two equations are called normal equations. Solving these two equations for a and b, we get,

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

and

$$a = \frac{\sum_{i=1}^{n} y_i}{n} - b\frac{\sum_{i=1}^{n} x_i}{n} = \overline{y} - b\overline{x}.$$

Algorithm (Linear least squares method):

INPUT: Set of n data pairs (x_i, y_i), $i = 1, 2, \cdots, n$.

PROCESS:

SET $sumx = 0, sumy = 0, sumx2 = 0, sumxy = 0$

FOR $i = 1$ TO n {

SET $sumx = sumx + x_i$

SET $sumy = sumy + y_i$

SET $sumx2 = sumx2 + (x_i * x_i)$

SET $sumxy = sumxy + (x_i * y_i)$

}

SET $b = \dfrac{n * sumxy - sumx * sumy}{n * sumx2 - sumx * sumx}$

SET $a = \dfrac{sumy - b * sumx}{n}$

OUTPUT: The straight line equation $y = a + bx$

2.5.1 Exercise (Fitting a Linear Equation)

1. Fit a straight line to the following set of data:

x	4	7	11	13	17
y	2	0	2	6	7

Solution: Here we have $n = 5$. We make the table from the given data as follows where the last row corresponds to the sum of the columns.

x_i	y_i	x_i^2	$x_i y_i$
4	2	16	8
7	0	49	0
11	2	121	22
13	6	169	78
17	7	289	119
52	17	644	227

So
$$b = \frac{5 \times 227 - 52 \times 17}{5 \times 644 - 52^2} = \frac{251}{516} = 0.4864$$
and
$$a = \frac{17}{5} - 0.4864 \times \frac{52}{5} = -1.6589.$$
Hence the straight line that can be fit is $y = -1.6589 + 0.4864x$.

2. Find the least-squares line that fits the following data:

x	1	3	4	6	8	9	11	14
y	1	2	4	4	5	7	8	9

3. Find the least-squares line that fits the following data:

x	1	2	3	4	5	6	7	8	9	10
y	52.5	58.7	65.0	70.2	75.4	81.1	87.2	95.5	102.2	108.4

4. Find the least-squares line that fits the following data:

x	1	2	3	4	5	6
y	5.04	8.12	10.64	13.18	16.20	20.04

2. Method of nonlinear least squares:
In this method, given a set of n data points (x_i, f_i), $i = 1, 2, \cdots, n$, we fit a nonlinear equation to the given data. We study two methods of fitting a non-linear equation to the given data set:

a. Fitting a polynomial equation

b. Fitting an exponential equation

a. Fitting a polynomial equation: Let (x_i, y_i), $i = 1, 2, \cdots, n$ be a given set of n data pairs. Let
$$y = f(x) = a_0 + a_1 x + \cdots + a_m x^m$$
be a polynomial of degree m that is fitted to the above data. Then the sum of the squares of the errors is given by
$$S = \sum_{i=1}^{n} [y_i - f(x_i)]^2.$$
We compute the coefficients $a_0, a_1, \cdots, a_m$ so that S is minimum. We know that the necessary condition for S to be minimum is
$$\frac{\partial S}{\partial a_0} = 0, \frac{\partial S}{\partial a_1} = 0, \cdots, \frac{\partial S}{\partial a_m} = 0.$$

For any $j = 0, 1, \cdots, m$, we have

$$\begin{aligned}
\frac{\partial S}{\partial a_j} &= \frac{\partial}{\partial a_j}\left[\sum_{i=1}^{n}(y_i - f(x_i))^2\right] \\
&= \sum_{i=1}^{n}\frac{\partial}{\partial a_j}[y_i - f(x_i)]^2 \\
&= \sum_{i=1}^{n} 2(y_i - f(x_i))\frac{\partial}{\partial a_j}(y_i - f(x_i)) \\
&= 2\sum_{i=1}^{n}(y_i - f(x_i))x_i^j.
\end{aligned}$$

So the condition for minimality becomes, for $j = 0, 1, \cdots, m$,

$$\begin{aligned}
\sum_{i=1}^{n}(y_i - f(x_i))x_i^j = 0 \\
\text{or, } \sum_{i=1}^{n}(y_i x_i^j - x_i^j f(x_i)) = 0 \\
\text{or, } \sum_{i=1}^{n} x_i^j(a_0 + a_1 x_i + \cdots + a_m x_i^m) = \sum_{i=1}^{n} y_i x_i^j \\
\text{or, } a_0\sum_{i=1}^{n} x_i^j + a_1\sum_{i=1}^{n} x_i^{j+1} + \cdots + a_m\sum_{i=1}^{n} x_i^{j+m} = \sum_{i=1}^{n} y_i x_i^j
\end{aligned}$$

Setting $j = 0$ to m, we get a system of $m+1$ linear equations as follows:

$$\begin{aligned}
a_0 n + a_1\sum x_i + \cdots + a_m\sum x_i^m &= \sum y_i \\
a_0\sum x_i + a_1\sum x_i^2 + \cdots + a_m\sum x_i^{m+1} &= \sum y_i x_i \\
\vdots \qquad\qquad & \vdots \qquad \vdots \\
a_0\sum x_i^m + a_1\sum x_i^{m+1} + \cdots + a_m\sum x_i^{2m} &= \sum y_i x_i^m
\end{aligned}$$

where each summation above is taken from $i = 1$ to n.

This set of $m+1$ equations can be represented in matrix notation as $AX = b$ where

$$A = \begin{pmatrix} n & \sum x_i & \cdots & \sum x_i^m \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^m & \sum x_i^{m+1} & \cdots & \sum x_i^{2m} \end{pmatrix}, X = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}, b = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum y_i x_i^m \end{pmatrix}.$$

Solving this system for X, we obtain the required values of a_i, $i = 0, 1, \cdots, m$.

Algorithm (Fitting a quadratic polynomial):

INPUT: Set of n data pairs (x_i, y_i), $i = 1, 2, \cdots, n$.

PROCESS:

SET $sumx = 0, sumx2 = 0, sumx3 = 0, sumx4 = 0$

SET $sumy = 0, sumxy = 0, sumx2y = 0$

FOR $i = 1$ TO n {

SET $sumx = sumx + x_i$
SET $sumx2 = sumx2 + x_i^2$
SET $sumx3 = sumx3 + x_i^3$
SET $sumx4 = sumx4 + x_i^4$
SET $sumy = sumy + y_i$
SET $sumxy = sumxy + x_i * y_i$
SET $sumx2y = sumx2y + x_i^2 * y_i$

}

Solve the following for a_0, a_1, a_2:

$$\begin{aligned} a_0 n + a_1 sumx + a_2 sumx2 &= sumy \\ a_0 sumx + a_1 sumx2 + a_2 sumx3 &= sumxy \\ a_0 sumx2 + a_1 sumx3 + a_2 sumx4 &= sumx2y \end{aligned}$$

OUTPUT: The quadratic polynomial equation $y = a_0 + a_1 x + a_2 x^2$.

Algorithm (Polynomial fitting):

1. Read number of data points n and order of polynomial m.
2. Read data values.
3. If $n \leq m$ print out 'regression is not possible' and stop else continue.
4. Compute coefficients of C matrix.
5. Compute coefficients of B matrix.
6. Solve for the coefficients $a_0, a_1, \cdots, a_m$.
7. Write the coefficients.
8. Estimate the function value at the given value of independent variable.
9. Stop.

2.5.2 Exercise (Fitting a Polynomial Equation)

1. Fit a polynomial of degree 2 to the data in the following table:

x	1.0	2.0	3.0	4.0
y	6.0	11.0	18.0	27.0

Solution: Let $y = a_0 + a_1 x + a_2 x^2$ be the required polynomial of degree 2. To determine a_0, a_1, a_2, we have to solve the following three simultaneous equations:

$$a_0 n + a_1 \sum x_i + a_2 \sum x_i^2 = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 = \sum y_i x_i$$
$$a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 = \sum y_i x_i^2$$

where all the summations above runs from $i = 1$ to 4. Now we make the following table:

x_i	y_i	x_i^2	x_i^3	x_i^4	$y_i x_i$	$y_i x_i^2$
1	6	1	1	1	6	6
2	11	4	8	16	22	44
3	18	9	27	81	54	162
4	27	16	64	256	108	432
10	62	30	100	354	190	644

Substituting these values, we get,

$$4a_0 + 10a_1 + 30a_2 = 62$$
$$10a_0 + 30a_1 + 100a_2 = 190$$
$$30a_0 + 100a_1 + 354a_2 = 644$$

Solving the above equations we get, $a_0 = 3$, $a_1 = 2$, and $a_2 = 1$. Thus the required polynomial is $y = 3 + 2x + x^2$.

2. Fit a quadratic polynomial to the following data:

x	1.0	1.5	2.0	2.5	3.0	3.5	4.0
y	1.1	1.2	1.5	2.6	2.8	3.3	4.1

3. Fit a quadratic polynomial to the following data:

x	0	1	2	3	4
y	1	1.8	1.3	2.5	6.3

4. Fit a quadratic polynomial to the following data:

x	1	2	3	4	5	6
y	0	16	30	70	110	155

b. Fitting an exponential equation: Let (x_i, y_i), $i = 1, 2, \cdots, n$ be a given set of n data pairs. Let

$$y = f(x) = ae^{bx}$$

be an exponential function that is fitted to the given data. Taking log on both sides, we get,

$$\log y = \log ae^{bx} = \log a + \log e^{bx} = \log a + bx$$

$$\text{or, } \log y = \log a + bx \cdots\cdots (1)$$

Let $u = \log y$ and $A = \log a$. Then (1) becomes

$$u = A + bx \cdots\cdots (2)$$

which is a linear equation. By the method of linear least squares approximation, we have

$$b = \frac{n\sum_{i=1}^{n} x_i u_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} u_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

and

$$A = \frac{\sum_{i=1}^{n} u_i}{n} - \frac{b\sum_{i=1}^{n} x_i}{n},$$

so $a = e^A$.

Algorithm (Fitting an exponential equation):

INPUT: Set of n data pairs (x_i, y_i), $i = 1, 2, \cdots, n$.

PROCESS:

SET $sumx = 0, sumu = 0, sumx2 = 0, sumxu = 0$

FOR $i = 1$ TO n {

SET $u_i = \log y_i$

SET $sumx = sumx + x_i$

SET $sumu = sumu + u_i$

SET $sumx2 = sumx2 + x_i^2$

SET $sumxu = sumxu + x_i * u_i$

}

SET $b = \dfrac{n * sumxu - sumx * sumu}{n * sumx2 - sumx * sumx}$

SET $A = \dfrac{sumu - b * sumx}{n}$

SET $a = e^A$

OUTPUT: The exponential equation $y = ae^{bx}$.

2.5.3 Exercise (Fitting an Exponential Equation)

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

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

Solution: Here we have $n = 5$. Let $u = \log y$ and $A = \log a$.

x_i	y_i	$u_i = \log y_i$	x_i^2	$x_i u_i$
2	4.077	1.405	4	2.810
4	11.084	2.405	16	9.620
6	30.128	3.405	36	20.430
8	81.897	4.405	64	35.240
10	222.62	5.405	100	54.050
30		17.025	220	122.150

So

$$b = \frac{5 \times 122.15 - 30 \times 17.025}{5 \times 220 - 30^2} = \frac{100}{200} = 0.5$$

and

$$A = \frac{17.025}{5} - 0.5 \times \frac{30}{5} = 0.405.$$

So $a = e^A = e^{0.405} = 1.499$. Thus $y = 1.499e^{0.5x}$ is the required exponential equation.

2. The curve $y = ae^{bx}$ is fitted to the following data:

x	1	2	3	4	5	6
y	1.5	4.6	13.9	40.1	125.1	299.5

Find the best values of a and b.

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

x	1	2	3	4	5	6
y	120	90	60	20	11	5

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

x	0	1	2	3	4	5	6	7	8
y	20	30	52	77	135	211	326	550	1052

Index