Interpolation

P. Sam Johnson

January 30, 2015

Overview

One of the basic ideas in Mathematics is that of a function and most useful tool of numerical analysis is **interpolation**.

According to Thiele (a numerical analyst), "**Interpolation is the art of reading between the lines of the table.**"

Broadly speaking, interpolation is the problem of obtaining the value of a function for any given functional information about it.

Interpolation technique is used in various disciplines like economics, business, population studies, price determination etc. It is used to fill in the gaps in the statistical data for the sake of continuity of information.

Overview

The concept of interpolation is the selection of a function $p(x)$ from a given class of functions in such a way that the graph of

$$y = p(x)$$

passes through a finite set of given data points. The function $p(x)$ is known as the **interpolating function** or **smoothing function**.

If $p(x)$ is a polynomial, then it is called the **interpolating polynomial** and the process is called the **polynomial interpolation**.

Similarly, if $p(x)$ is a finite trigonometric series, **we have trigonometric interpolation. But we restrict the interpolating function $p(x)$ to being a polynomial.**

Overview

The study of interpolation is based on the calculus of finite differences.

Polynomial interpolation theory has a number of important uses. Its primary uses is to furnish some mathematical tools that are used in developing methods in the areas of approximation theory, numerical integration, and the numerical solution of differential equations.

We discuss Newtons forward/backward formulae (for equally spaced nodes), Lagrange's formula, Newton's divided difference formulae (for unequally spaced nodes) and error bounds in **three** lectures.

Introduction

A census of the population of the India is taken every 10 years. The following table lists the population, in thousands of people, from 1951 to 2011.

Year	1951	1961	1971	1981	1991	2001	2011
Population (in thousands)	361,088	439,235	548,160	683,329	846,388	1,028,737	1,210,193

In reviewing these data, we might ask whether they could be used to provide a reasonable estimate of the population, say, in 1996, or even in the year 2014. Predictions of this type can be obtained by using a function that fits the given data.

This process is called **interpolation / extrapolation**.

Introduction

If y is a function of x, then the functional relation may be denoted by the equation

$$y = f(x).$$

The forms of $f(x)$ can, of course, be very diverse, but we consider $f(x)$ as a polynomial of the nth degree in x

$$y = a_0 + a_1 x + \cdots + a_n x^n \quad (a_n \neq 0).$$

We call x as the **independent variable** and y as the **dependent variable**. It is usual to call x as **argument** and y as function of the argument or **entry**.

Since the polynomials are relatively simple to deal with, **we interpolate to the data by polynomials.**

Weierstrass Approximation Theorem

If the value of x whose corresponding value y is to be estimated lies within the given range of x, then it is a problem of **interpolation**. On the other hand, if the value lies outside the range, then it is a problem of **extrapolation**.

Thus for the theory of interpolation, it is not esssential that the functional form of $f(x)$ be known. The only information needed is the values of the function given for some values of the argument.

In the method of interpolation, it is assumed that the function is capable of being expressed as a polynomial. This assumption is based on **Weierstrass approximation theorem**. That is, the existence of an interpolating polynomial is supported by the theorem.

Weierstrass Approximation Theorem

Given any function, defined and continuous on a closed and bounded interval, there exists a polynomial that is as "close" to the given function as desired. This result is expressed precisely in the following theorem.

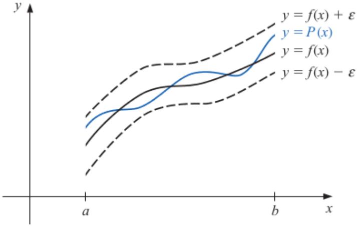
Theorem (Weierstrass Approximation Theorem)

Suppose that f is defined and continuous on $[a, b]$. For each $\varepsilon > 0$, there exists a polynomial $p(x)$, with the property that

$$|f(x) - p(x)| < \varepsilon, \text{ for all } x \in [a, b].$$

Why polynomials are important?

Weierstrass approximation theorem is illustrated in the figure.



In science and engineering, polynomials arise everywhere.

An important reason for considering the class of polynomials in the approximation of functions is that the "derivative and indefinite integral of a polynomial" are easy to determine and they are also polynomials.

For these reasons, **polynomials are often used for approximating continuous functions.** We introduce various interpolating polynomials using the concepts of forward, backward and central differences.

Main Asssumption for Interpolation

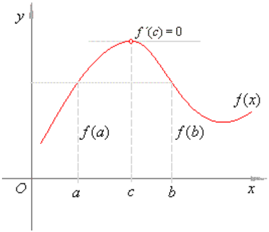
There are no sudden jumps or falls in the values of the function from one period to another. This assumption refers to the smoothness of $f(x)$ i.e., the shape of the curve $y = f(x)$ changes gradually over the period under consideration.

For example, if the population figures are given for, 1931, $1951, 1961, 1971$ and figures for 1941 are to be interpolated, we shall have to assume that the year 1941 **was not an exceptional year**, such as that affected by epidemics, war or other calamity or large scale immigration.

We recall Rolle's theorem, which is useful in evaluating error bounds.

Theorem (Rolle's Theorem)

Let f be continuous on $[a, b]$ and differentiable in (a, b). If $f(a) = f(b)$, then there is at least one point $c \in (a, b)$ such that $f'(c) = 0$.



Theorem (Generalized Rolle's Theorem)

Let f be continuous on $[a, b]$ and n times differentiable in (a, b). If $f(x)$ is zero at the $n + 1$ distinct numbers $c_0, c_1, \ldots, c_n$ in $[a, b]$, then a number c in (a, b) exists with $f^{(n)}(c) = 0$.

Error in Polynomial Interpolation

Let the function $y(x)$, defined by the $(n+1)$ points

$$(x_i, y_i), \qquad i = 0, 1, 2, \ldots, n$$

be continuous and differentiable $(n+1)$ times, and let $y(x)$ be approximated by a polynomial $p_n(x)$ of degree not exceeding n such that

$$p_n(x_i) = y_i$$

for $i = 0, 1, 2, \ldots, n$.

Using the polynomial $p_n(x)$ of degree n, we can obtain approximate values of $y(x)$ at some points other $x_i,\ 0 \leq i \leq n$.

Since the expression $y(x) - p_n(x)$ vanishes for $x = x_0, x_1, \ldots, x_n$ we put

$$y(x) - p_n(x) = L\pi_{n+1}(x) \tag{1}$$

where

$$\pi_{n+1}(x) = (x - x_0)(x - x_1)\cdots(x - x_n)$$

and L is to be determined such that the equation (1) holds for any intermediate value of $x' \in (x_0, x_n)$. Clearly

$$L = \frac{y(x') - p_n(x')}{\pi_{n+1}(x')}. \tag{2}$$

We construct a function $F(x)$ such that

$$F(x) = y(x) - p_n(x) - L\pi_{n+1}(x) \tag{3}$$

where L is given by the equation (2) above.
It is clear that

$$F(x_0) = F(x_1) = \cdots = F(x_n) = F(x') = 0$$

that is, $F(x)$ vanishes $(n+2)$ times in the interval $x_0 \leq x \leq x_n$.

Consequently, by the repeated application of Rolle's theorem, $F'(x)$ must vanish $(n+1)$ times, $F''(x)$ must vanish n times, etc,. in the interval $x_0 \leq x \leq x_n$. In particular, $F^{(n+1)}(x)$ must vanish once in the interval.

Let this point be given by $x = \xi, x_0 < \xi < x_n$. On differentiating the equation (3) $(n+1)$ times with respect to x and putting $x = \xi$, we obtain

$$F^{(n+1)}(\xi) = 0 = y^{(n+1)}(\xi) - L(n+1)!$$

so that

$$L = \frac{y^{(n+1)}(\xi)}{(n+1)!}. \tag{4}$$

Error in Polynomial Interpolation

Comparison of (2) and (4) yields the results

$$y(x') - p_n(x') = \frac{y^{(n+1)}(\xi)}{(n+1)!}\pi_{n+1}(x').$$

Dropping the prime on x', we obtain, for some $x_0 < \xi < x_n$,

$$y(x) - p_n(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_n)}{(n+1)!}y^{(n+1)}(\xi) \tag{5}$$

which is the required expression for the error. Since $y(x)$ is, generally, unknown and hence we do not have any information concerning $y^{(n+1)}(x)$, formula (5) is **almost useless in practical computations.**

On the other hand, **it is extremely useful in theroetical work in different branches of numerical analysis.**

Newton's Interpolation Formulae for Equally Spaced Points

Given the set of $(n+1)$ values,

$$(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n),$$

of x and y, it is required to find $p_n(x)$, a polynomial of the nth degree such that y and $p_n(x)$ agree at the tabulated points.

Let the values of x be equidistant,

$$x_i = x_0 + ih, i = 0, 1, 2, \ldots, n.$$

Since $p_n(x)$ is a polynomial of the nth degree, it may be written as

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

Polynomial Coefficients

Imposing the condition that y and $p_n(x)$ should agree at the set of tabulated points, we obtain

$$\begin{aligned}
a_0 &= y_0 \\
a_1 &= \frac{y_1 - y_0}{x_1 - x_0} = \frac{\Delta y_0}{h} \\
a_2 &= \frac{\Delta^2 y_0}{h^2 2!} \\
a_3 &= \frac{\Delta^3 y_0}{h^3 3!} \\
&\vdots \\
a_n &= \frac{\Delta^n y_0}{h^n n!}.
\end{aligned}$$

Interpolating Polynomial

Therefore

$$p_n(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{h^2 2!}(x - x_0)(x - x_1) + \cdots$$

$$\cdots + \frac{\Delta^n y_0}{h^n n!}(x - x_0)(x - x_1)\cdots(x - x_{n-1})$$

is the polynomial of degree n agreeing with the (unknown) function y at the tabulated points.

Remainder Term (Error) in Polynomial Interpolation

Theorem

Let $f(x)$ be a function defined in (a, b) and suppose that $f(x)$ have $n+1$ continuous derivatives on (a, b). If $a \leq x_0 < x_1 < \cdots < x_n \leq b$, then

$$f(x) - p_n(x) = \frac{(x - x_0)(x - x_1)\cdots(x - x_n)}{(n+1)!} f^{(n+1)}(\xi),$$

for some ξ between x and x_0 depending on $x_0, x_1, \ldots, x_n$ and f.

Newton's Forward Difference Interpolation Formula

Setting $x = x_0 + ph$ and substituting for $a_0, a_1, \ldots, a_n$, the above equation becomes

$$p_n(x) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \cdots$$

$$\cdots + \frac{p(p-1)(p-2)\cdots(p-n+1)}{n!}\Delta^n y_0$$

which is **(Gregory)-Newton's forward difference interpolation formula** and is useful for interpolation **near the beginning** of a set of tabular values and is useful for extrapolating the values of y (to the left of y_0).

Newton's Forward Difference Interpolation Formula

The first two terms of Newton's forward formula give the **linear interpolation** while the first three terms give a **parabolic interpolation** and so on.

Exercise

Let the function $y = f(x)$ take the values $y_0, y_1, \ldots, y_n$ corresponding to the values $x_0, x_0 + h, \ldots, x_0 + nh$ of x. Suppose $f(x)$ is a polynomial of degree n and it is required to evaluate $f(x)$ for $x = x_0 + ph$, where p is a any real number. Derive Newton's forward difference interpolation formula, by using shift operator E.

[Hint : $y_p = f(x) = f(x_0 + ph) = E^p f(x_0) = (1 + \Delta)^p y_0$.]

Forward Difference Table

The values $\boxed{\text{inside the boxes}}$ of the following difference table are used in deriving the Newton's forward difference interpolation formula.

Value of x	Value of $y = f(x)$	First Difference $\Delta f(x)$	Second Difference $\Delta^2 f(x)$	Third Difference $\Delta^3 f(x)$	Fourth Difference $\Delta^4 f(x)$
x_0	$\boxed{y_0}$				
		$\boxed{\Delta y_0}$			
$x_0 + h$	y_1		$\boxed{\Delta^2 y_0}$		
		Δy_1		$\boxed{\Delta^3 y_0}$	
$x_0 + 2h$	y_2		$\Delta^2 y_1$		$\boxed{\Delta^4 y_0}$
		Δy_2		$\Delta^3 y_1$	
$x_0 + 3h$	y_3		$\Delta^2 y_2$		
		Δy_3			
$x_0 + 4h$	y_4				

Error in Newton's Forward Difference Interpolation Formula

To find the error committed in replacing the function $y(x)$ by means of the polynomial $p_n(x)$, we obtain

$$y(x) - p_n(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_n)}{(n+1)!} y^{(n+1)}(\xi)$$

for some $\xi \in (x_0, x_n)$.

The error in the Newton's forward difference interpolation formula is

$$y(x) - p_n(x) = \frac{p(p-1(p-2)\cdots(p-n)}{(n+1)!} h^{n+1} y^{(n+1)}(\xi)$$

for some $\xi \in (x_0, x_n)$, and $x = x_0 + ph$.

Error in Newton's Forward Difference Interpolation Formula

As remarked earlier we do not have any information concerning $y^{(n+1)}(x)$, and therefore the above formula is useless in practice.

Neverthless, if $y^{(n+1)}(x)$ does not vary too rapidly in the interval, a useful estimate of the derivative can be obtained in the following way. Expanding $y(x+h)$ by Taylor's series, we obtain

$$y(x+h) = y(x) + hy'(x) + \frac{h^2}{2!}y''(x) + \cdots .$$

Neglecting the terms containing h^2 and higher powers of h, this gives

$$y'(x) \approx \frac{y(x+h) - y(x)}{h} = \frac{\Delta y(x)}{h}$$

Error in Newton's Forward Difference Interpolation Formula

Writing $y'(x)$ as $Dy(x)$ where $D \equiv d/dx$, the **differentiation operator**, the above equation gives the operator relations

$$D \equiv \frac{1}{h}\Delta \text{ and so } D^{n+1} \equiv \frac{1}{h^{n+1}}\Delta^{n+1}.$$

We thus obtain

$$y^{(n+1)}(x) \approx \frac{1}{h^{n+1}}\Delta^{n+1}y(x).$$

Hence the equation can be written as (equally spaced nodes, $x = x_0 + ph$)

$$y(x) - p_n(x) = \frac{p(p-1)(p-2)\cdots(p-n)}{(n+1)!}\Delta^{n+1}y(\xi)$$

for some $\xi \in (x_0, x_n)$, which is suitable for computation.

Newton's Backward Difference Interpolation Formula

Suppose we assume $p_n(x)$ in the following form

$$p_n(x) = a_0 + a_1(x - x_n) + a_2(x - x_n)(x - x_{n-1}) + \cdots$$

$$\cdots + a_n(x - x_n)(x - x_{n-1}) \cdots (x - x_1)$$

and then impose the condition that y and $p_n(x)$ should agree at the tabulated points $x_n, x_{n-1}, \ldots, x_2, x_1, x_0$, we obtain (after some simplification)

$$p_n(x) = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \cdots + \frac{p(p+1)\cdots(p+n-1)}{n!}\nabla^n y_n$$

where $p = (x - x_n)/h$.

This is **(Gregory)-Newton's backward difference interpolation formula** and it uses tabular values to the left of y_n. This formula is therefore useful for interpolation **near the end of** the tabular values and is useful for extrapolating values of y (to the right of y_n).

Backward Difference Table

The values inside the boxes of the following difference table are used in deriving the Newton's backward difference interpolation formula.

Value of x	Value of $y = f(x)$	First Difference $\nabla f(x)$	Second Difference $\nabla^2 f(x)$	Third Difference $\nabla^3 f(x)$	Fourth Difference $\nabla^4 f(x)$
x_0	y_0				
		∇y_1			
$x_0 + h$	y_1		$\nabla^2 y_2$		
		∇y_2		$\boxed{\nabla^3 y_3}$	
$x_0 + 2h$	y_2		$\nabla^2 y_3$		$\boxed{\nabla^4 y_4}$
		∇y_3		$\boxed{\nabla^3 y_4}$	
$x_0 + 3h$	y_3		$\boxed{\nabla^2 y_4}$		
		$\boxed{\nabla y_4}$			
$x_0 + 4h$	$\boxed{y_4}$				

Error in Newton's Backward Difference Interpolation Formula

It can be shown that the error in this formula may be written as

$$y(x) - p_n(x) = \frac{p(p+1)(p+2)\cdots(p+n)}{(n+1)!} h^{n+1} y^{(n+!)}(\xi)$$

where $x_0 < \xi < x_n$ and $x = x_n + ph$.

Taylor's Theorem

Theorem

Let $f(x)$ have $n+1$ continuous derivatives on $[a,b]$ for some $n \geq 0$, and let $x, x_0 \in [a,b]$. Then $f(x) = p_n(x) + R_n(x)$
where

$$p_n(x) = \sum_{k=0}^{n} \frac{(x-x_0)^k}{k!} f^{(k)}(x_0) \qquad \textbf{(}n\textbf{-degree polynomial)}$$

and

$$R_n(x) = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi) \qquad \textbf{(error term)}$$

for some ξ between x and x_0.

Hence "ξ between x and x_0" means that either $x_0 < \xi < x$ or $x < \xi < x_0$ depending on the particular values of x and x_0 involved.

Taylor Polynomials

Here $p_n(X)$ is called the n**th Taylor polynomial** for f about x_0, and R_n is called the **remainder term** (or **truncation error**) associated with $p_n(x)$. Since the number ξ in the truncation error R_n depends on the value of x at which the polynomial $p_n(x)$ is being evaluated, it is a function of the variable x.

Taylor's theorem simply ensures that such a function exists, and that its value lies between x and x_0.

In fact, one of the common problems in numerical methods is to try to determine a realistic bound for the value of $f^{(n+1)}(\xi)$ when x is within some specified interval.

Taylor polynomials are not useful for interpolation

The Taylor polynomials are one of the fundamental building blocks of numerical analysis.

The Taylor polynomials agree as closely as possible with a given function at a specific point, but they concentrate their accuracy near that point.

A good interpolation needs to provide a relatively accurate approximation over an entire interval, and Taylor polynomials do not generally do this.

Example

Taylor polynomials of various degree for $f(x) = 1/x$ *about* $x_0 = 1$ *are*

$$p_n(x) = \sum_{k=0}^{n} (-1)^k (x-1)^k.$$

When we approximate $f(3) = 1/3$ *by* $p_n(3)$ *for larger values of* n*, the approximations become increasingly inaccurate, as shown in the following table.*

n	0	1	2	3	4	5	6	7
$p_n(3)$	1	-1	3	-5	11	-21	43	-85

Taylor polynomials are not appropriate for interpolation

Since the Taylor polynomials have the property that all the information used in the approximation is concentrated at the single point x_0, it is not uncommon for these polynomials to give inaccurate approximations as we move away from x_0. This limits Taylor polynomial approximation to the situation in which approximations are needed only at points close to x_0.

For ordinary computational purposes it is more efficient to use methods that include information at various points.

The primary use of Taylor polynomials in numerical analysis is not for approximation purposes, but for the derivation of numerical techniques and for error estimation.

Since the Taylor polynomials are not appropriate for interpolation, alternative methods are needed.

The problem of determining a polynomial of degree one that passes through the distinct points (x_0, y_0) and (x_1, y_1) is the same as approximating a function f for which $f(x_0) = y_0$ and $f(x_1) = y_1$ by means of a first-degree polynomial interpolating, or agreeing with, the values of f at the given points.

We first define the functions

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

and then define

$$p_1(x) = L_0(x)y_0 + L_1(x)y_1 = \frac{x - x_1}{x_0 - x_1}y_0 + \frac{x - x_0}{x_1 - x_0}y_1.$$

Since $L_0(x_0) = 1, L_0(x_1) = 0, L_1(x_0) = 0,$ and $L_1(x_1) = 1$, we have $p_1(x_0) = y_0$ and $p_1(x_1) = y_1$. So p_1 is the unique linear function passing (x_0, y_0) and (x_1, y_1).

Exercises

1. *The table gives the distance in nautical miles of the visible horizon for the given heights (in feet) above the earth's surface.*

x	100	150	200	250	300	350	400
$y = f(x)$	10.63	13.03	15.04	16.81	18.42	19.90	21.27

Find the values of y when $x = 160$ and $x = 410$.

2. *From the following table, estimate the number of students who obtained marks between 40 and 45.*

Marks	30-40	40-50	50-60	60-70	70-80
No. of Students	31	42	51	35	31

3. *Find the cubic polynomial which takes the following values.*

x	0	1	2	3
$y = f(x)$	1	2	1	10

Also compute $f(4)$.

Exercises

4. *In the table below, the values of y are consecutive terms of a series of which 23.6 is the 6th term. Find the first and tenth terms of the series.*

x	3	4	5	6	7	8	9
$y = f(x)$	4.8	8.4	14.5	23.6	36.2	52.8	73.9

5. *Using Newton's forward interpolation formula, show that*

$$\sum_{k=1}^{n} k^3 = \left\{ \frac{n(n+1)}{2} \right\}^2 .$$

Lagrange Interpolating Polynomial

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 $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$.

In this case we need to construct, for each $k = 0, 1, 2, \ldots, n$, a function $L_k(x)$ (called **Lagrange basis**, also called the nth **Lagrange interpolating polynomial**) with the property that $L_k(x_i) = \begin{cases} 0 & \text{when } i \neq k \\ 1 & \text{when } i = k \end{cases}$ hence

$$L_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{(x - x_i)}{(x_k - x_i)}.$$

Lagrange Interpolating Polynomial

The interpolating polynomial is easily described once the form of L_k is known, by the following theorem.

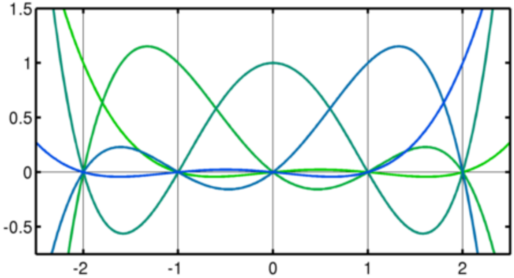
Theorem

If $n+1$ points $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ are given, then a unique polynomial $p_n(x)$ of degree at most n exists with $f(x_k) = p_n(x_k)$ for each $k = 0, 1, \ldots, n$. This polynomial is given by

$$p_n(x) = \sum_{k=0}^{n} f(x_k) L_k(x).$$

Graphs of Lagrange Interpolating Polynomials

Given 5 points $(x_0, y_0), (x_1, y_1), \ldots, (x_4, y_4)$, a sketch of the graph of a typical L_k is shown in the following figure.



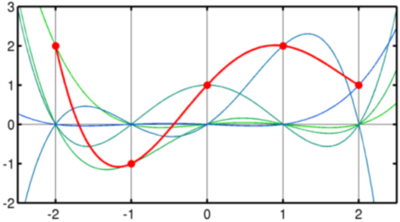
Note how each basis polynomial has a value of 1 for $x = x_k$ $(0 \leq k \leq 4)$, and a value of 0 at all other locations.

Example

Simply multiplying each basis with the corresponding sample value, and adding them all up yields the interpolating polynomial

$$p(x) = \sum_{k=0}^{4} f(x_k) L_k(x).$$

The 5 weighted polynomials are $L_k(x)f(x_k)$ $(0 \le k \le 4)$ and their sum (red line) is the interpolating polynomial $p(x)$ (red line) which is shown in the following figure.



How to calculate error bound?

The next step is to calculate a remainder term or bound for the error involved in approximating a function by an interpolating polynomial. This is done in the following theorem.

Theorem (An Important Result for Error Formula)

Suppose $x_0, x_1, \ldots, x_n$ *are distinct numbers in the interval* $[a, b]$ *and* $f \in C^{n+1}[a, b]$*. Then, for each* $x \in [a, b]$*, a number* $\xi(x)$ *(generally unknown) in* (a, b) *exists with*

$$f(x) = p(x) + \frac{(x - x_0)(x - x_1)\cdots(x - x_n)}{(n+1)!} f^{(n+1)}(\xi(x)),$$

where $p(x)$ *is the interpolating polynomial given by* $p(x) = \sum_{k=0}^{n} f(x_k)L_k(x)$.

The above formula is also called **'Lagrange Error Formula'**.

Error Analysis

The error formula is an important theoretical result because Lagrange polynomials are used extensively for deriving numerical differentiation and integration methods.

Error bounds for these techniques are obtained from the "Lagrange error formula".

Note that the error for the Lagrange polynomial is quite similar to that for the Taylor polynomial.

Comparison of Error Bounds in Taylor and Lagrange Polynomials

The nth **Taylor polynomial** about x_0 concentrates all the known information at x_0 and has an error term of the form

$$\frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi).$$

The **Lagrange polynomial** of degree n uses information at the distinct numbers $x_0, x_1, \ldots, x_n$ and, in place of $(x - x_0)^{n+1}$, its error formula uses a product of the $n + 1$ terms $(x - x_0)(x - x_1) \cdots (x - x_n)$

$$\frac{(x - x_0)(x - x_1) \cdots (x - x_n)}{(n+1)!} f^{(n+1)}(\xi).$$

Double Interpolation

We have so for derived interpolation formulae to approximate a function of a single variable.

In case of functions of two variables, we interpolate with respect to the first variable keeping the other variable constant. Then interpolate with respect to the second variable.

Similarly, we can extend the said procedure for functions of three variables.

Inverse Interpolation

We have been finding the value of y corresponding to a certain value of x from a given set of values of x and y.

On the other hand, the process of estimating the value of x for a value of y is called **inverse interpolation**. When the values of y are unequally spaced, Lagrange's method is used and when the values of y are equally spaced, the following iterative method is used.

In the procedure, x is assumed to be expressible as a polynomial in y.

Newton's forward interpolation formula is

$$p_n(x) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \cdots. \quad (6)$$

From (6) we get

$$p = \frac{1}{\Delta y_0}\left\{y_p - y_0 - \frac{p(p-1)}{2!}\Delta^2 y_0 - \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 - \cdots\right\}.$$

Neglecting the second and higher differences, we obtain the **first approximation** to p as

$$p_1 = \frac{y_p - y_0}{\Delta y_0}.$$

To find the **second approximation**, retaining the term with second differences in (6) and replacing p by p_1, we get

$$p_2 = \frac{1}{\Delta y_0}\left\{y_p - y_0 - \frac{p_1(p_1-1)}{2!}\Delta^2 y_0\right\}.$$

Iterative Method

To find the **third approximation**, retaining the term with third differences in (6) and replacing p by p_2, we get

$$p_3 = \frac{1}{\Delta y_0}\left\{ y_p - y_0 - \frac{p_2(p_2-1)}{2!}\Delta^2 y_0 - \frac{p_2(p_2-1)(p_2-2)}{3!}\Delta^3 y_0 \right\}$$

and so on. This process is continued till two successive approximations of p agree with each other.

This technique can equally well be applied by any other interpolation formula. This method is a powerful iterative procedure for finding the roots of an equation to a good degree of accuracy.

We shall discuss later some more formulae for finding roots of an equation.

Exercises

6. *Find the polynomial* $f(x)$ *by using Lagrange's formula and hence find* $f(3)$ *for*

x	0	1	2	3
$y = f(x)$	2	3	12	147

7. *A curve passes through the points* $(0, 18), (1, 10), (3, -18)$ *and* $(6, 90)$. *Find the slope of the curve at* $x = 2$.

8. *Using Lagrange's formula, express the function*

$$\frac{3x^2 + x + 1}{(x-1)(x-2)(x-3)}$$

as a sum of partial fractions.

9. *Find the missing term in the following table using interpolation.*

x	0	1	2	3	4
$y = f(x)$	1	3	9	-	81

Exercises

10. *Find the distance moved by a particle and its acceleration at the end of 4 seconds, if the time verses velocity data is as follows.*

t	0	1	3	4
v	21	15	12	10

11. *Using Lagrange's formula prove that*

$$y_0 = \frac{y_1 + y_{-1}}{2} - \frac{1}{8}\left\{\frac{1}{2}(y_3 - y_1) - \frac{1}{2}(y_{-1} - y_{-3})\right\}.$$

[Hint : Here $x_0 = -3, x_1 = -1, x_2 = 1, x_2 = 3$.]

12. *Given*

$$\begin{aligned} \log_{10} 654 &= 2.8156, \quad \log_{10} 658 = 2.8182, \\ \log_{10} 659 &= 2.8189, \quad \log_{10} 661 = 2.8202 \end{aligned}$$

find by using Lagrange's formula, the value of $\log_{10} 656$.

Exercises

13. *The following table gives the viscosity of an oil as a function of temperature. Use Lagrange's formula to find viscosity of oil at a temperature of* 140°.

Temperature	110°	130°	160°	190°
Viscosity	*10.8*	*8.1*	*5.5*	*4.8*

14. *Given* $u_1 = 40, u_3 = 45, u_5 = 54$, *find* u_2 *and* u_4.

15. *Given* $y_0 = 3, y_1 = 12, y_2 = 81, y_3 = 200, y_4 = 100, y_5 = 8$, *without forming the difference table, find* $\Delta^5 y_0$.

16. *From the data given below, find the number of students whose weight is between* 60 *and* 70.

Weight	*0-40*	*40-60*	*60-80*	*80-100*	*100-120*
No. of Students	*250*	*120*	*100*	*70*	*50*

Exercises

17. *The values of $U(x)$ are known at a, b, c. Show that maximum or minimum of Lagrange's interpolation formula is attained at*

$$x = \frac{\sum U_a(b^2 - c^2)}{2\sum U_a(b - c)}.$$

18. *By iterative method, tabulate $y = x^3$ for $x = 2, 3, 4, 5$ and calculate the cube root of 10 correct to 3 decimal places.*

19. *The following values of $y = f(x)$ are given*

x	10	15	20
y	1754	2648	3564

Find the value of x for $y = 3000$ by iterative method.

20. *Using inverse interpolation, find the real root of the equation $x^3 + x - 3 = 0$ which is close to 1.2.*

Exercises

21. *Solve the equation* $x = 10 \log x$, *by iterative method, given that*

x	1.35	1.36	1.37	1.38
$\log x$	0.1303	0.1355	0.1367	0.1392

22. *Apply Lagrange's method, to find the value of* x *when* $f(x) = 15$ *from the given data.*

x	5	6	9	11
$f(x)$	12	13	14	16

23. *The equation* $x^3 - 15x + 4$ *has a root close to* 0.3, *obtain this root upto* 4 *decimal places using inverse interpolation.*

Divided Differences

The Lagrange's formula has the drawback that if another interpolation values were inserted, then the interpolation coefficients are required to be calculated.

This labour of recomputing the interpolation coefficients is saved by using Newton's general interpolation formula which employs what are called "**divided differences**".

Newton's Divided Difference Interpolation

Suppose that $p_n(x)$ is the nth degree Lagrange polynomial that agrees with the function f at the distinct numbers $x_0, x_1, \ldots, x_n$. Although **this polynomial is unique**, alternate algebraic representations are useful in certain situations.

The divided differences of f with respect to $x_0, x_1, \ldots, x_n$ are used to express $p_n(x)$ in the form

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

for appropriate constants $a_0, a_1, \ldots, a_n$.

Evaluating Coefficients

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 x_0 leaves only the constant term a_0. That is, $a_0 = p_n(x_0) = f(x_0)$.

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

$$f(x_0) + a_1(x_1 - x_0) = p_n(x_1) = f(x_1)$$

so

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

Divided Difference Notation

The divided-difference notation, is introduced, which is similar to Aitkens Δ^2 notation.

The **zeroth divided difference** of the function f with respect to x_i, denoted by $f[x_i]$, is simply the value of f at x_i,

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

The remaining divided differences are defined inductively.

The **first divided difference** of f with respect to x_i and x_{i+1} is denoted by $f[x_i, x_{i+1}]$ and is defined as

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

Newton's General Interpolation Formula with Divided Differences

The **second divided difference**, $f[x_i, x_{i+1}, x_{i+2}]$, is defined as

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

The process ends with single ***n*th divided difference**,

$$f[x_0, x_1, \ldots, x_n] = \frac{f[x_1, x_2, \ldots, x_n] - f[x_0, x_1, \ldots, x_{n-1}]}{x_n - x_0}.$$

Hence $p_n(x)$ can be rewritten as

$$p_n(x) = f[x_0] + \sum_{k=1}^{n} f[x_0, x_1, \ldots, x_k](x - x_0) \cdots (x - x_{k-1}),$$

which is called **Newton's general interpolation formula with divided differences**.

Properties of Divided Differences

- The divided differences are symmetrical in their arguments – the value of $f[x_0, x_1, \ldots, x_k]$ is independent of the order of the numbers $x_0, x_1, \ldots, x_k$. That is,

$$f[x_0, x_1] = f[x_1, x_0],$$

$$f[x_0, x_1, x_2] = f[x_1, x_2, x_0] = f[x_2, x_0, x_1] \quad \text{and so on.}$$

- The nth divided differences of a polynomial of degree n are constants.

Divided Difference Table

x	$f(x)$	First Divided Differences	Second Divided Differences
x_0	$f[x_0]$		
		$f[x_0, x_1] = \frac{f[x_1]-f[x_0]}{x_1-x_0}$	
x_1	$f[x_1]$		$f[x_0, x_1, x_2] = \frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0}$
		$f[x_1, x_2] = \frac{f[x_2]-f[x_1]}{x_2-x_1}$	
x_2	$f[x_2]$		

Suppose f is continuously differentiable on $[x_0, x_1]$. By the mean value theorem, there exists $\xi \in [x_0, x_1]$ such that

$$f'(\xi) = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f[x_0, x_1].$$

The following theorem generalizes this result.

Theorem

Suppose that $f \in C^n[a, b]$ and $x_0, x_1, \ldots, x_n$ are distinct numbers in $[a, b]$. Then a number ξ (generally unknown) exists in (a, b) with

$$f[x_0, x_1, \ldots, x_n] = \frac{f^{(n)}(\xi)}{n!}.$$

Nodes with equal spacing : Let $h = x_{i+1} - x_i, \quad i = 0, 1, \ldots, n-1$. Hence

$$p_n(x) = p_n(x_0 + ph) = f[x_0] + \sum_{k=1}^{p} \binom{p}{k} k! h^k f[x_0, x_1, \ldots, x_k].$$

Remainder Term (Error) in Newton's Divided Difference Formula

The Newton's Divided Difference Formula is given by

$$f(x) = f[x_0] + \sum_{k=1}^{n} f[x_0, x_1, \ldots, x_k](x - x_0)\cdots(x - x_{k-1}) + R_n(x),$$

with the error term as

$$R_n(x) = \frac{(x - x_0)(x - x_1)\cdots(x - x_n)}{(n+1)!} f^{(n+1)}(\xi)$$

for some $\xi \in (x_0, x_n)$.

The error term as in case of Lagrange's formula.

Relation between Dividied and Forward Differences

Forward difference operator is defined by

$$\Delta f(x_i) = f(x_{i+1}) - f(x_i), \qquad \text{for} \quad i \geq 0.$$

Higher powers are defined recursively by

$$\Delta^k f(x_i) = \Delta(\Delta^{k-1} f(x_i)), \qquad \text{for} \quad i \geq 0.$$

With this notation,

$$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0} = \frac{1}{h}\Delta f(x_0)$$

$$f[x_0, x_1, x_2] = \frac{1}{2h}\left[\frac{\Delta f[x_1] - \Delta f[x_0]}{h}\right] = \frac{1}{2h^2}\Delta^2 f(x_0)$$

and, in general

$$f[x_0, x_1, \ldots, x_k] = \frac{1}{k!h^k}\Delta^k f(x_0).$$

This is the relation between divided and forward differences.

Newton's Backward Divided Difference

Backward difference operator is defined by

$$\nabla f(x_i) = f(x_i) - f(x_{i-1}), \qquad \text{for} \quad i \geq 1.$$

Higher powers are defined recursively by

$$\nabla^k f(x_i) = \nabla(\nabla^{k-1} f(x_i)), \qquad \text{for} \quad i \geq 2.$$

With this notation,

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

$$f[x_n, x_{n-1}, x_{n-2}] = \frac{1}{2h^2}\nabla^2 f(x_n)$$

and, in general

$$f[x_n, x_{n-1}, \ldots, x_{n-k}] = \frac{1}{k!h^k}\nabla^k f(x_n).$$

Exercises

24. *If* $f(x) = \frac{1}{x^2}$, *find the first divided differences*

    (a) $[a, b]$ (b) $[a, b, c]$.

    Here a, b, c *are arguments for* $f(x) = \frac{1}{x^2}$.

25. *Find the divided difference table for the function* $f(x) = x^2 + 2x + 2$ *whose arguments are* $1, 2, 4, 7, 10$.

26. *Find the following divided differences of* $f(x) = \frac{1}{x^2}$ *whose arguments are*

    (a) $[1, 2]$ (b) $[1, 2, 4]$ (c) $[1, 2, 4, 5]$ (d) $[2, 4, 5]$.

27. *If* $f(x) = \frac{1}{x}$ *whose arguments are* a, b, c, d *in this order, prove that*

$$[a, b, c, d] = \frac{-1}{abcd}.$$

Exercises

28. *Using Newton's divided difference formula, find the equation of the cubic curve which passes through the points* $(4, -43), (7, 83), (9, 327)$ *and* $(12, 1053)$*. Hence find* $f(10)$*.*

29. *Given the values*

x	5	7	11	13	17
$y = f(x)$	150	392	1452	2366	5202

evaluate $f(9)$*, using*

(a) *Lagrange's formula*

(b) *Newton's divided difference formula.*

30. *Find the value of* x *correct to one decimal place for which* $y = 7$ *given*

x	1	3	4
$y = f(x)$	4	12	19

31. *Tabulate* $y = x^3$ *for* $x = 2, 3, 4, 5$ *and calculate the cube root of* 10 *correct to* 3 *decimal places.*

Exercises

32. *Using the Newton's divided difference formula, evaluate* $f(8)$ *and* $f(15)$ *given*

x	4	5	7	10	11	13
$y = f(x)$	48	100	294	900	1210	2028

33. *Determine* $f(x)$ *as a polynomial in* x *for the following data.*

x	-4	-1	0	2	5
$y = f(x)$	1245	33	5	9	1335

34. *Using Newton's divided difference formula, find the missing value from the following table.*

x	1	2	4	5	6
$y = f(x)$	14	15	5	-	9

Exercises

35. *Interpolate $f(2)$ from the following data*

x	1	2	3	4	5
$f(x)$	7	?	13	21	37

and explain why the values obtained is different from the obtained by putting $x = 2$ in the expression $2^x + 5$.

36. *From the following table of yearly premiums for policies maturity at quinquennial (recurring every five years) ages, interpolate the premiums for policies maturity at the age of 12 years.*

Age (years) x	10	15	20	25	30	35
Premimum $f(x)$	3.54	3.22	2.91	2.60	2.31	2.04

37. *The population of a country is given below. Estimate the population for the year 1965.*

Year (t)	1931	1941	1951	1961	1971
Population (U_t) (in thousands)	46	66	81	93	101

Exercises

38. *The following are the marks obtained by* 492 *candidates in a certain examination.*

x	0-40	40-45	45-50	50-55	55-60	60-65
$f(x)$	210	43	54	74	32	79

Find out the number of candidates

(a) *who secured more than* 48 *but not more than* 50 *marks*

(b) *less than* 48 *but not less than* 45 *marks.*

[Hint: For the marks-range $a-b$*, define* x *as* $\frac{b-40}{5}$*.]*

39. *If* p, q, r, s *are the successive entries corresponding to equidistant arguments in a table, show that when* 3*rd differences are taken into account, the entry corresponding to the argument half-way between the arguments of* q *and* r *is* $A + \frac{1}{24}B$*, where* A *is the arthimetic mean of* q *and* r*, and* B *is the arithmetic mean of* $3q - 2p - s$ *and* $3r - 2s - p$*.*

Exercises

40. *The following are the mean temperatures (Fahreheit) on three days, 30 days apart round the periods of summer and winter. Estimate the appropriate dates and values of the maximum and minimum temperatures.*

Day	**Summer**		**Winter**	
	Date	*Temperature*	*Date*	*Temperature*
0	*15th June*	*58.8*	*16th Dec*	*40.7*
30	*15th July*	*63.4*	*15th Jan*	*38.1*
60	*14th Aug*	*62.4*	*14th Feb*	*39.3*

[Hint : Form difference tables for summer and winter separately, by considering the transformation $d(day) \to d/30 = x$*.]*

Exercises

41. *Use Newton's forward difference formula to obtain the interpolating polynomial $f(x)$ satisfying the following data.*

x	1	2	3	4
$f(x)$	26	18	4	1

If another point $x = 5, f(x) = 26$, is added to the above data, will the interpolating polynomial, be the same as before or different. Explain why?

42. *Given $\sum_{x=1}^{10} f(x) = 500426$, $\sum_{x=4}^{10} f(x) = 329240$, $\sum_{x=7}^{10} f(x) = 175212$ and $f(10) = 40365$. Find $f(1)$. [Hint : Define $u_t = \sum_{x=t}^{10} f(x)$, for $t = 1, 4, 7, 10$. Then $f(1) = u_1 - u_2$. Find u_2.]*

43. *Given $f(0) = 1, f(1) + f(2) = 10$ and $f(3) + f(4) + f(5) = 65$, find $f(4)$. [Hint: Consider $f(x) = a + bx + cx^2$.]*

44. *Find $f[3, 4, 5, 6]$ when $f(x) = x^3 - x$.*

45. *Given that $f(0) = 8, f(1) = 68, f(5)123$. Construct a divided difference table. Using the table determine the value of $f(2)$.*

Exercises

46. *Form the divided difference table and find* $f[a, b, c]$ *for* $f(x) = 1/x$.
47. *Find the polynomial of the lowest degree which assumes the values* $3, 12, 15, -21$ *when* x *has the values* $3, 2, 1, -1$ *respectively.*
48. *Given* $f(0) = -18, f(1) = 0, f(3) = 0,\ f(5) = -248,\ f(6) = 0$ *and* $f(9) = 13104$, *find the form of* $f(x)$, *assuming it to be a polynomial in* x.
49. *The values of* $f(x)$ *are known at points* x_0, x_1, x_2. *Prove that the second divided difference is equal to*

$$\frac{f(x_0)}{(x_0 - x_1)(x_0 - x_2)} + \frac{f(x_1)}{(x_1 - x_0)(x_1 - x_2)} + \frac{f(x_2)}{(x_2 - x_0)(x_2 - x_1)}.$$

Write down the equal form of nth divided difference.

50. *Show that Lagrange's formula can be evolved by equating* $(n+1)$*th divided differences of* $f(x)$ *to zero if* $f(x)$ *is a polynomial of degree* n. *[Hint:* $0 = f[x, x_0, x_1, \cdots, x_n]$ *and the previous result for nth divided difference]*

Exercises

51. *Prove that Lagrange's formula can be put in the form*

$$p_n(x) = \sum_{k=0}^{n} \frac{\pi(x) f(x_k)}{(x - x_k)\pi'(x_k)}$$

where

$$\pi(x) = \Pi_{r=0}^{n}(x - x_r) \text{ and } \pi'(x_k) = \left[\frac{d}{dx}\pi(x)\right]_{x=x_k}.$$

52. *The following values of the function $f(x)$ for values of x are given $f(1) = 4, f(2) = 5, f(7) = 5$ and $f(8) = 4$. Find the values of $f(6)$ and also the value of x for which $f(x)$ is maximum or minimum.*

53. *Apply Newton's divided difference formula (inversely) to find, to two decimal places, the value of x when $y = f(x) = 19$, from the following data.*

x	0	1	2
$f(x)$	0	1	20

Exercises

54. *The mode of a certain frequency curve* $y = f(x)$ *is very near to* $x = 9$ *and the values of the frequency density* $f(x)$ *for* $x = 8.9, 9$ *and* 9.3 *are respectively equal to* $0, 30, 0.35$ *and* 0.25*. Calculate the approximate value of mode.*

55. *The points* $(7, 3), (8, 1), (9, 1), (10, 6)$ *satisfy the function* $y = f(x)$*. Use Lagrange's interpolation formula, to find* y *for* $x = 9.5$*, and also find the interpolating polynomial.*

56. *Obtain the value of* x *for* $y = 30$ *by successive approximation method from the following data.*

x	10	12	14	16
$f(x)$	25	32	40	50

57. *Apply Lagrange's formula (inversely) to find the value of* x *when* $y = 6$*, given the following table.*

x	168	120	72	63
$f(x)$	3	7	9	10

Central Differences

We derived Newton's forward and backward interpolation formulae which are applicable for interpolation near the beginning and end of tabulated values.

The following formulae are based on central differences which are best suited for interpolation near the middle of the table.

- Gauss's forward interpolation formula
- Gauss's backward interpolation formula
- Stirling's formula
- Bessel's formula
- Laplace-Everett's formula.

The coefficients in the above central difference formulae are smaller and converge faster than those in Newton's formulae.

References

1. **Richard L. Burden and J. Douglas Faires**, *Numerical Analysis - Theory and Applications*, Cengage Learning, Singapore.
2. **Kendall E. Atkinson**, *An Introduction to Numerical Analysis*, Wiley India.
3. **David Kincaid and Ward Cheney**, *Numerical Analysis - Mathematics of Scientific Computing*, American Mathematical Society, Providence, Rhode Island.
4. **S.S. Sastry**, *Introductory Methods of Numerical Analysis*, Fourth Edition, Prentice-Hall, India.