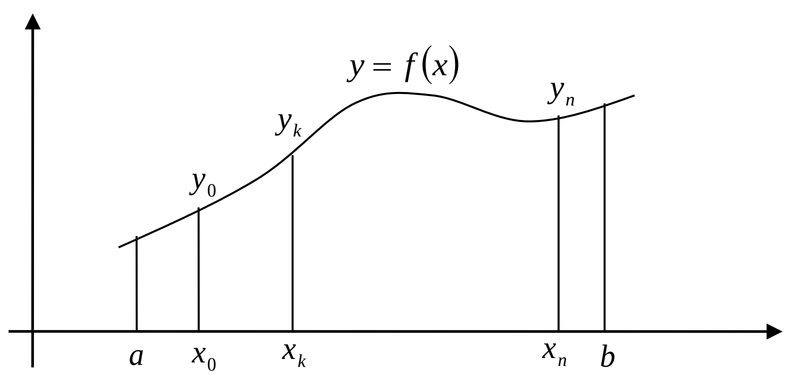
CHAPTER 6 LEAST SQUARES PROBLEMS, INTERPOLATION AND POLYNOMIAL APPROXIMATION

In this Chapter, you will learn:

- interpolation and extrapolation,
- three types of least squares approximation,
- Taylor polynomial,
- Lagrange's polynomial,
- Newton's divided-difference polynomial.

1. INTERPOLATION AND EXTRAPOLATION



Suppose that the function $y = f(x)$ is known at the (n+1) data points $(x_0, y_0),\ldots,(x_n, y_n)$, where the values x_k are spread out over the interval $[a,b]$ and satisfy $a \le x_0 < x_1 < \ldots < x_n \le b,$ and $y_k = f(x_k)$.

A polynomial $P_n(x)$ of degree n shall be constructed which passes through these (n+1) data points. In the construction, only the numerical values x_k and y_k are needed. Hence, the higher-order derivatives are not necessary. The polynomial $P_n(x)$ can be used to approximate over the interval $[a,b]$. The function $y_k = f(x_k)$ is available only at (n+1) tabulated data points and a method is needed to approximate $f(x)$ at non-tabulated abscissas.

Interpolation

Interpolation is a method of constructing new data points within the range of a discrete set of known data points. When $x_0 < x_k < x_n$, the approximation $y_k = P_n(x_k)$ is called an *interpolated* value.

Extrapolation

Extrapolation is a method of estimating, beyond the original interval $[a,b]$. If either $x_k < x_0$ or $x_k > x_n$, then the approximation $y_k = P_n(x_k)$ is called an *extrapolated* value. Extrapolation assumes that the behavior of $f(x)$ outside the range $[a,b]$ is identical to that inside the range and this may not always be valid.

2. LEAST SQUARES APPROXIMATION

Function approximation is closely related to the idea of function interpolation. In function approximation, we do not require the approximating function to match the given data exactly. The most common method of approximating data is the *least squares approximation*.

The method of *least squares* seeks to minimize the sum (all the tabulated data points) of the squares of the differences between the function value and the data value (total squared error). The minimum of the total squared error is attained when its partial derivatives are zero.

Linear Least Squares

There are *n* set of observations of related data, $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Let

$$y = a + bx \qquad (1)$$

be the equation to the linear line of best for them. We have to find the constants *a* and *b*. For any x_k, the *expected value* of *y* is y_k, the value calculated from the Equation (1) is

$$y_k = a + bx_k$$

and the observed value of *y* is y_i. The deviation (error) is

$$d_i = y_i - (a + bx_i)$$

By giving the values $i = 1,2,3,\ldots n$, we get the various of deviation.

Let $E(a,b)$ be the sum of squares of the deviations:

$$E(a,b) = \sum \left[y_i - (a + bx_i)\right]^2$$

For $E(a,b)$ to be minimum, the conditions are:

$$\frac{\partial E(a,b)}{\partial a} = 0 \text{ and } \frac{\partial E(a,b)}{\partial b} = 0$$

From the above conditions, we have the following *least squares normal equations*:

$$na + \left(\sum_{i=1}^{n} x_i\right)b = \sum_{i=1}^{n} y_i$$

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

We may get the values of *a* and *b* by solving the *least squares normal equations* as shown below:

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

By having the values of *a* and *b*, we get the equation of the *best fit linear line*.

Example 1: Consider the following data:

x_i	-1	0	1	2	3	4	5	6
y_i	10	9	7	5	4	3	0	-1

(a) Use the method of *least squares* to find the equation of the *best fit linear line*.

x_i	y_i	x_i^2	$x_i y_i$
-1	10	1	-10
0	9	0	0
1	7	1	7
2	5	4	10
3	4	9	12
4	3	16	12
5	0	25	0
6	-1	36	-6
20	**37**	**92**	**25**

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

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

$a = 8.6428571$ $b = -1.6071429$

The equation of the *best fit linear line* is ___________________________

(b) Estimate the value of y when $x = 1.5$.

Polynomial Least Squares

The method of least squares approximation data fitting is not restricted to linear function, $f(x) = a + bx$ only. As a matter of fact, in many cases data from experimental results are not linear, so we need to consider some other guess functions. Suppose that the guess function for the data is a polynomial. The m-th degree *polynomial* for n data pairs is expressed in the following.

$$\begin{aligned} P_m(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots + a_m x^m \\ &= \sum_{k=0}^{m} a_k x^k \end{aligned}$$

According to the least squares approximation, we need to find the coefficients $a_0, a_1, \ldots, a_m$ that minimized

$$E(a_0, a_1, \ldots, a_m) = \sum_{i=1}^{m} \left[y_i - P_m(x_i) \right]^2$$

$E(a_0, a_1, \ldots, a_m)$ is minimum if

$$\frac{\partial}{\partial a_j} E(a_0, a_1, \ldots, a_m) = 0; \quad j = 0,1,\ldots,m$$

From the above conditions (minimizing the sum of squares), we have m+1 *least squares normal equations*

$$\begin{array}{ccccccccc} a_0 n & + & a_1 \sum x_i & + & \cdots & + & a_m \sum x_i^m & = & \sum y_i \\ a_0 \sum x_i & + & a_1 \sum x_i^2 & + & \cdots & + & a_m \sum x_i^{m+1} & = & \sum x_i y_i \\ a_0 \sum x_i^2 & + & a_1 \sum x_i^3 & + & \cdots & + & a_m \sum x_i^{m+2} & = & \sum x_i^2 y_i \\ \vdots & + & \vdots & + & \cdots & + & \vdots & = & \vdots \\ a_0 \sum x_i^m & + & a_1 \sum x_i^{m+1} & + & \cdots & + & a_m \sum x_i^{2m} & = & \sum x_i^m y_i \end{array} \qquad (2)$$

The coefficients $a_0, a_1, \ldots, a_m$ can be found by solving the matrix of the linear system in Equation (2).

Exponential Least Squares

In many cases data from experimental tests are not linear. So, we need to fit the data with some functions other than polynomial function. Here, we broaden the *least squares approximation* to a popular form, *exponential* form.

Suppose we want to fit the data by a non-linear function, *exponential function* that is expressed as

$$y = ae^{bx} \quad \text{or} \quad y = ax^{b} \tag{3}$$

The non-linear function in Equation (3) is usually linearized by taking logarithm before determining the parameters:

$$\ln y = \ln a + bx \quad \text{or} \quad \ln y = \ln a + b \ln x$$

We now set $Y = \ln y$; $A = \ln a$; $X = \ln x$ to get a linear function of x or $\ln x$ as described earlier:

$$Y = A + bx \quad \text{or} \quad Y = A + bX$$

Finally, transform the points (x_i, y_i) to the points $(x_i, \ln y_i)$ or $(\ln x_i, \ln y_i)$ and use the *linear least squares* described in Section 2.1 to get A and b. Having obtained A and b, we use the relations

$$a = e^{A} \quad \text{and} \quad b = b$$

to obtain a and b.

In order to obtain the values of A and b, we must have the following *least squares normal equations*:

Case 1: $y = ae^{bx}$

$$\ln y = \ln a + bx \Rightarrow Y = A + bx$$

$$nA + \left(\sum_{i=1}^{n} x_i\right) b = \sum_{i=1}^{n} Y_i$$

$$\left(\sum_{i=1}^{n} x_i\right) A + \left(\sum_{i=1}^{n} x_i^2\right) b = \sum_{i=1}^{n} x_i Y_i$$

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

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

Case 2: $y = ax^b$

$$\ln y = \ln a + b\ln x \Rightarrow Y = A + bX$$

$$nA + \left(\sum_{i=1}^{n} X_i\right) b = \sum_{i=1}^{n} Y_i$$

$$\left(\sum_{i=1}^{n} X_i\right) A + \left(\sum_{i=1}^{n} X_i^2\right) b = \sum_{i=1}^{n} X_i Y_i$$

$$A = \frac{\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i Y_i \sum_{i=1}^{n} X_i}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}$$

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

Example 2: Consider the following data, find an equation of the form $y = ae^{bx}$

x_i	0	5	10	15	20
y_i	7	11	16	20	26

$y = ae^{bx}$

$\ln y = \ln a + bx$

Set $Y = \ln y; \quad A = \ln a : Y = A + bx$

x_i	0	5	10	15	20	**50**
Y_i	1.945910	2.397895	2.772589	2.995732	3.258097	**13.370223**
x_i^2	0	25	100	225	400	**750**
$x_i Y_i$	0	11.989475	27.72589	44.93598	65.16194	**149.813285**

3. TAYLOR POLYNOMIAL

Suppose that we want the best n^{th} degree approximation to $f(x)$ at $x = a$.
We compare $f(x)$ to

$$P_n(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + \ldots + a_n(x-a)^n$$

We make the following observations:

$$f(a) = P_n(a) = a_0 \quad \Rightarrow \quad a_0 = f(a)$$

$$f'(a) = P_n'(a) = a_1 + 2a_2(x-a) + 3a_3(x-a)^2 + \ldots + na_n(x-a)^{n-1}$$

At $x = a \Rightarrow a_1 = f'(a)$

$$f''(a) = P_n''(a) = 2a_2 + (3)(2)a_3(x-a) + \ldots + n(n-1)a_n(x-a)^{n-2}$$

At $x = a \Rightarrow a_2 = \frac{1}{2} f''(a)$

Note: each time we take a derivative we pick up the next integer in other words

$$a_3 = \frac{1}{(2)(3)} f^{(3)}(a)$$

If we define $f^{(k)}(a)$ to mean the k^{th} derivative of f evaluated at $x = a$ then

$$a_k = \frac{1}{k!} f^{(k)}(a)$$

In General

$$\begin{aligned} P_n(x) &= a_0 + a_1(x-a) + a_2(x-a)^2 + a_3(x-a)^3 + \ldots + a_n(x-a)^n \\ &= f(a) + f'(a)(x-a) + \frac{1}{2!} f''(a)(x-a)^2 + \frac{1}{3!} f^{(3)}(a)(x-a)^3 + \ldots + \frac{1}{n!} f^{(n)}(a)(x-a)^n \end{aligned}$$

$$P_n(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a)(x-a)^k$$

This is called the n^{th} degree *Taylor polynomial* at $x = a$.
When $a = 0$, $P_n(x)$ is called the n^{th} *McLaurin Polynomial*.

$$P_n(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0)(x)^k$$

Example 3: Find the *fifth degree McLaurin Polynomial* for **sin *x***

$f(x) = \sin(x)$ → $f(0) = \sin(0) = 0$

$f^{(1)}(x) = \cos(x)$ → $f^{(1)}(0) = \cos(0) = 1$

$f^{(2)}(x) = -\sin(x)$ → $f^{(2)}(0) = -\sin(0) = 0$

$f^{(3)}(x) = -\cos(x)$ → $f^{(3)}(0) = -\cos(0) = -1$

$f^{(4)}(x) = \sin(x)$ → $f^{(4)}(0) = \sin(x) = 0$

$f^{(5)}(x) = \cos(x)$ → $f^{(5)}(x) = \cos(x) = 1$

So that $P_5(x) = 0 + \frac{1}{1!}x + \frac{0}{2!}x^2 + \frac{-1}{3!}x^3 + \frac{0}{4!}x^4 + \frac{1}{5!}x^5 = x - \frac{1}{6}x^3 + \frac{1}{120}x^5$

4. LAGRANGE'S POLYNOMIAL

In this section, we find approximating polynomials that are determined by specifying certain points on the plane through which they must pass.

Consider the linear polynomial

$$P(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1; \; x \neq x_0$$

The linear polynomial passing through $(x_0, f(x_0))$ and $(x_1, f(x_1))$ was constructed using the quotients

$$L_0(x) = \frac{x - x_1}{x_0 - x_1} \quad \text{and} \quad L_1(x) = \frac{x - x_0}{x_1 - x_0}$$

To generalize the concept of linear interpolation, consider the construction of a polynomial of degree at most *n* that passes through the *n+1* points $(x_0, f(x_0))$, $(x_1, f(x_1)), \ldots, (x_n, f(x_n))$. We can construct the genetalized quotient:

$$L_i(x) = \frac{(x - x_0)\cdot(x - x_1)\cdots\cdots(x - x_{i-1})\cdot(x - x_{i+1})\cdots\cdots(x - x_n)}{(x_i - x_0)(x_i - x_1)\cdots(x_i - x_{i-1})\cdot(x_i - x_{i+1})\cdots(x_i - x_n)}$$

$$= \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j} \text{ for each } i = 0, 1, \ldots., n.$$

The interpolating polynomial is easily described if the form of $L_i(x)$ is known. This polynomial, called the n^{th} *Lagrange interpolating polynomial*, is defined in the following form:

$$P_n(x) = \sum_{i=0}^{n} f(x_i) L_i(x)$$

Example 4: By taking the numbers x_0 = 2, x_1 = 2.5 and x_2 = 4, find the *second interpolating polynomial* for $f(x) = \frac{1}{x}$ by using *Lagrange interpolating polynomial*.

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} =$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_1)(x_1 - x_2)} =$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} =$$

$$f(x_0) =$$

$$f(x_1) =$$

$$f(x_2) =$$

$$P_2(x) = \sum_{i=0}^{2} f(x_i) L_i(x)$$

$$=$$

5. NEWTON'S DIVIDED-DIFFERENCE POLYNOMIAL

It is useful to find several approximating polynomials $P_1(x), P_2(x), \ldots, P_n(x)$ and then choose the one that suits our needs. If the Lagrange polynomials are used, there is no constructive relationship between $P_{n-1}(x)$ and $P_n(x)$. Each polynomial has to be constructed individually, and the work required to compute the higher-degree polynomials involves many computations. We take a new approach and construct *Newton polynomials* that have the recursive pattern.

$$P_1(x) = a_0 + a_1(x - x_0)$$

$$P_2(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1)$$

$$P_3(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2)$$

$$\vdots$$

$$\boxed{\begin{aligned} P_n(x) = {} & a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2) + \ldots + \\ & a_n(x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{n-1}) \end{aligned}}$$

Suppose that $x_0, x_1, \ldots, x_n$ are $n+1$ distinct numbers in $[a,b]$. There exists a unique polynomial $P_n(x)$ of degree at most *n* with the property that

$$f(x_j) = P_n(x_j) \quad \text{for } j = 0,1,\ldots.n.$$

The coefficient $a_k; k = 0,1,2,\ldots,n$ of $P_n(x)$ depends on the values $f(x_j)$ for $j = 0, 1, \ldots, k$. It can be computed using *divided-difference*.
The *divided-difference* for a function $f(x)$ are defined as follows:

$$f[x_k] = f(x_k)$$

$$f[x_k, x_{k+1}] = \frac{f[x_{k+1}] - f[x_k]}{x_{k+1} - x_k}$$

$$f[x_k, x_{k+1}, x_{k+2}] = \frac{f[x_{k+1}, x_{k+2}] - f[x_k, x_{k+1}]}{x_{k+2} - x_k}$$

The recursive rule for constructing higher-order divided differences is

$$f[x_k, x_{k+1}, \ldots, x_{k+j}] = \frac{f[x_{k+1}, \ldots, x_{k+j}] - f[x_k, \ldots, x_{k+j-1}]}{x_{k+j} - x_k}$$

and is used to construct the divided-difference in table as follow:

x_k	$y_k = f[x_k]$	**First divided difference**	**Second divided difference**	**Third divided difference**	**Fourth divided Difference**
x_0	y_0				
		$f[x_0,x_1] = \frac{f[x_1]-f[x_0]}{x_1-x_0}$			
x_1	y_1		$f[x_0,x_1,x_2] = \frac{f[x_1x_2]-f[x_0x_1]}{x_2-x_0}$		
		$f[x_1,x_2] = \frac{f[x_2]-f[x_1]}{x_2-x_1}$		$f[x_0,x_1,x_2,x_3]$	
x_2	y_2		$f[x_1,x_2,x_3] = \frac{f[x_2x_3]-f[x_1x_2]}{x_3-x_1}$		$f[x_0,x_1,x_2,x_3,x_4]$
		$f[x_2,x_3] = \frac{f[x_3]-f[x_2]}{x_3-x_2}$		$f[x_1,x_2,x_3,x_4]$	
x_3	y_3		$f[x_2,x_3,x_4] = \frac{f[x_3x_4]-f[x_2x_3]}{x_4-x_2}$		
		$f[x_3,x_4] = \frac{f[x_4]-f[x_3]}{x_4-x_3}$			
x_4	y_4				

Example 5: Let $f(x) = x^3 - 4x$. Construct the divided-difference table based on the nodes $x_0 = 1$, $x_1 = 2$, ……., $x_5 = 6$, and find the *Newton polynomial* $P_3(x)$ based on x_0, x_1, x_2, x_3.

Divided difference table:

x_k	$y_k = f[x_k]$	First divided difference	Second divided difference	Third divided difference	Fourth divided Difference	Fifth divided difference
$x_0 = 1$						
$x_1 = 2$						
$x_2 = 3$						
$x_3 = 4$						
$x_4 = 5$						
$x_5 = 6$						

The *Newton polynomial*:

$$P_3(x) = -3 + 3(x - 1) + 6(x - 1)(x - 2) + (x - 1)(x - 2)(x - 3)$$

$$= x^3 - 4x$$