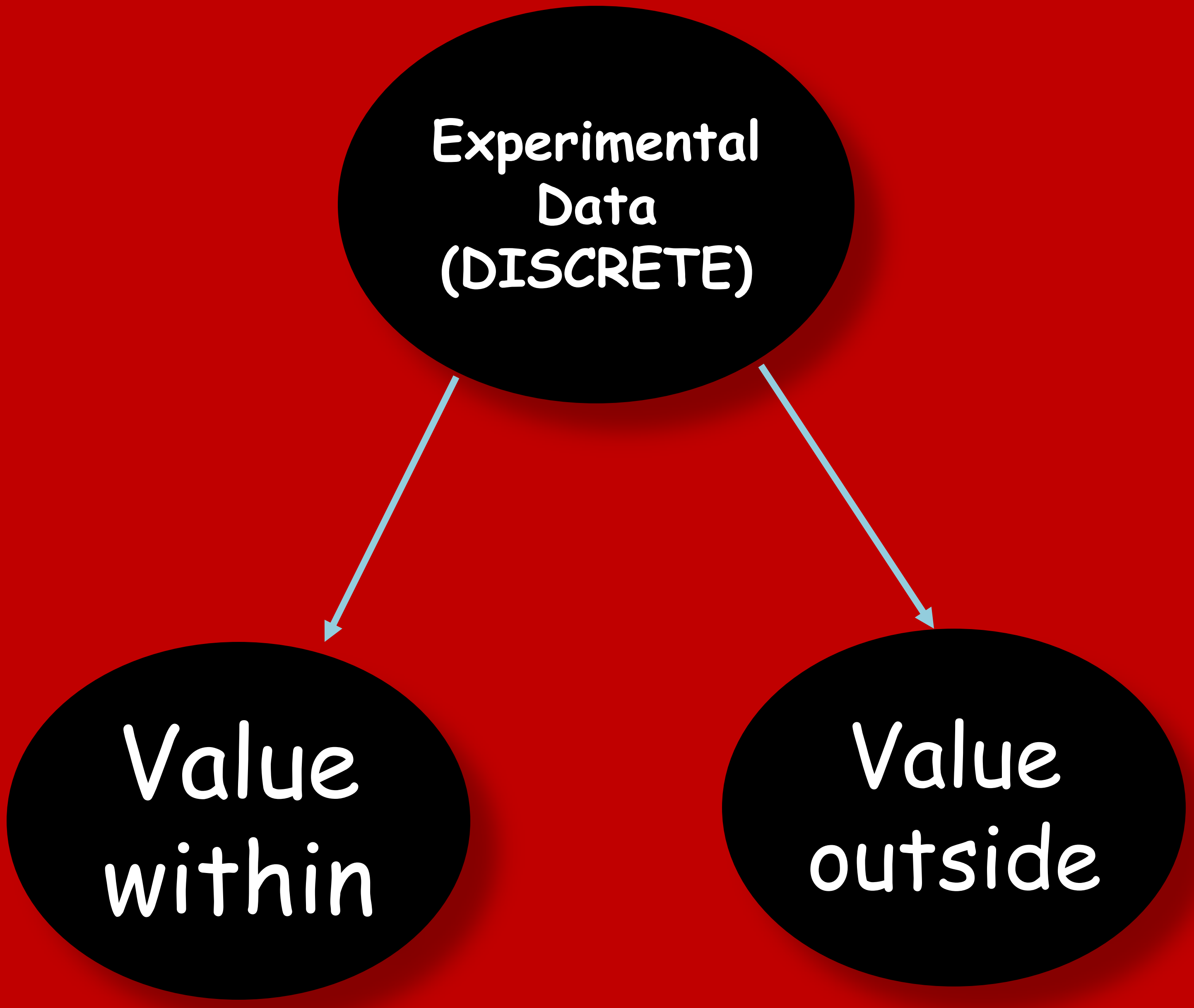
Interpolation & Extrapolation

Vinamrita Singh
Department of Physics & Astrophysics
University of Delhi

Introduction…

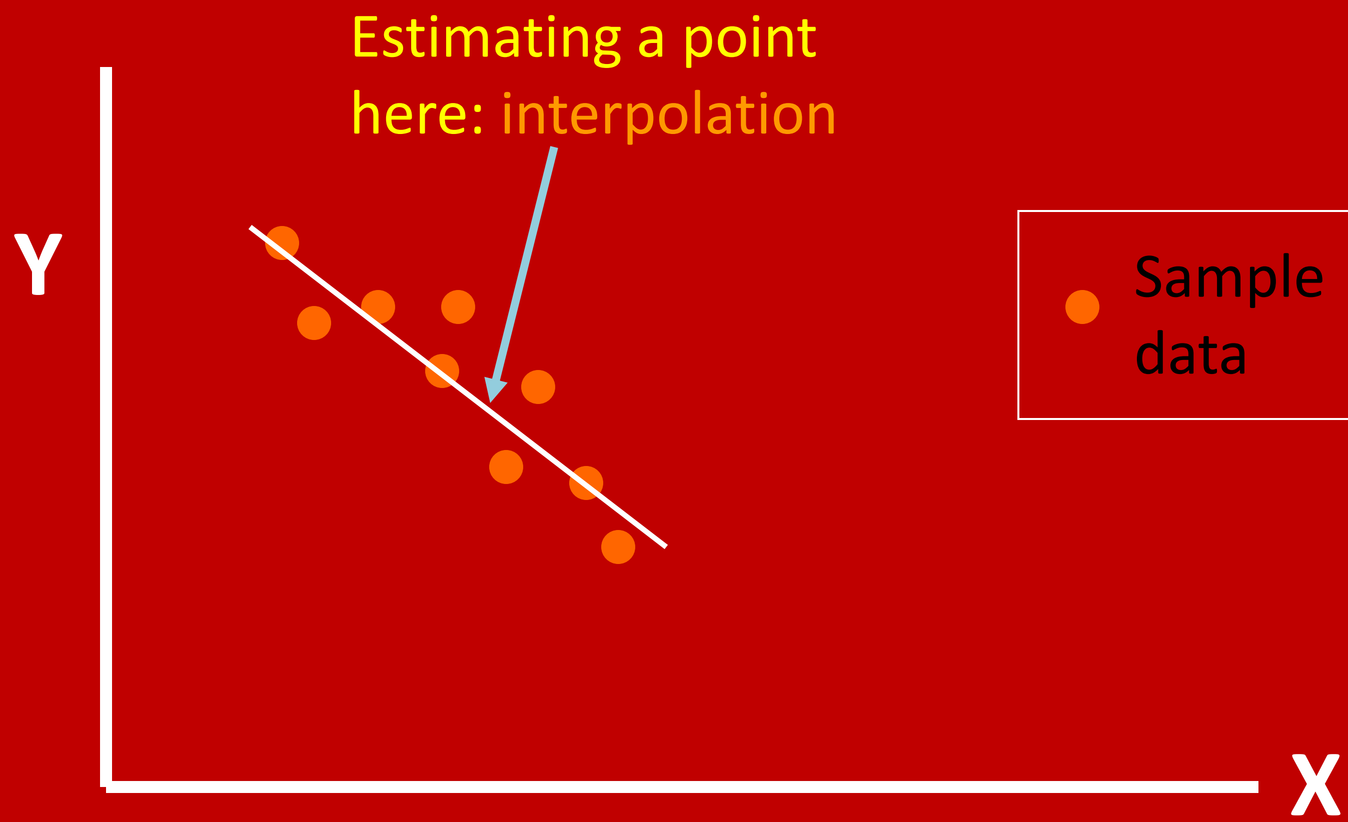
- An important part in a scientist's life is the interpretation of measured data or theoretical calculations.
- Usually when you do a measurement you will have a ***discrete*** set of points representing your experiment.
- Assume that the data is represented by pairs of values:
  - an independent variable ***"x,"*** which you vary
  - quantity ***"y,"*** which is the measured value at the point *x*.

x_0	x_1	x_2	x_3
y_0	y_1	y_2	y_3


Experimental Data (DISCRETE)
Value within
Value outside

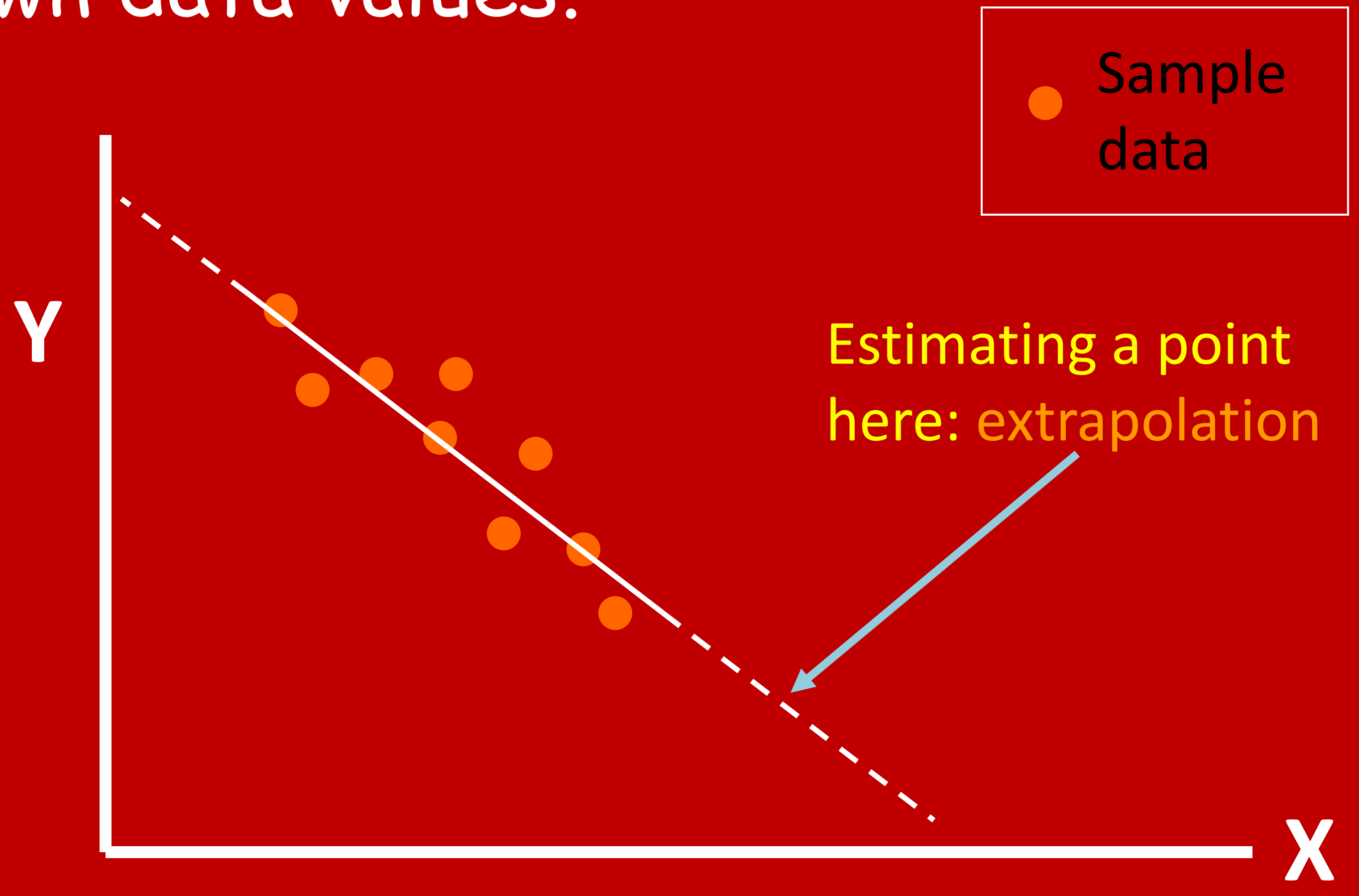
Definition: Interpolation

- ✓ Estimating the attribute values of locations that are **within** the range of available data using known data values.



Definition: Extrapolation

- ✓ Estimating the attribute values of locations **outside** the range of available data using known data values.



INTERPOLATION

Interpolation

Interpolation is carried out using approximating functions such as:

1. Polynomials
2. Trigonometric functions
3. Exponential functions
4. Fourier methods

Interpolating Polynomials

Following interpolating methods are most polular:

1. Lagrange Interpolation (unevenly spaced data)
2. Newton's Divided Difference (evenly spaced data)
3. Central difference method

LANGRANGE’S INTERPOLATION

Lagrange Polynomials

- The formula used to interpolate between data pairs $(x_0,f(x_0))$, $(x_1,f(x_1))$,…, $(x_n,f(x_n))$ is given by,

$$P(x) = \sum_{j=1}^{n} P_j(x)$$

- Where the polynomial $P_j(x)$ is given by,

$$P_j(x) = y_j \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{x - x_k}{x_j - x_k}$$

Lagrange Polynomials

EXPANDING, we get

$$P(x) = y_1 \frac{(x-x_2)(x-x_3)\dots(x-x_n)}{(x_1-x_2)(x_1-x_3)\dots(x_1-x_n)}$$

$$+ y_2 \frac{(x-x_1)(x-x_3)\dots(x-x_n)}{(x_2-x_1)(x_2-x_3)\dots(x_2-x_n)} + \dots$$

$$+ y_n \frac{(x-x_1)(x-x_2)\dots(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)\dots(x_n-x_{n-1})}$$

Lagrange Polynomials

- **Consider the table of interpolating points we wish to fit.**

i	x	y
0	x_0	$f(x_0)$
1	x_1	$f(x_1)$
2	x_2	$f(x_2)$
3	x_3	$f(x_3)$

Lagrange Polynomials

Putting in the following equation

$$P(x) = y_1 \frac{(x-x_2)(x-x_3)\ldots(x-x_n)}{(x_1-x_2)(x_1-x_3)\ldots(x_1-x_n)}$$

$$+ y_2 \frac{(x-x_1)(x-x_3)\ldots(x-x_n)}{(x_2-x_1)(x_2-x_3)\ldots(x_2-x_n)} + \ldots$$

$$+ y_n \frac{(x-x_1)(x-x_2)\ldots(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)\ldots(x_n-x_{n-1})}$$

We get:

i	x	y
0	x_0	$f(x_0)$
1	x_1	$f(x_1)$
2	x_2	$f(x_2)$
3	x_3	$f(x_3)$

- The interpolation polynomial as,

$$P(x)=\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}f(x_0)+\frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}f(x_1)$$

$$+\frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}f(x_2)+\frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}f(x_3)$$

- Note that the Lagrangian polynomial passes through each of the points used in its construction.

Advantages:

- The Lagrange formula is popular because it is easy to code.
- Also, the data are not required to be specified with *x* in ascending or descending order.
- This method can be used for ***unequally*** spaced data.

Disadvantages:

- Although the computation of $P_n(x)$ is simple, the method is still not particularly efficient for large values of n.
- When n is large and the data for x is ordered, some improvement in efficiency can be obtained by considering only the data pairs in the vicinity of the x value for which $P_n(x)$ is sought.
- The price of this improved efficiency is the possibility of a poorer approximation to $P_n(x)$.

Newton’s Formulae

Newton's Method:

1. Forward difference interpolation formula
2. Backward difference interpolation formula
3. Divided difference interpolation formula

Newton's Interpolation Method

- The nth degree polynomial may be written in the special form:

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

- If we take a_i such that $P_n(x) = f(x)$ at n+1 known points so that $P_n(x_i) = f(x_i)$, $i=0,1,\ldots,n$, then $P_n(x)$ is an interpolating polynomial.

Newton's Forward Difference Method

Let $(X_{0,},Y_{0}), (X_{1,},Y_{1}), \ldots, (X_{n,},Y_{n})$ be the given points with

$X_{i+1} = X_i + h,\ i = 0,1,2,\ldots,(n-1).$

Finite Difference Operators

- Forward difference operator
  $\Delta f(x_i) = f(x_i + h) - f(x_i)$

Forward Difference Table

i	x_i	f_i	Δf	$\Delta^2 f$	$\Delta^3 f$	Δ^4
0	x_0	f_0				
			Δf_0			
1	x_1	f_1		$\Delta^2 f_0$		
			Δf_1		$\Delta^3 f_0$	
2	x_2	f_2		$\Delta^2 f_1$		$\Delta^4 f_0$
			Δf_2		$\Delta^3 f_1$	
3	x_3	f_3		$\Delta^2 f_2$		
			Δf_3			
4	x_4	f_4				

NEWTON GREGORY FORWARD INTERPOLATION

For convenience we put $p = \frac{x - x_0}{h}$ and $f_0 = y_0$. Then we have

$$P(x_0 + ph) = y_o + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \cdots +$$

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

Example

Estimate f (3.17)from the data using Newton Forward Interpolation.

x:	3.1	3.2	3.3	3.4	3.5
f(x):	0	0.6	1.0	1.2	1.3

Solution

First let us form the difference table

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
3.1	**0**				
		0.6			
3.2	0.6		**- 0.2**		
		0.4		**0**	
3.3	1.0		- 0.2		**0.1**
		0.2		0.1	
3.4	1.2		-0.1		
		0.1			
3.5	1.3				

Here $x_0 = 3.1$, $x = 3.17$, $h = 0.1$.

Solution

$$p = \frac{x - x_0}{h} = \frac{0.07}{0.1} = 0.7$$

Newton forward formula is:

$$P(x) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \frac{p(p-1)(p-2)(p-3)}{4!}\Delta^4 y_0$$

$$P(3.17) = 0 + 0.7 \times 0.6 + \frac{0.7(0.7-1)}{2} \times (-0.2) + \frac{0.7(0.7-1)(0.7-2)}{6} \times 0 + \frac{0.7(0.7-1)(0.7-2)(0.7-3)}{24} \times 0.1$$

$$= 0.4384$$

Thus $f(3.17) = 0.4384$.

Newton's Backward Difference Method

Let (X_0, Y_0), (X_1, Y_1), ..., (X_n, Y_n) be the given points with

$X_{i+1} = X_i + h$, $i = 0, 1, 2, \ldots, (n-1)$.

Finite Difference Operators

- Backward difference operator
  $\nabla f(x_i) = f(x_i) - f(x_i - h)$

NEWTON GREGORY BACKWARD INTERPOLATION FORMULA

Taking $p = \frac{x - x_n}{h}$, we get the interpolation formula as:

$$P(x_n+ph) = y_0 + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \cdots + \frac{p(p+1)(p+2)\ldots(p+n-1)}{n!}\nabla^n y_n$$

Example

Estimate f(42) from the following data using newton backward interpolation.

x:	20	25	30	35	40	45
f(x):	354	332	291	260	231	204

Solution

The difference table is:

x	f	∇f	$\nabla^2 f$	$\nabla^3 f$	$\nabla^4 f$	$\nabla^5 f$
20	354					
		- 22				
25	332		- 19			
		- 41		29		
30	291		10		-37	
		- 31		- 8		45
35	260		2		**8**	
		- 29		**0**		
40	231		**2**			
		- 27				
45	**204**					

Here $x_n = 45$, $h = 5$, $x = 42$

and $p = -0.6$

Solution

Newton backward formula is:

$$P(x) = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_n + \frac{p(p+1)(p+2)(p+3)(p+4)}{5!}\nabla^5 y_n$$

$$P(42) = 204 + (-0.6)(-27) + \frac{(-0.6)(0.4)}{2}\times 2 + \frac{(-0.6)(0.40(1.4)}{6}\times 0 + \frac{(-0.6)(0.4)(1.4)(2.4)}{24}\times 8 + \frac{(-0.6)(0.4)(1.4)(2.4)(3.4)}{120}\times 45 = 219.1430$$

Thus, f(42) = 219.143

Newton's Divided differences

- A divided difference is defined as the difference in the function values at two points, divided by the difference in the values of the corresponding independent variable.
- Thus, the first divided difference at point is defined as

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

- Thus, the first divided difference at point is defined as

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

- The second difference is given as:

$$f[x_0, x_1, x_2] = \frac{f[x_0, x_1] - f[x_1, x_2]}{x_0 - x_2}$$

- In general,

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

Newton's Forward Divided Difference Table

x_i	f_i	$f[x_i, x_{i+1}]$	$f[x_i, x_{i+1}, x_{i+2}]$	$f[x_i, x_{i+1}, x_{i+2}, x_{i+3}]$
x_0	f_0	$f[x_0, x_1]$	$f[x_0, x_1, x_2]$	$f[x_0, x_1, x_2, x_3]$
x_1	f_1	$f[x_1, x_2]$	$f[x_1, x_2, x_3]$	$f[x_1, x_2, x_3, x_4]$
x_2	f_2	$f[x_2, x_3]$	$f[x_2, x_3, x_4]$	
x_3	f_3	$f[x_3, x_4]$		
x_4	f_4			

where

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

One with actual values.

x_i	f_i	$f[x_i, x_{i+1}]$	$f[x_i, \ldots, x_{i+2}]$	$f[x_i, \ldots, x_{i+3}]$	$f[x_i, \ldots, x_{i+4}]$
3.2	22.0	8.400	2.856	−0.528	0.256
2.7	17.8	2.118	2.012	0.0865	
1.0	14.2	6.342	2.263		
4.8	38.3	16.750			
5.6	51.7				

Example

x_i	f_i	$f[x_i, x_{i+1}]$	$f[x_i, \ldots, x_{i+2}]$	$f[x_i, \ldots, x_{i+3}]$	$f[x_i, \ldots, x_{i+4}]$
3.2	22.0	8.400	2.856	−0.528	0.256
2.7	17.8	2.118	2.012	0.0865	
1.0	14.2	6.342	2.263		
4.8	38.3	16.750			
5.6	51.7				

- The 3rd degree polynomial fitting all points from $x_0 = 3.2$ to $x_3 = 4.8$ is given by
- $P_3(x) = 22.0 + 8.400(x - 3.2) + 2.856(x - 3.2)(x - 2.7) – 0.528(x - 3.2)(x - 2.7)(x - 1.0)$
- The 4th degree polynomial fitting all points is given by
- $P_4(x) = P_3(x) + 0.256(x - 3.2)(x - 2.7)(x - 1.0)(x - 4.8)$
- The interpolated value at **x = 3.0** gives $P_3(x) = 20.2120$.

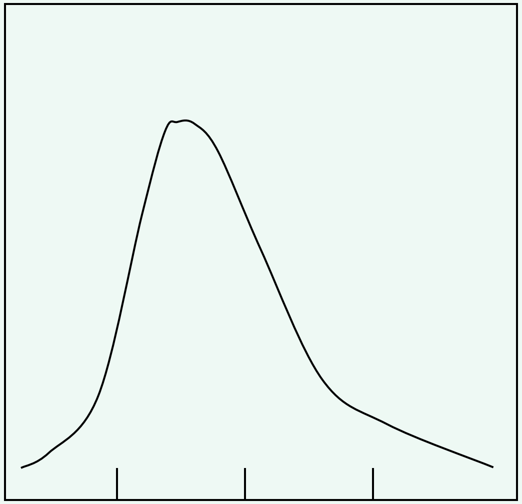
Newton's Divided differences

There are two disadvantages to using the Lagrangian interpolation polynomial for interpolation.

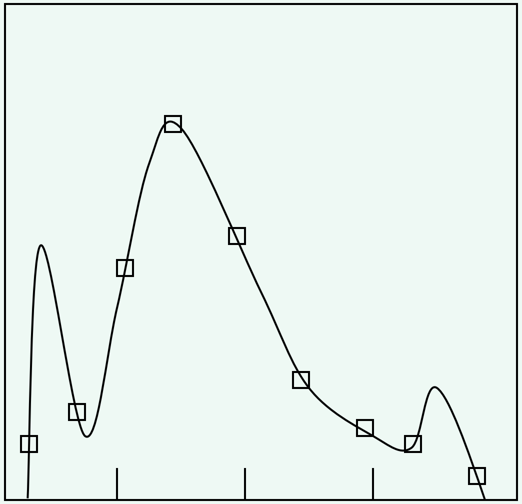
1. It involves more arithmetic operations than does the divided differences.
2. If we desire to add or subtract a point from the set to construct the polynomial, we essentially have to start over in the computations.

   The divided difference avoids this.

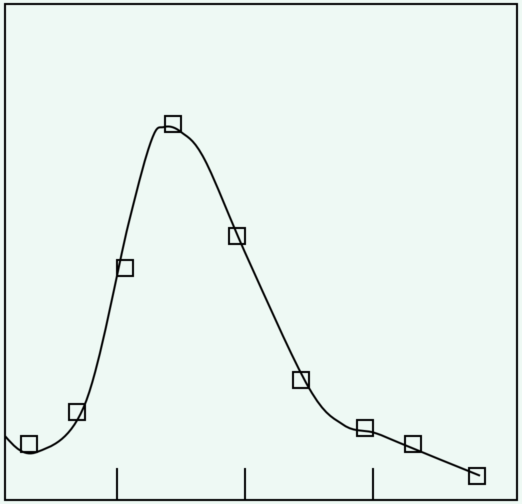
3. Tabular data have a finite number of digits. The last digit is typically rounded off. Round off has an effect on the accuracy of the higher-order differences.



True Curve



Fit using an eighth-degree polynomial



Fit using a series of 3rd degree polynomials

Although it is tempting, higher order polynomials should not be used unless there is reason to believe that using one polynomial will give a good fit.

EXTRAPOLTION

METHODS

Linear extrapolation

Polynomial extrapolation

Conic extrapolation

Linear Extrapolation

- Linear Extrapolation means creating a tangent line at the end of the known data and extending it beyond that limit.
- Linear extrapolation will provide good results only when used to extend the graph of an approximately linear function or not too far beyond the known data.
- If the two data points nearest to the point x_* to be extrapolated are (x_k, y_k) and (x_{k-1}, y_{k-1}), linear extrapolation gives the function

$$y(x_*) = y_{k-1} + \frac{x_* - x_{k-1}}{x_k - x_{k-1}}(y_k - y_{k-1}).$$

Polynomial Extrapolation

A polynomial curve can be created through the entire known data or just near the end. The resulting curve can then be extended beyond the end of the known data. Polynomial extrapolation is typically done by means of Lagrange interpolation or using Newton's method of finite differences to create a Newton series that fits the data. The resulting polynomial may be used to extrapolate the data.

Conic Extrapolation

A conic section can be created using five points near the end of the known data. If the conic section created is an ellipse or circle, it will loop back and rejoin itself. A parabolic or hyperbolic curve will not rejoin itself, but may curve back relative to the X-axis. This type of extrapolation could be done with a conic sections template (on paper) or with a computer.

Thank You