Numerical Methods

18KP2PELP2

K. NITHYA DEVI

Assistant Professor

Department of Physics

Kunthavai Naacchiyaar Government Arts College for Women (A)

Thanjavur – 613007

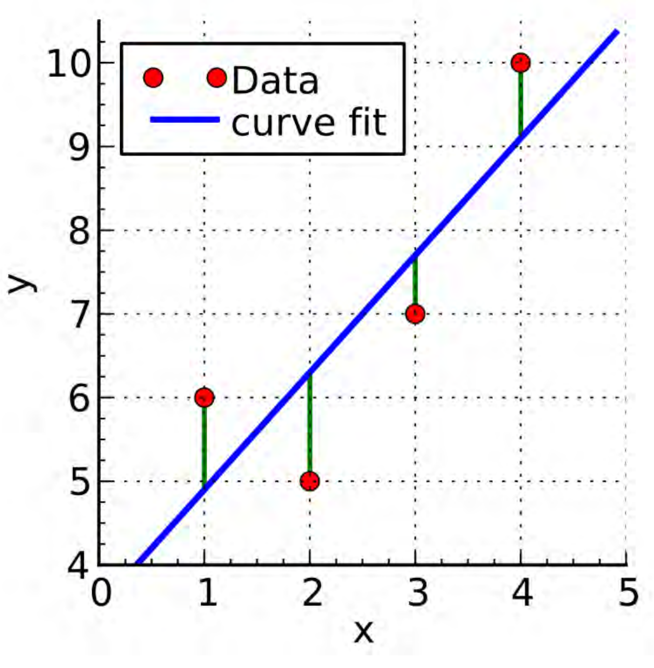
Unit – 1

Methods of curve Fitting

Curve Fitting

In regression analysis, curve fitting is the process of specifying the model that provides the best fit to the specific curves in your dataset. Curved relationships between variables are not as straightforward to fit and interpret as linear relationships. Curve fitting is one of the most powerful and most widely used analysis tools in Origin. Curve fitting examines the relationship between one or more predictors (independent variables) and a response variable (dependent variable), with the goal of defining a "best fit" model of the relationship.

The Line of Best Fit Line of best fit refers to a line through a scatter plot of data points that best expresses the relationship between those points. ... A straight line will result from a simple linear regression analysis of two or more independent variables.



Principle of least square methods

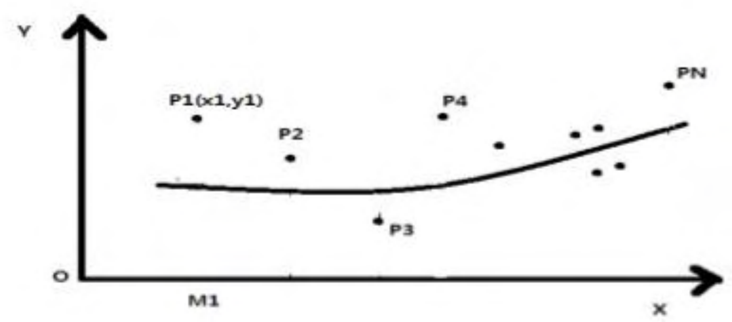
The least squares principle states that the SRF should be constructed (with the constant and slope values) so that the sum of the squared distance between the observed values of your dependent variable and the values estimated from your SRF is minimized (the smallest possible value). a method of fitting a curve to a set of points representing statistical data in such a way that the sum of the squares of the distances of the points from the curve is a minimum.

The least squares approach limits the distance between a function and the data points that the function explains. It is used in regression analysis, often in nonlinear regression modelling in which a curve is fit into a set of data. Mathematicians use the least squares method to arrive at a maximum-likelihood estimate.

After the mean for each cell is calculated, the least squares means are simply the average of these means. For treatment A, the LS mean is (3+7.5)/2 = 5.25; for treatment B, it is (5.5+5)/2=5.25. The LS Mean for both treatment groups are identical. We use the least squares criterion to pick the regression line.

The regression line is sometimes called the "line of best fit" because it is the line that fits best when drawn through the points. It is a line that minimizes the distance of the actual scores from the predicted scores.

What is a Least Squares Regression Line? fits that relationship. That line is called a Regression Line and has the equation $\hat{y} = a + b x$.
The Least Squares Regression Line is the line that makes the vertical distance from the data point



s

to the regression line as small as possible.

In order to obtain most probable values (MPVs), the sum of squares of the residuals must be minimized. (See book for derivation.) In the weighted case, the weighted squares of the residuals must be minimized.

$$\sum v^2 = v_1^2 + v_2^2 + v_3^2 + \cdots + v_n^2 = \text{minimum}$$

$$\sum wv^2 = w_1v_1^2 + w_2v_2^2 + w_3v_3^2 + \cdots + w_nv_n^2 = \text{minimum}$$

Technically the weighted form shown assumes that the measurements are independent, but we can handle the general case involving covariance.

Fitting a straight line

The equation of the straight line trend is Y= a + bX Table 4 shows that the trend value of the total publications, calculated year wise which is increasing trend during the study period. The graph of a linear equation of the form y = a + bx is a straight line. Any line that is not vertical can be described by this equation.

A line or curve of best fit on each graph. Lines of best fit can be straight or curved. Some will pass through all of the points, while others will have an even spread of points on either side.

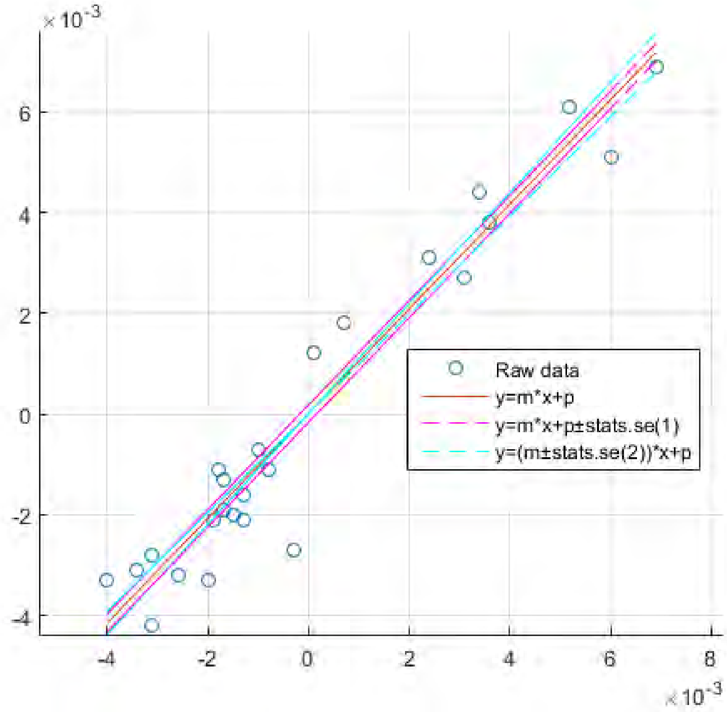
There is usually no right or wrong line, but the guidelines below will help you to draw the best. Line of best fit refers to a line through a scatter plot of data points that best expresses the relationship between those points. ... A regression involving multiple related variables can produce a curved line in some cases.

Trend forecasting gives the best forecasting reliability when the driving factors of your business affect your measures in a linear fashion. ... For example, when your historic revenue increases or decreases at a constant rate, you are seeing a linear effect.

The line of best fit is determined by the correlation between the two variables on a scatter plot. In the case that there are a few outliers (data points that are located far away from the rest of the data) the line will adjust so that it represents those points as well.

The line's slope equals the difference between points' y-coordinates divided by the difference between their x-coordinates. Select any two points on the line of best fit. These points may or may not be actual scatter points on the graph. Subtract the first point's y-coordinate from the second point's y-coordinate.

A line of best fit is drawn through a scatterplot to find the direction of an association between two variables. This line of best fit can then be used to make predictions. To draw a line of best fit, balance the number of points above the line with the number of points below the line.



1. Fit a Straight line and find x, Y values?

x_i	y_i	x_iy_i	x_i^2
1	0.5	0.5	1
2	2.5	5	4
3	2	6	9
4	4	16	16
5	3.5	17.5	25
6	6	36	36
7	5.5	38.5	49
28	**24**	**119.5**	**140**

$$\sum x_i = 28 \quad \sum y_i = 24.0$$

$$\sum x_i^2 = 140 \quad \sum x_i y_i = 119.5$$

$$\bar{x} = \frac{28}{7} = 4$$

$$\bar{y} = \frac{24}{7} = 3.428571$$

2.Calculate the value of a and b using a fit curve method

Time (t)	Yt	Ty	t^2	Trend Value
1	2	2	1	2.858
2	3	6	4	3.046
3	2	6	9	3.234
4	5	20	16	3.422
5	7	35	25	3.610
6	1	6	36	3.798
7	4	28	49	3.986
8	6	48	64	4.174
9	2	18	81	4.362
10	5	50	100	4.550
Total =55	**37**	**219**	**385**	**37.04**

$$b = \frac{n\sum ty - \sum t \sum y}{n\sum t^2 - (\sum t)^2} = \frac{10(219) - 55(37)}{10(385) - (55)^2} = \frac{2190 - 2035}{3850 - 3025} = \frac{155}{825} = 0.188$$

$$a = \frac{\sum y}{n} - \frac{b\sum t}{n} = \frac{37}{10} - \frac{(0.188)(55)}{10} = 3.7 - 1.034 = 2.67$$

To find the **best-fit line** $y = a + bx$ we minimize χ^2

$$\chi^2 = \sum_{i=1}^{N} (y_i - a - bx_i)^2/\sigma_i^2$$

We thus have to find a and b such that

$$\begin{aligned} \frac{\partial \chi^2}{\partial a} &= 0 \\ \frac{\partial \chi^2}{\partial b} &= 0 \end{aligned}$$

This gives

$$\begin{aligned} a\Sigma(1) + b\Sigma(x) - \Sigma(y) &= 0 \\ a\Sigma(x) + b\Sigma(x^2) - \Sigma(xy) &= 0 \end{aligned}$$

where we have defined

$$\begin{aligned} \Sigma(1) &= \sum_i 1/\sigma_i^2 \\ \Sigma(x) &= \sum_i x_i/\sigma_i^2 \\ \Sigma(y) &= \sum_i y_i/\sigma_i^2 \\ \Sigma(x^2) &= \sum_i x_i^2/\sigma_i^2 \\ \Sigma(xy) &= \sum_i x_i y_i/\sigma_i^2 \end{aligned}$$

Solving for a and b gives

$$\begin{aligned} b &= \frac{\Sigma(xy)\Sigma(1) - \Sigma(x)\Sigma(y)}{\Sigma(x^2)\Sigma(1) - [\Sigma(x)]^2} \\ a &= \frac{\Sigma(y) - b\Sigma(x)}{\Sigma(1)} \end{aligned}$$

Estimation of Trend by the Method of Least Squares

Q. The annual sales of a company are as follows:

Year	1991	1992	1993	1994	1995
Sales '000	45	56	58	46	75

Using the method of least squares, fit a st. line trend and estimate the annual sales of 1997.

Year	Sales y	1990 = 0 Time-Deviation x	x^2	xy	Estimated Trend'000 $Y=45 + 5x$
1991	45	1	1	45	50
1992	56	2	4	112	55
1993	78	3	9	234	60
1994	46	4	16	184	65
1995	75	5	25	375	70

$n = 5$ $\quad \Sigma\ y = 300$ $\quad \Sigma\ x = 15$ $\quad \Sigma\ x^2 = 55$ $\quad \Sigma\ xy = 950$

After the mean for each cell is calculated, the least squares means are simply the average of these means. For treatment A, the LS mean is (3+7.5)/2 = 5.25; for treatment B, it is (5.5+5)/2=5.25. The LS Mean for both treatment groups are identical. We use the least squares criterion to pick the regression line.

The regression line is sometimes called the "line of best fit" because it is the line that fits best when drawn through the points. It is a line that minimizes the distance of the actual scores from the predicted scores.

3.Calculate the value of a and b using a fit curve method

Direct Labor-Hours (x)	Factory Overhead (y)	xy	x^2
9 hours	$ 15	135	81
19	20	380	361
11	14	154	121
14	16	224	196
23	25	575	529
12	20	240	144
12	20	240	144
22	23	506	484
7	14	98	49
13	22	286	169
15	18	270	225
17	18	306	289
174 hours	$225	3,414	2,792

From the table above:

$\Sigma x = 174 \quad \Sigma y = 225 \quad \Sigma xy = 3{,}414 \quad \Sigma x^2 = 2{,}792$

$\bar{x} = \Sigma x/n = 174/12 = 14.5 \quad \bar{y} = \Sigma y/n = 225/12 = 18.75$

Substituting these values into the formula for b first:

$$b = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{(12)(3{,}414) - (174)(225)}{(12)(2{,}792) - (174)^2} = \frac{1{,}818}{3{,}228} = \underline{\underline{0.5632}}$$

$$a = \bar{y} - b\bar{x}$$

$$= (18.75) - (0.5632)(14.5) = 18.75 - 8.1664 = \underline{\underline{10.5836}}$$

Therefore, $y' = 10.5836 + 0.5632\,x$

Fitting a Parabola

A quadratic regression is the process of finding the equation of the parabola that best fits a set of data. As a result, we get an equation of the form: y=ax2+bx+c where a≠0. The best way to find this equation manually is by using the least squares method. the parabola has a downward opening.

The presumption that the axis is parallel to the y axis allows one to consider a parabola as the graph of a polynomial of degree 2, and conversely: the graph of an arbitrary polynomial of degree 2 is a parabola

The process of obtaining the equation is similar, but it is more algebraically intensive. Given the focus (h, k) and the directrix $y=mx+b$, the equation for a parabola is $(y - mx - b)^2 / (m^2 +1) = (x - h)^2 + (y - k)^2$.

When liquid is rotated, the forces of gravity result in the liquid forming a parabola-like shape. The most common example is when you stir up orange juice in a glass by rotating it round its axis.
... Parabolas are also used in satellite dishes to help reflect signals that then go to a receiver. Second degree polynomials are also known as quadratic polynomials. Their shape is known as a parabola.

If you have the equation of a parabola in vertex form $y=a(x-h)2+k$, then the vertex is at (h, k) and the focus is (h, k+14a). Notice that here we are working with a parabola with a vertical axis of symmetry, so the x-coordinate of the focus is the same as the x-coordinate of the vertex.

If you have the equation of a parabola in vertex form $y=a(x-h)2+k$, then the vertex is at (h, k) and the focus is (h, k+14a). Notice that here we are working with a parabola with a vertical axis of symmetry, so the x-coordinate of the focus is the same as the x-coordinate of the vertex.

In this equation, the vertex of the parabola is the point (h, k). You can see how this relates to the standard equation by multiplying it out: $y=a(x-h)(x-h)+ky=ax2-2ahx+ah2+k$. This means that in the standard form, $y=ax2+bx+c$, the expression $-b2a$ gives the x -coordinate of the vertex.

When liquid is rotated, the forces of gravity result in the liquid forming a parabola-like shape. The most common example is when you stir up orange juice in a glass by rotating it round its axis.
... Parabolas are also used in satellite dishes to help reflect signals that then go to a receiver. Second degree polynomials are also known as quadratic polynomials.

The regression line is sometimes called the "line of best fit" because it is the line that fits best when drawn through the points. It is a line that minimizes the distance of the actual scores from the predicted scores.

The equation of line **$y = ax^2 + bx + c$**

$$\text{Normal equations:}\quad \begin{array}{llll} a\sum x^2 & + \; b\sum x & + cn & = \sum y \\ a\sum x^3 & + \; b\sum x^2 & + c\sum x & = \sum xy \\ a\sum x^4 & + \; b\sum x^3 & + c\sum x^2 & = \sum x^2 y \end{array}$$

Solving the above equations simultaneously we get the values of a, b and c then we can write the relation between y and x as the line of equation $y = ax^2 + bx + c$.

Using Cramer's rule the value of a, b and c can be obtained as follows:

$$\Delta = \begin{vmatrix} \sum x^2 & \sum x & n \\ \sum x^3 & \sum x^2 & \sum x \\ \sum x^4 & \sum x^3 & \sum x^2 \end{vmatrix} \quad \Delta a = \begin{vmatrix} \sum y & \sum x & n \\ \sum xy & \sum x^2 & \sum x \\ \sum x^2 y & \sum x^3 & \sum x^2 \end{vmatrix}$$

$$\Delta b = \begin{vmatrix} \sum x^2 & \sum y & n \\ \sum x^3 & \sum xy & \sum x \\ \sum x^4 & \sum x^2 y & \sum x^2 \end{vmatrix} \quad \Delta c = \begin{vmatrix} \sum x^2 & \sum x & \sum y \\ \sum x^3 & \sum x^2 & \sum xy \\ \sum x^4 & \sum x^3 & \sum x^2 y \end{vmatrix}$$

$$a = \frac{\Delta a}{\Delta}, \qquad b = \frac{\Delta b}{\Delta}, \qquad c = \frac{\Delta c}{\Delta}$$

In this equation, the vertex of the parabola is the point (h, k). You can see how this relates to the standard equation by multiplying it out: y=a(x−h) (x−h) +ky=ax2−2ahx+ah2+k. This means that in the standard form, y=ax2+bx+c, the expression −b2a gives the x -coordinate of the vertex.

When liquid is rotated, the forces of gravity result in the liquid forming a parabola-like shape. The most common example is when you stir up orange juice in a glass by rotating it round its axis.
... Parabolas are also used in satellite dishes to help reflect signals that

then go to a receiver. Second degree polynomials are also known as quadratic polynomials.

The process of obtaining the equation is similar, but it is more algebraically intensive. Given the focus (h, k) and the directrix y=mx+b, the equation for a parabola is (y - mx - b) ^2 / (m^2 +1) = (x - h) ^2 + (y - k) ^2.

Problems

Find the value ofaParabola for a given table

X	1.0	1.6	2.5	4.0	6.0
Y	9.4	11.8	14.7	18.0	23.0

x	y	x^2	x^3	x^4	xy	x^2y
1.0	9.4	1.0	1.000	1.0000	9.40	9.400
1.6	11.8	2.56	4.096	6.5536	18.88	30.208
2.5	14.7	6.25	15.625	39.0625	36.75	91.875
4.0	18.0	16.00	64.00	256.0000	72.00	288.000
6.0	23.0	36.00	216.00	1296.0000	138.00	828.000
$\sum_i$ = 15.1	76.9	61.81	110.721	1598.6161	275.03	1247.483

$$b = \frac{n\sum ty - \sum t \sum y}{n\sum t^2 - (\sum t)^2} = \frac{10(219) - 55(37)}{10(385) - (55)^2} = \frac{2190 - 2035}{3850 - 3025} = \frac{155}{825} = 0.188$$

$$a = \frac{\sum y}{n} - \frac{b\sum t}{n} = \frac{37}{10} - \frac{(0.188)(55)}{10} = 3.7 - 1.034 = 2.67$$

2. Find the value ofaParabola for a given table

Direct Labor-Hours (x)	Factory Overhead (y)	xy	x^2
9 hours	$ 15	135	81
19	20	380	361
11	14	154	121
14	16	224	196
23	25	575	529
12	20	240	144
12	20	240	144
22	23	506	484
7	14	98	49
13	22	286	169
15	18	270	225
17	18	306	289
174 hours	$225	3,414	2,792

From the table above:

$\Sigma x = 174 \quad \Sigma y = 225 \quad \Sigma xy = 3{,}414 \quad \Sigma x^2 = 2{,}792$

$\bar{x} = \Sigma x/n = 174/12 = 14.5 \quad \bar{y} = \Sigma y/n = 225/12 = 18.75$

Substituting these values into the formula for b first:

$$b = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{(12)(3{,}414) - (174)(225)}{(12)(2{,}792) - (174)^2} = \frac{1{,}818}{3{,}228} = \underline{\underline{0.5632}}$$

$$a = \bar{y} - b\bar{x}$$

$$= (18.75) - (0.5632)(14.5) = 18.75 - 8.1664 = \underline{\underline{10.5836}}$$

Therefore, $y' = 10.5836 + 0.5632\,x$

Fitting an Exponential Curve

An exponential regression is the process of finding the equation of the exponential function that fits best for a set of data. As a result, we get an equation of the form y=abx where a≠0. The relative predictive power of an exponential model is denoted by R2. o finds the curve of best fit, you will need to do exponential regression. Press STAT, then right arrow to highlight CALC, and then press 0: ExpReg. The correlation coefficient is r, which is 0.994 in this case. That means that the equation is a 99.4% match to the data.

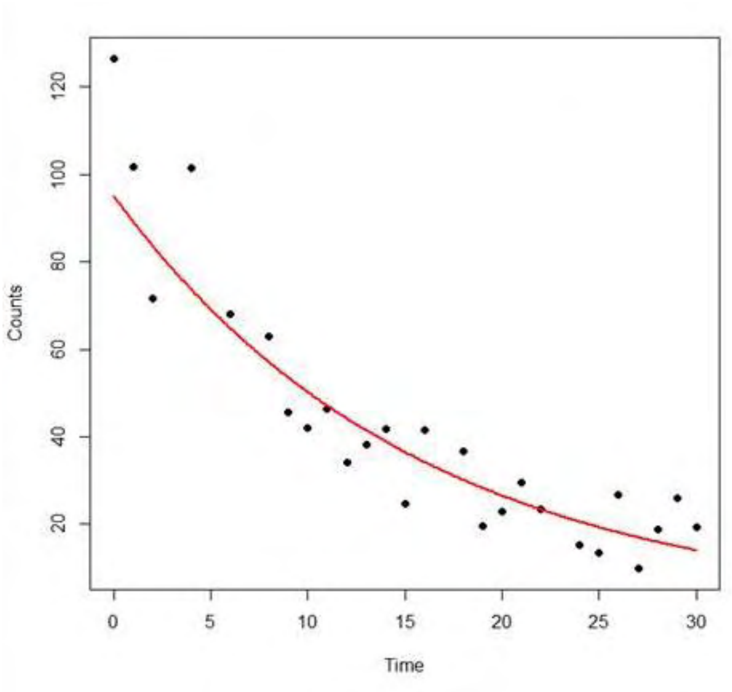
In other words, when an exponential equation has the same base on each side, the exponents must be equal. ... For example, consider the equation 34x−7=32x3 3 4 x − 7 = 3 2 x 3. To solve for x, we use the division property of exponents to rewrite the right side so that both sides have the common base 3.

Write a multiplication sign between each of the base numbers that you have just written. An exponent is a number being multiplied by itself a certain number of times, and this is what you are representing when you write the multiplication signs between base numbers. Multiply out your new equation.

Exponential equations are equations in which variables occur as exponents. For example, exponential equations are in the form ax=by. ... In other words, if the bases are the same, then the exponents must be equal

When dividing two terms with the same base, subtract the exponent in the denominator from the exponent in the numerator: Power of a Power: To raise a power to a power, multiply the exponents. The rules of exponents provide accurate and efficient shortcuts for simplifying variables in exponential notation.

Start with m=1 and n=1, then slowly increase n so that you can see 1/2, 1/3 and 1/4. Then try m=2 and slide n up and down to see fractions like 2/3 etc. Now try to make the exponent -1. Lastly try increasing m, then reducing n, then reducing m, then increasing n: the curve should go around and around.

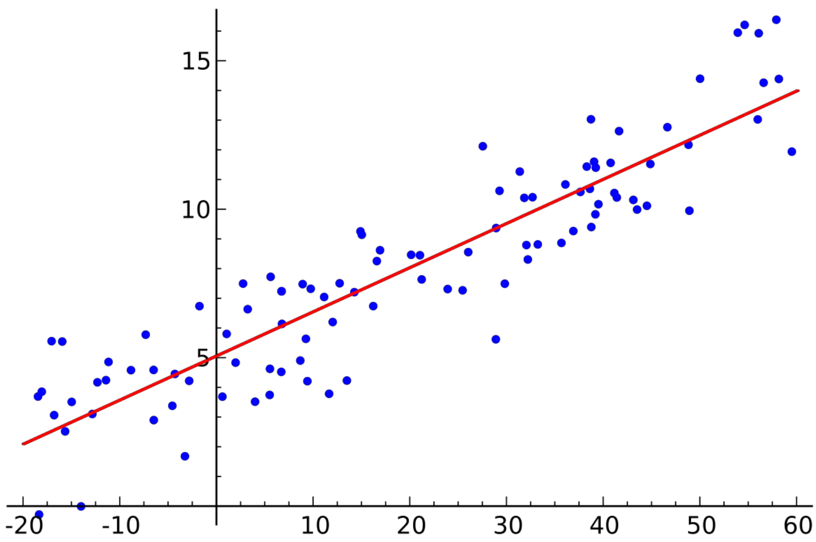


Linear Regression

Linear regression is the next step up after correlation. It is used when we want to predict the value of a variable based on the value of another variable. The variable we want to predict is called the dependent variable. For simple linear regression, the least squares estimates of the model parameters β_0 and β_1 are denoted b_0 and b_1. Using these estimates, an estimated regression equation is constructed: $\hat{y} = b_0 + b_1x$.

Linear Regression is the process of finding a line that best fits the data points available on the plot, so that we can use it to predict output values for inputs that are not present in the data set we have, with the belief that those outputs would fall on the line. Simple linear regression is a regression model that estimates the relationship between one independent variable and one dependent variable using a straight line. Both variables should be quantitative

Multiple linear regression refers to a statistical technique that uses two or more independent variables to predict the outcome of a dependent variable. The technique enables analysts to determine the variation of the model and the relative contribution of each independent variable in the total variance.



$$A = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$B = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Regression Formula:

Y= a + bX

where slope of trend line is calculated as:

$$b_1 = \frac{\sum (x - \bar{x}) * (y - \bar{y})}{\sum (x - \bar{x})^2}$$

and the intercept is computed as:

$$b_0 = y - (b_1 * X)$$

Linear regression is the next step up after correlation. It is used when we want to predict the value of a variable based on the value of another variable. The variable we want to predict is called the dependent variable. For simple linear regression, the least squares estimates of the model parameters β_0 and β_1 are denoted b_0 and b_1. Using these estimates, an estimated regression equation is constructed: $\hat{y} = b_0 + b$

Unit – IV

Numerical Integration and Differentiation

Finite Difference

A finite difference is a mathematical expression of the form $f(x + b) - f(x + a)$. If a finite difference is divided by $b - a$, one gets a difference quotient. The approximation of derivatives by finite differences plays a central role in finite difference methods for the numerical solution of differential equations, especially boundary value problems.

Certain recurrence relations can be written as difference equations by replacing iteration notation with finite differences.

Today, the term "finite difference" is often taken as synonymous with finite difference approximations of derivatives, especially in the context of numerical methods.[1][2][3] Finite difference approximations are finite difference quotients in the terminology employed above.

Three basic types are commonly considered: *forward*, *backward*, and *central* finite differences.

A forward difference is an expression of the form

Depending on the application, the spacing h may be variable or constant. When omitted, h is taken to be 1: $\Delta[f](x) = \Delta_1[f](x)$.

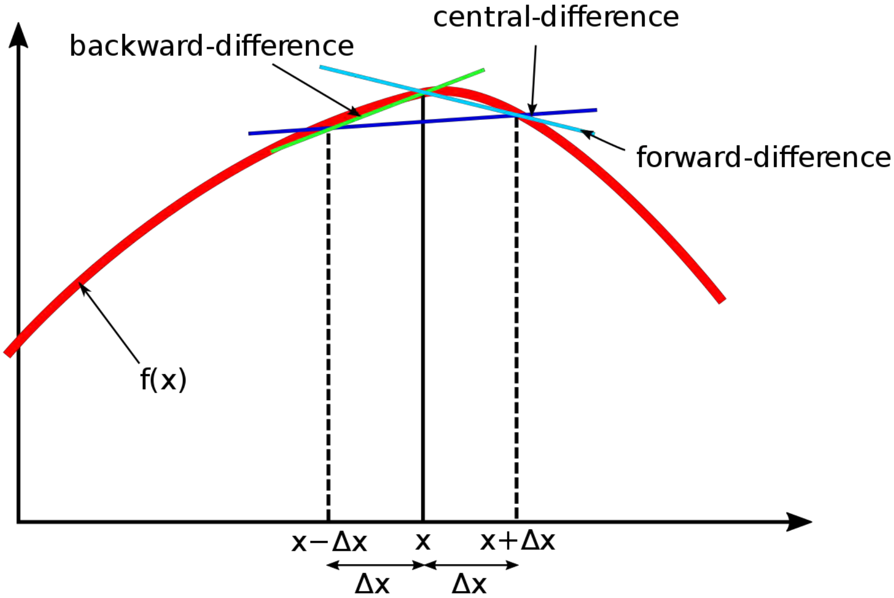
A backward difference uses the function values at x and $x - h$, instead of the values at $x + h$ and x. Finally, the central difference is given by

Finite difference is often used as an approximation of the derivative, typically in numerical differentiation.

The derivative of a function f at a point x is defined by the limit.

If h has a fixed (non-zero) value instead of approaching zero, then the right-hand side of the above equation would be written
Hence, the forward difference divided by h approximates the derivative when h is small. The error in this approximation can be derived from Taylor's theorem. Assuming that f is differentiable, we have
The same formula holds for the backward difference:

However, the central (also called centred) difference yields a more accurate approximation. If f is twice differentiable,



Newton Forward Difference Formula

Forward differences are useful in solving ordinary differential equations by single-step predictor-corrector methods (such as Euler and Runge-Kutta methods). For instance, the forward difference above predicts the value of I_1 from the derivative $I'(t_0)$ and from the value I_0.

Making use of forward difference operator and forward difference table (will be defined a little later) this scheme simplifies the calculations involved in the polynomial approximation of functions which are known at equally spaced data points.

So if we know the forward difference values of f at x_0 until order n then the above formula is very easy to use to find the function values of f at any non-tabulated value of x in the internal [a, b]. The higher order forward differences can be obtained by making use of forward difference table.

Newton's forward difference formula is a finite difference identity giving an interpolated value between tabulated points $\{f_p\}$ in terms of the first value f_0 and the powers of the forward difference Δ. For $a \in [0, 1]$, the formula states

$$f_a = f_0 + a\,\Delta + \frac{1}{2!}\,a\,(a-1)\,\Delta^2 + \frac{1}{3!}\,a\,(a-1)\,(a-2)\,\Delta^3 + \ldots. \qquad (1)$$

When written in the form

$$f(x+a)=\sum_{n=0}^{\infty}\frac{(a)_n\,\Delta^n f(x)}{n!} \tag{2}$$

with $(a)_n$ the falling factorial, the formula looks suspiciously like a finite analog of a Taylor series expansion. This correspondence was one of the motivating forces for the development of umbral calculus.

An alternate form of this equation using binomial coefficients is

$$f(x+a)=\sum_{n=0}^{\infty}\binom{a}{n}\Delta^n f(x),$$

Forward difference table

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
x_0	y_0					
		Δy_0				
x_1	y_1		$\Delta^2 y_0$			
$(= x_0 + h)$		Δy_1		$\Delta^3 y_0$		
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$	
$(= x_0 + 2h)$		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$	
$= (x_0 + 3h)$		Δy_3		$\Delta^3 y_2$		
x_4	y_4		$\Delta^2 y_3$			
$= (x_0 + 4h)$		Δy_4				
x_5	y_5					
$= (x_0 + 5h)$						

Newton Backward Difference Formula

. This is another way of approximating a function with an n^{th} degree polynomial passing through (n+1) equally spaced points. where s = $(x - x_1) / (x_1 - x_0)$ and Ñf_1 is the backward difference of f at x_1.

For interpolating the value of the function $y = f(x)$ near the end of table of values, and to extrapolate value of the function a short distance forward from y_n, Newton's backward interpolation formula is used

Derivation

Let $y = f(x)$ be a function which takes on values $f(x_n), f(x_n\text{-h}), f(x_n\text{-2h}), \ldots, f(x_0)$ corresponding to equispaced values $x_n, x_n\text{-}h, x_n\text{-}2h, \ldots, x_0$. Suppose, we wish to evaluate the function $f(x)$ at $(x_n + ph)$, where p is any real number, then we have the shift operator E, such that

$$f(x_n + ph) = E^p f(x_n) = (E^{-1})^{-p} f(x_n) = (1-\nabla)^{-p} f(x_n)$$

Binomial expansion yields,

$$f(x_n + ph) = \left[1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 + \frac{p(p+1)(p+2)}{3!}\nabla^3 + \cdots + \frac{p(p+1)(p+2)\cdots(p+n-1)}{n!}\nabla^n + \text{Error}\right] f(x_n)$$

That is

$$f(x_n + ph) = f(x_n) + p\nabla f(x_n) + \frac{p(p+1)}{2!}\nabla^2 f(x_n) + \frac{p(p+1)(p+2)}{3!}\nabla^3 f(x_n) + \cdots + \frac{p(p+1)(p+2)\cdots(p+n-1)}{n!}\nabla^n f(x_n) + \text{Error}$$

This formula is known as Newton's backward interpolation formula. This formula is also known as Newton's-Gregory backward difference interpolation formula.

If we retain $(r + 1)$terms, we obtain a polynomial of degree r agreeing with $f(x)$ at x_n, $x_{n-1}, \ldots, x_{n-r}$. Alternatively, this formula can also be written as

$$y_x = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \cdots + \frac{p(p+1)(p+2)\cdots(p+n-1)}{n!}\nabla^n y_n + \text{Error}$$

Here $\quad p = \dfrac{x - x_n}{h}$

For interpolating the value of the function $y = f(x)$ near the end of table of values, and to extrapolate value of the function a short distance forward from y_n, Newton's backward interpolation formula is used

Derivation

Let $y = f(x)$ be a function which takes on values $f(x_n), f(x_n - h), f(x_n - 2h), \ldots, f(x_0)$ corresponding to equispaced values $x_n, x_n - h, x_n - 2h, \ldots, x_0$. Suppose, we wish to evaluate the function $f(x)$ at $(x_n + ph)$, where p is any real number, then we have the shift operator E, such that

$$f(x_n + ph) = E^p f(x_n) = (E^{-1})^{-p} f(x_n) = (1 - \nabla)^{-p} f(x_n)$$

Binomial expansion yields,

$$f(x_n + ph) = \left[1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 + \frac{p(p+1)(p+2)}{3!}\nabla^3 + \cdots + \frac{p(p+1)(p+2)\cdots(p+n-1)}{n!}\nabla^n + \text{Error}\right] f(x_n)$$

That is

$$f(x_n + ph) = f(x_n) + p\nabla f(x_n) + \frac{p(p+1)}{2!}\nabla^2 f(x_n) + \frac{p(p+1)(p+2)}{3!}\nabla^3 f(x_n) + \cdots + \frac{p(p+1)(p+2)\cdots(p+n-1)}{n!}\nabla^n f(x_n) + \text{Error}$$

This formula is known as Newton's backward interpolation formula. This formula is also known as Newton's-Gregory backward difference interpolation formula.

If we retain $(r + 1)$terms, we obtain a polynomial of degree r agreeing with $f(x)$ at x_n, $x_{n-1}, \ldots, x_{n-r}$. Alternatively, this formula can also be written as

$$y_x = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \cdots + \frac{p(p+1)(p+2)\cdots(p+n-1)}{n!}\nabla^n y_n + \text{Error}$$

Here $\quad p = \dfrac{x - x_n}{h}$

Newton's backward difference formula is a finite difference identity giving an interpolated value between tabulated points $\{f_p\}$ in terms of the first value f_0 and the powers of the forward difference Δ. For $a \in [0, 1]$, the formula states

$$f_a = f_0 + a\,\Delta + \frac{1}{2!}\,a\,(a-1)\,\Delta^2 + \frac{1}{3!}\,a\,(a-1)\,(a-2)\,\Delta^3 + \ldots. \qquad (1)$$

When written in the form

$$f(x+a)=\sum_{n=0}^{\infty}\frac{(a)_n\,\Delta^n f(x)}{n!} \qquad (2)$$

with $(a)_n$ the falling factorial, the formula looks suspiciously like a finite analog of a Taylor series expansion. This correspondence was one of the motivating forces for the development of umbral calculus.

An alternate form of this equation using binomial coefficients is

$$f(x+a)=\sum_{n=0}^{\infty}\binom{a}{n}\Delta^n f(x),$$

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$	$\nabla^4 y$	$\nabla^5 y$
x_0	y_0					
		∇y_1				
x_1	y_1		$\nabla^2 y_2$			
$(= x_0 + h)$		∇y_2		$\nabla^3 y_3$		
x_2	y_2		$\nabla^2 y_3$		$\nabla^4 y_4$	
$(= x_0 + 2h)$		∇y_3		$\nabla^3 y_4$		$\nabla^5 y_5$
x_3	y_3		$\nabla^2 y_4$		$\nabla^4 y_5$	
$(= x_0 + 3h)$		∇y_4		$\nabla^3 y_5$		
x_4	y_4		$\nabla^2 y_5$			
$(= x_0 + 4h)$		∇y_5				
x_5	y_5					
$(= x_0 + 5h)$						

Problems:

1. Find the difference table for the given table

x	0	1	2	3
F(x)	1	2	11	34

Sol.

Given

x	0	1	2	3
$f(x)$	1	2	11	34

Find $f(2.8)$

Since the required value 2.8 is at the end of the table, apply Newton's backward interpolation formula.

$x_n + nh = 2.8 \Rightarrow 3 + n\,(1) = 2.8$

$\Rightarrow n = 2.8 - 3 = -0.2$

$[\because h = 1]$

The difference table is

x	y	∇y	$\nabla^2 y$	$\nabla^3 y$
0	1			
1	2	1		
2	11	9	8	
3	34	23	14	6

Newton's backward interpolation formula is

$$y_{(x = x_n + nh)} = y_n + \frac{n}{1!}\nabla y_n + \frac{n(n+1)}{2!}\nabla^2 y_n + \frac{n(n+1)(n+2)}{3!}\nabla^3 y_n$$

$$\Rightarrow y(2.8) = 34 + (-0.2)\,(23) + \frac{(-0.2)(-0.2+1)}{2}\,(14) + \frac{(-0.2)(-0.2+1)(-0.2+2)}{6}\,(16)$$

$$\Rightarrow y(2.8) = 34 - 4.6 + (-0.2)\,(0.8)\,(7) + (-0.2)\,(0.8)\,(1.8)$$

$$\Rightarrow y(2.8) = 34 - 4.6 - 1.12 - 0.288$$

$$\Rightarrow y(2.8) = 27.992$$

Problem-02: Using Newton's divided difference estimate $f(x)$ from the following table.

x	-1	0	2	3	4
$f(x)$	-16	-7	-1	8	29

Solution: The divided difference table is as follows:

x	$f(x)$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$	$\Delta^4 f(x)$
-1	-16				
		$\frac{-7+16}{0+1}=9$			
0	-7		$\frac{3-9}{2+1}=-2$		
		$\frac{-1+7}{2-0}=3$		$\frac{2+2}{3+1}=1$	
2	-1		$\frac{9-3}{3-0}=2$		0
		$\frac{8+1}{3-1}=9$		$\frac{6-2}{4-0}=1$	
3	8		$\frac{21-9}{4-2}=6$		
		$\frac{29-8}{4-3}=21$			
4	29				

Here, $x_0=-1,\ x_1=0,\ x_2=2,\ x_3=3,\ x_4=4$

$$f(x_0)=-16,\ f(x_0,x_1)=9,\ f(x_0,x_1,x_2)=-2,\ f(x_0,x_1,x_2,x_3)=1$$

By Newton's divided difference formula we get,

$$f(x)=f(x_0)+(x-x_0)f(x_0,x_1)+(x-x_0)(x-x_1)f(x_0,x_1,x_2)+(x-x_0)(x-x_1)(x-x_2)f(x_0,x_1,x_2,x_3)$$
$$\cdots\cdots+(x-x_0)(x-x_1)(x-x_2)\cdots\cdots(x-x_{n-1})f(x_0,x_1,x_2,\cdots\cdots x_n)$$

$$=-16+(x+1)\times 9+x(x+1)\times(-2)+x(x+1)(x-2)\times 1$$

$$=-16+9x+9-2x^2-2x+x^3-2x^2+x^2-2x$$

$$\therefore\ f(x)=x^3-3x^2+5x-7 \quad (Ans.)$$

Problem-03: Using Lagrange's formula estimate $f(x)$ from the following table.

x	0	1	2	5
$f(x)$	2	3	12	147

Solution: Here, $x_0=0,\ x_1=1,\ x_2=2,\ x_3=5$

$$f(x_0)=2,\ f(x_1)=3,\ f(x_2)=12,\ f(x_3)=147$$

By Lagrange's formula we get,

$$f(x)=\frac{(x-x_1)(x-x_2)\cdots\cdots(x-x_n)}{(x_0-x_1)(x_0-x_2)\cdots\cdots(x_0-x_n)}f(x_0)+\frac{(x-x_0)(x-x_2)\cdots\cdots(x-x_n)}{(x_1-x_0)(x_1-x_2)\cdots\cdots(x_1-x_n)}f(x_1)+\frac{(x-x_0)(x-x_1)\cdots\cdots(x-x_n)}{(x_2-x_0)(x_2-x_1)\cdots\cdots(x_2-x_n)}f(x_2)$$
$$\cdots\cdots+\frac{(x-x_0)(x-x_1)\cdots\cdots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\cdots\cdots(x_n-x_{n-1})}f(x_n)$$

$$=\frac{(x-1)(x-2)(x-5)}{(0-1)(0-2)(0-5)}\times 2+\frac{(x-0)(x-2)(x-5)}{(1-0)(1-2)\cdots\cdots(1-5)}\times 3+\frac{(x-0)(x-1)(x-5)}{(2-0)(2-1)(2-5)}\times 12+\frac{(x-0)(x-1)(x-2)}{(5-0)(5-1)(5-2)}\times 147$$

$$=\frac{x^3-8x^2+17x-10}{-10}\times 2+\frac{x^3-7x^2+10x}{4}\times 3+\frac{x^3-6x^2+5x}{-6}\times 12+\frac{x^3-3x^2+2x}{60}\times 147$$

$$=\frac{-x^3+8x^2-17x+10}{5}+\frac{3x^3-21x^2+30x}{4}-2x^3+12x^2-10x+\frac{49x^3-147x^2+98x}{20}$$

Problem-05: The population of a town in the last six censuses was as given below. Estimate the population for the year 1946.

Year(x)	1911	1921	1931	1941	1951	1961
Population in thousands(y)	12	15	20	27	39	52

Solution: Since $x = 1946$ is nearer at the end of the set of given tabular values, so we have to construct the backward difference table

The backward difference table of the given data is as follows:

Year(x)	*Populations*(y)	∇	∇^2	∇^3	∇^4	∇^5
1911	12					
		3				
1921	15		2			
		5		0		
1931	20		2		3	
		7		3		**-10**
1941	27		5		**-7**	
		12		**-4**		
1951	39		**1**			
		13				
1961	**52**					

Here $x = 1946,\ x_n = 1961,\ h = 10,\ y_n = 52,\ \nabla y_n = 13,\ \nabla^2 y_n = 1,\ \nabla^3 y_n = -4,\ \nabla^4 y_n = -7, \nabla^5 y_n = -10$.

$$\therefore\ p = \frac{x - x_n}{h} = \frac{1946 - 1961}{10} = -\frac{15}{10} = -1.5$$

By Newton's backward formula we get,

$$y(x) = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n \cdots\cdots + \frac{p(p+1)\cdots\cdots(p+n-1)}{n!}\nabla^n y_n$$

$$\therefore\ y(1946) = 52 + (-1.5)\times 13 + \frac{(-1.5)(-1.5+1)}{2!}\times 1 + \frac{(-1.5)(-1.5+1)(-1.5+2)}{3!}\times(-4) + \frac{(-1.5)(-1.5+1)(-1.5+2)(-1.5+3)}{4!}\times(-7)$$
$$+ \frac{(-1.5)(-1.5+1)(-1.5+2)(-1.5+3)(-1.5+4)}{5!}\times(-10)$$

$$= 32.3438\,(approx.)$$

Trapezoidal rule

Trapezoidal Rule is a rule that evaluates the area under the curves by dividing the total area into smaller trapezoids rather than using rectangles. This integration works by approximating the region under the graph of a function as a trapezoid, and it calculates the area.

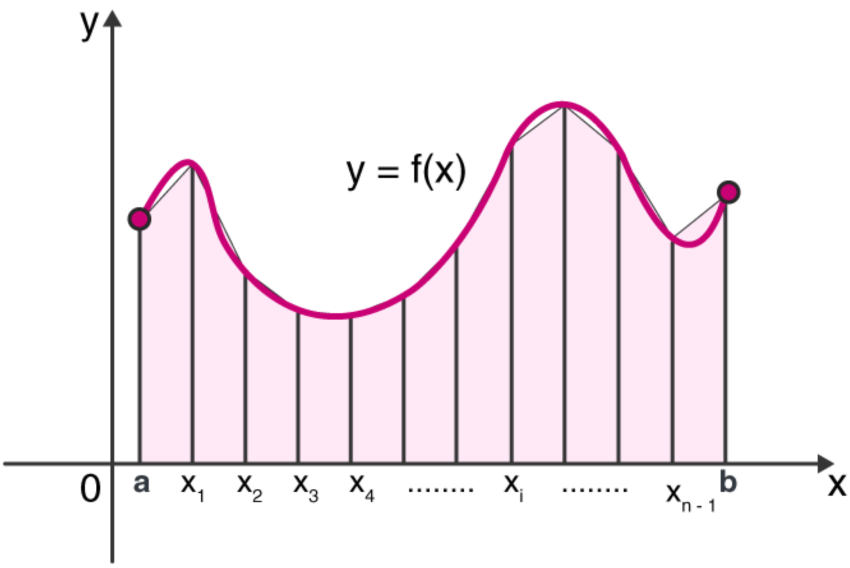
By now you know that we can use Riemann sums to approximate the area under a function. ... Key idea: By using trapezoids (aka the "trapezoid rule") we can get more accurate approximations than by using rectangles. t follows that if the integrand is concave up (and thus has a positive second derivative), then the error is negative and the trapezoidal rule overestimates the true value. The trapezoidal rule is not as accurate as Simpson's Rule when the underlying function is smooth, because Simpson's rule uses quadratic approximations instead of linear approximations. The formula is usually given in the case of an odd number of equally spaced points.

The Midpoint rule is always more accurate than the Trapezoid rule. ... For example, make a function which is linear except it has nar row spikes at the midpoints of the subdivided intervals. Then the approximating rectangles for the midpoint rule will rise up to the level of the spikes, and be a huge overestimate. The trapezoidal rule tends to overestimate the value of a definite integral systematically over intervals where the function is concave up and to underestimate the value of a definite integral systematically over intervals where the function is concave down.

Write A=h(b1+b2)/2, where A represents the trapezoid's area, b1 represents one of the base lengths, b2 represents the other base length and h represents the height. Rearrange the equation to get h alone. Multiply both sides of the equation by 2 to get. 2A=h(b1+b2).
By dividing the interval [a, b] into many smaller intervals, and applying the trapezoidal rule to each, this allows us to find a better approximation the integral. In order to find the length for one of the two equivalent nonparallel legs of the trapezoid (side), first use the height of the trapezoid to form right triangles on the interior of the trapezoid that each have a base length.

the graph of a function as a trapezoid, and it calculates the area.

By now you know that we can use Riemann sums to approximate the area under a function. ... Key idea: By using trapezoids (aka the "trapezoid rule") we can get more accurate approximations than by using rectangles. t follows that if the integrand is concave up (and thus has a positive second derivative), then the error is negative and the trapezoidal rule overestimates the true value. The trapezoidal rule is not as accurate as Simpson's Rule when the underlying function is smooth, because Simpson's rule uses quadratic approximations instead of linear approximations. The formula is usually given in the case of an odd number of equally spaced points.



$$\int_a^b f(x)dx$$
$$\approx\left(\frac{b-a}{n}\right)\left[\frac{1}{2}(y_1+y_2)+\frac{1}{2}(y_2+y_3)+\frac{1}{2}(y_3+y_4)+..\frac{1}{2}(y_n+y_{n+1})\right]$$
$$\approx\left(\frac{b-a}{n}\right)\left[\frac{1}{2}(y_1+y_{n+1})+(y_2+y_3+y_4+\ldots+y_n)\right]$$

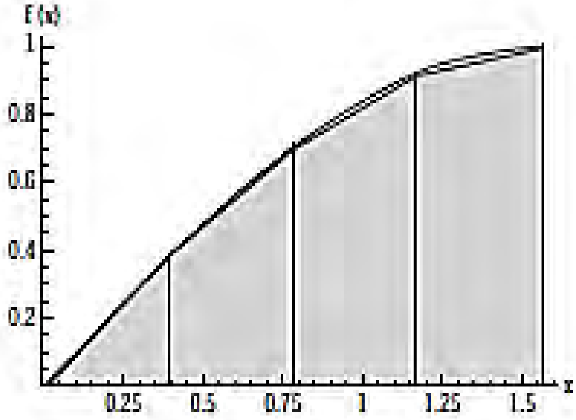
The Midpoint rule is always more accurate than the Trapezoid rule. ... For example, make a function which is linear except it has nar row spikes at the midpoints of the subdivided intervals. Then the approximating rectangles for the midpoint rule will rise up to the level of the spikes, and be a huge overestimate. The trapezoidal rule tends to overestimate the value of a definite integral systematically over intervals where the function is concave up and to underestimate the value of a definite integral systematically over intervals where the function is concave down.

By using trapezoids (aka the "trapezoid rule") we can get more accurate approximations than by using rectangles. t follows that if the integrand is concave up (and thus has a positive second derivative), then the error is negative and the trapezoidal rule overestimates the true value. The trapezoidal rule is not as accurate as Simpson's Rule when the underlying function is smooth, because Simpson's rule uses quadratic approximations instead of linear approximations. The formula is usually given in the case of an odd number of equally spaced points.

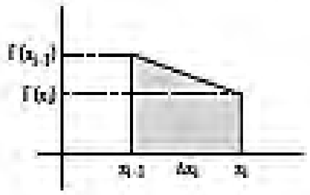
$$\int_{x_0}^{x_n} f(x)dx = \sum_{i=1}^{n} S_i = \frac{h}{2}(f(x_0)+f(x_1)) + \frac{h}{2}(f(x_1)+f(x_2)) + \dots$$
$$+ \frac{h}{2}(f(x_{n-1})+f(x_n)) = \frac{h}{2}\left[f(x_0)+f(x_n)+\sum_{i=1}^{n-1} f(x_i) \right]$$

we can get more accurate approximations than by using rectangles. t follows that if the integrand is concave up (and thus has a positive second derivative), then the error is negative and the trapezoidal rule overestimates the true value. The trapezoidal rule is not as accurate as Simpson's Rule when the underlying function is smooth, because Simpson's rule uses quadratic approximations instead of linear approximations. The formula is usually given in the case of an odd number of equally spaced points.

The Midpoint rule is always more accurate than the Trapezoid rule. ... For example, make a function which is linear except it has nar row spikes at the midpoints of the subdivided intervals. Then the approximating rectangles for the midpoint rule will rise up to the level of the spikes, and be a huge overestimate. The trapezoidal rule tends to overestimate the value of a definite integral systematically over intervals where the function is concave up and to underestimate the value of a definite integral systematically over intervals where the function is concave down.



Trapezoidal approximation to $\int_0^{\pi/2} \sin x \, dx$ using four trapezoids.



The area enclosed by a trapezoid with base Δx and sides A and is $\frac{\Delta x}{2}(A+\quad)$. Thus, the area enclosed by the trapezoid constructed in the ith interval, $[x_{i-1}, x_i]$, is $\frac{\Delta x_i}{2}\left[f(x_{i-1}) + f(x_i)\right]$.

The total trapezoidal area, obtained by adding the individual areas, is

$$\frac{\Delta x_1}{2}[f(x_0) + f(x_1)] + \frac{\Delta x_2}{2}[f(x_1) + f(x_2)] + \frac{\Delta x_3}{2}[f(x_2) + f(x_3)] + \cdots + \frac{\Delta x_n}{2}[f(x_{n-1}) + f(x_n)]$$

If all intervals have the same length, Δx, this reduces to

$$\frac{\Delta x}{2}\Big[[f(x_0) + f(x_1)] + [f(x_1) + f(x_2)] + [f(x_2) + f(x_3)] + \cdots + [f(x_{n-1}) + f(x_n)]\Big]$$

or

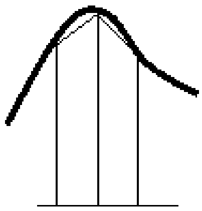
$$\frac{\Delta x}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + 2f(x_3) + \cdots + 2f(x_{n-1}) + f(x_n)\right]$$

By using trapezoids (aka the "trapezoid rule") we can get more accurate approximations than by using rectangles. t follows that if the integrand is concave up (and thus has a positive second derivative), then the error is negative and the trapezoidal rule overestimates the true value. The trapezoidal rule is not as accurate as Simpson's Rule when the underlying function is smooth, because Simpson's rule uses quadratic approximations instead of linear approximations.

The trick is how to minimize discretization & roundoff errors!

+ Infinite sums take infinite time to evaluate
+ Roundoff errors put lower bounds on "h"

The usual approach is to cleverly fit the integrand to a form which you can integrate analytically, e.g.

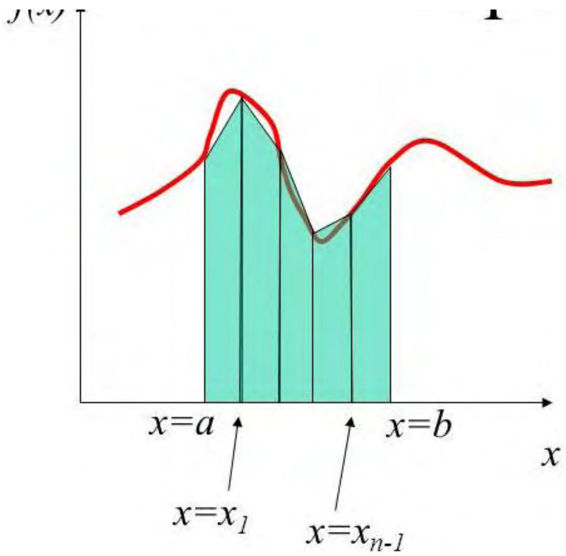


$$f(x) \cong \begin{cases} m_1 x + b_1 & \text{if } a < x < x_1 \\ m_2 x + b_2 & \text{if } x_1 < x < x_2 \\ \ldots & \ldots \end{cases}$$

$$m_1 = \frac{f(x_1) - f(a)}{x_1 - a} \qquad b_1 = f(x_1) - m_1 x_1$$

$$\int_a^{x_1} m_1 x + b_1 \, dx = \left(\frac{m_1 x^2}{2} + b_1 x\right)\Big|_a^{x_1} = \frac{(f(x_1) + f(a))(x_1 - a)}{2}$$

Trapezoidal Rule: $I = \int_a^b f(x)dx \cong h(f(a) + f(b))/2 + h\sum f(x_i)$



The rectangular rule can be made more accurate by using trapezoids to replace the rectangles as shown. A linear approximation of the function locally sometimes work much better than using the averaged value like the rectangular rule does.

$$\int_a^b f(x)dx = \frac{\Delta x}{2}[f(a) + f(x_1)] + \frac{\Delta x}{2}[f(x_1) + f(x_2)] + .. + \frac{\Delta x}{2}[f(x_{n-1}) + f(b)]$$

$$= \Delta x[\frac{1}{2} f(a) + f(x_1) + .. f(x_{n-1}) + \frac{1}{2} f(b)]$$

Example—Evaluate $\int_{-3}^{3} x^4 dx$ by using Trapezoidal rule. Verify result by actual integration.
Solution—given that $f(x) = x^4$
Interval length $(b - a) = (3 - (-3)) = 6$
So we divide 6 equal intervals with $h = 6/6 = 1.0$
And tabulate the values as below

x	-3	-2	-1	0	1	2	3
$y = x^4$	81	16	1	0	1	16	81
	y_0	y_1	y_2	y_3	y_4	y_5	y_6

We know that—

$$\int_a^b f(x)dx \approx \frac{h}{2}[(y_0 + y_n) + 2(y_1 + y_2 + \cdots + y_{n-1})]$$

$$\int_{-3}^{3} x^4 dx \approx \frac{1}{2}[(81 + 81) + 2(16 + 1 + 0 + 1 + 16)] = \frac{162 + 68}{2} = 115$$

By actual integration $\int_{-3}^{3} x^4 dx = \left[\left(\frac{3^5}{5}\right) - \left(-\frac{3^5}{5}\right)\right] = \left[\frac{243}{5} + \frac{243}{5}\right] = \frac{486}{5} = 97.5$

Example— Evaluate $\int_0^1 \frac{1}{(1+x^2)} dx$by using Trapezoidal rule with h = 0.2.
Solution— Given $f(x) = \frac{1}{(1+x^2)}$ and interval length $(b - a) = (1 - 0) = 1$.
So we divide 6 equal intervals with h= 0.2

x	0	0.2	0.4	0.6	0.8	1
$y = \frac{1}{(1 + x^2)}$	1	0.96154	0.86207	0.73529	0.60976	0.5000

We know $\int_a^b f(x)dx \approx \frac{h}{2}[(y_0 + y_n) + 2(y_1 + y_2 + \cdots + y_{n-1})]$

$$\int_0^1 \frac{1}{(1 + x^2)} dx \approx \frac{0.2}{2}[(1 + 0.5000) + 2(0.96154 + 0.86207 + 0.73529 + 0.60976)]$$

$$=(0.1)[\ (1.05) + 6.33732\ ]$$
$$= 0.783732$$

Problem:

Derive the trapezoidal rule of integration using method of undetermined coefficients.

Solution:

Approximate an integral by the formula given below with undetermined coefficients, c_1 and c_2 .

$$\int_a^b f(x)dx \approx c_1 f(a) + c_2 f(b)$$

We now are required to find these coefficients, c_1 and c_2 .

We choose, $f(x) = a_0 + a_1 x$ for which the formula is exact to determine c_1 and c_2 .

From exact integral calculus, we get

$$\int_a^b f(x)dx = \int_a^b (a_0 + a_1 x)dx$$

$$= a_0(b-a) + a_1\left(\frac{b^2 - a^2}{2}\right) \quad (1)$$

and from the formula, we get

$$c_1 f(a) + c_2 f(b) = c_1(a_0 + a_1 a) + c_2(a_0 + a_1 b)$$

$$= a_0(c_1 + c_2) + a_1(c_1 a + c_2 b) \quad (2)$$

So for the formula to give the same result as from the exact integral calculus, we equate equations (1) and (2)

$$a_0(b-a) + a_1(\frac{b^2 - a^2}{2}) = a_0(c_1 + c_2) + a_1(c_1 a + c_2 b)$$

We now have one equation and two unknowns. But, since the left and right hand sides have to be equal for arbitrary values of a_0 and a_1, that is possible only if the coefficients of a_0 and a_1 have to be equal.

$$c_1 + c_2 = b - a \quad (3)$$

$$c_1 a + c_2 b = \frac{b^2 - a^2}{2} \quad (4)$$

Multiply equation (3) by a gives

$$c_1 a + c_2 a = ab - a^2 \tag{5}$$

Subtracting equation (5) from equation (4) gives

$$(c_1 a + c_2 b) - (c_1 a + c_2 a) = \frac{b^2 - a^2}{2} - (ab - a^2)$$

$$c_2 (b - a) = \frac{b^2 - a^2 - 2ab + 2a^2}{2}$$

$$= \frac{b^2 + a^2 - 2ab}{2}$$

$$= \frac{(b-a)^2}{2}$$

$$c_2 = \frac{b-a}{2} \tag{6}$$

Using the obtained value of c_2 from equation (6) and substituting it in equation (3) gives

$$c_1 + \frac{b-a}{2} = b - a$$

$$c_1 = (b-a) - \frac{b-a}{2}$$

$$= \frac{2(b-a) - (b-a)}{2}$$

$$= \frac{b-a}{2} \tag{7}$$

Hence from equations (6) and (7), the approximate formula

$$\int_a^b f(x)dx \approx c_1 f(a) + c_2 f(b)$$

becomes

$$\int_a^b f(x)dx \approx \frac{b-a}{2} f(a) + \frac{b-a}{2} f(b)$$

This is same as the trapezoidal rule.

Example:

Use the Trapezoidal Rule with $n = 4$ to estimate $\int_1^2 \frac{dx}{x}$

Solution:

Each subinterval $= \frac{(b-a)}{n} = \frac{(2-1)}{4} = \frac{1}{4}$

$\therefore \quad x_0 = 1, \quad x_1 = 1\frac{1}{4}, \quad x_2 = 1\frac{1}{2}, \quad x_3 = 1\frac{3}{4}, \quad x_4 = 2$

$$y_0 = f(x_0) = \frac{1}{1} = 1$$

$$y_1 = f(x_1) = \frac{1}{\left(1\frac{1}{4}\right)} = \frac{4}{5}$$

$$y_2 = f(x_2) = \frac{1}{\left(1\frac{1}{2}\right)} = \frac{2}{3}$$

$$y_3 = f(x_3) = \frac{1}{\left(1\frac{3}{4}\right)} = \frac{4}{7}$$

$$y_4 = f(x_4) = \frac{1}{2}$$

Using: $\int_a^b f(x)dx = \left(\frac{b-a}{n}\right)\left(\frac{1}{2}y_0 + y_1 + y_2 + \ldots + y_{n-1} + \frac{1}{2}y_n\right)$

$$\int_1^2 \frac{dx}{x} = \left(\frac{2-1}{4}\right)\left(\frac{1}{2}(1) + \frac{4}{5} + \frac{2}{3} + \frac{4}{7} + \frac{1}{2}\left(\frac{1}{2}\right)\right)$$

$$\int_1^2 \frac{dx}{x} = \left(\frac{1}{4}\right)\left(\frac{1}{2} + \frac{4}{5} + \frac{2}{3} + \frac{4}{7} + \frac{1}{4}\right)$$

$$\int_1^2 \frac{dx}{x} = \frac{1}{8} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{16}$$

$$\int_1^2 \frac{dx}{x} = \approx 0.697023809$$

Simpson Rule

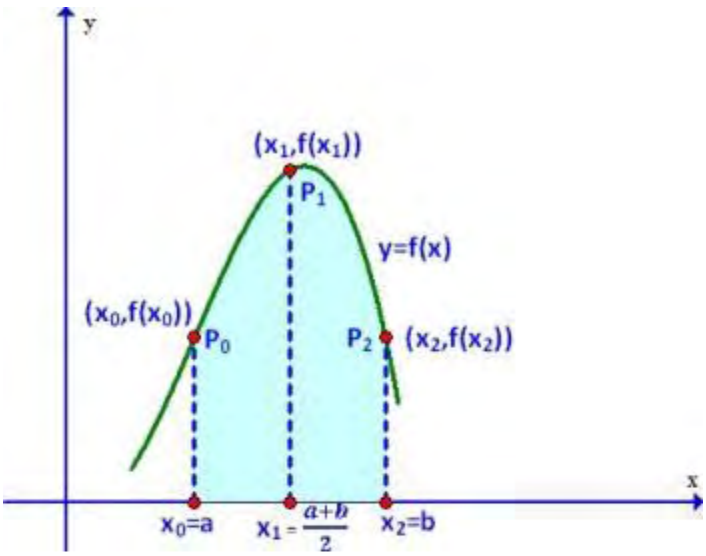
In numerical analysis, Simpson's 1/3 rule is a method for numerical approximation of definite integrals. Specifically, it is the following approximation: In Simpson's 1/3 Rule, we use parabolas to approximate each part of the curve. We divide. the area into n equal segments of width Δx. If we have f(x) = y, which is equally spaced between [a, b] and if a = x_0, x1 = x_0 + h, x2 = x_0 + 2h …., x_n = x_0 + nh, where h is the difference between the terms. Or we can say that $y_0 = f(x_0)$, $y_1 = f(x_1)$, $y_2 = f(x_2)$…,$y_n = f(x_n)$ are the analogous values of y with each value of x.

In Simpson's Rule, we will use parabolas to approximate each part of the curve. This proves to be very efficient since it's generally more accurate than the other numerical methods we've seen. (See more about Parabolas.) We divide the area into n equal segments of width Δx.

The Approximate Int(f(x), x = a... b, method = Simpson [3/8], opts) command approximates the integral of f(x) from a to b by using Simpson's 3/8 rule. This rule is also known as Newton's 3/8 rule. The first two arguments (function expression and range) can be replaced by a definite integral. Simpson's 3/8 rule, also called Simpson's second rule requests one more function evaluation inside the integration range, and is exact if f is a polynomial up to cubic degree. Simpson's 1/3 and 3/8 rules are two special cases of closed Newton–Cotes formulas.

The midpoint rule approximates the definite integral using rectangular regions whereas the trapezoidal rule approximates the definite integral using trapezoidal approximations. Simpson's rule approximates the definite integral by first approximating the original function using piecewise quadratic functions.

Its strength is that, although rectangles and trapezoids work better for linear functions, Simpson's Rule works quite well on curves. Simpson's Rule is based on the fact that given any three points, you can find the equation of a quadratic through those points. Whereas the main advantage of the Trapezoid rule is its rather easy conceptualization and derivation, Simpson's rule 2 Page 3 approximations usually achieve a given level of accuracy faster. Moreover, the derivation of Simpson's rule is only marginally more difficult.



The Approximate Int(f(x), x = a... b, method = Simpson [3/8], opts) command approximates the integral of f(x) from a to b by using Simpson's 3/8 rule. This rule is also known as Newton's 3/8 rule. The first two arguments (function expression and range) can be replaced by a definite integral. Simpson's 3/8 rule, also called Simpson's second rule requests one more function evaluation inside the integration range, and is exact if f is a polynomial up to cubic degree. Simpson's 1/3 and 3/8 rules are two special cases of closed Newton–Cotes formulas.

In Simpson's Rule, we will use parabolas to approximate each part of the curve. This proves to be very efficient since it's generally more accurate than the other numerical methods we've seen. (See more about Parabolas.) We divide the area into n equal segments of width Δx.

The Approximate Int(f(x), x = a... b, method = Simpson [3/8], opts) command approximates the integral of f(x) from a to b by using Simpson's 3/8 rule. This rule is also known as Newton's 3/8 rule. The first two arguments (function expression and range) can be replaced by a definite integral. Simpson's 3/8 rule, also called Simpson's second rule requests one more function evaluation inside the integration range, and is exact if f is a polynomial up to cubic degree. Simpson's 1/3 and 3/8 rules are two special cases of closed Newton–Cotes formulas.

Example:

Use Simpson's Rule with $n = 4$ to estimate $\int_1^2 \frac{dx}{x}$

Solution:

$$\text{Each subinterval} = \frac{(b-a)}{n} = \frac{(2-1)}{4} = \frac{1}{4}$$

$$\therefore \quad x_0 = 1, \quad x_1 = 1\frac{1}{4}, \quad x_2 = 1\frac{1}{2}, \quad x_3 = 1\frac{3}{4}, \quad x_4 = 2$$

$$y_0 = f(x_0) = \frac{1}{1} = 1$$

$$y_1 = f(x_1) = \frac{1}{\left(1\frac{1}{4}\right)} = \frac{4}{5}$$

$$y_2 = f(x_2) = \frac{1}{\left(1\frac{1}{2}\right)} = \frac{2}{3}$$

$$y_3 = f(x_3) = \frac{1}{\left(1\frac{3}{4}\right)} = \frac{4}{7}$$

$$y_4 = f(x_4) = \frac{1}{2}$$

Using:

$$\int_a^b f(x)dx = \frac{1}{3}\left(\frac{b-a}{n}\right)\left(y_0 + 4y_1 + 2y_2 + \ldots + 2y_{n-2} + 4y_{n-1} + y_n\right)$$

$$\int_1^2 \frac{dx}{x} = \frac{1}{3}\left(\frac{2-1}{4}\right)\left(1 + (4)\frac{4}{5} + (2)\frac{2}{3} + (4)\frac{4}{7} + \frac{1}{2}\right)$$

$$\int_1^2 \frac{dx}{x} = \left(\frac{1}{3}\right)\left(\frac{1}{4}\right)\left(1 + \frac{16}{5} + \frac{4}{3} + \frac{16}{7} + \frac{1}{2}\right)$$

2. Calculate the value of the Simpson rule for the below problem

$$\int_0^1 \frac{1}{1+x^2}\, dx \approx \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + y_4)$$

$$n = 4, \quad h = \frac{1-0}{4} = 0 \cdot 25$$

x	0	$0 \cdot 25$	$0 \cdot 5$	$0 \cdot 75$	1
y	1	$0 \cdot 94118$	$0 \cdot 8$	$0 \cdot 64$	$0 \cdot 5$

$$\int_0^1 \frac{1}{1+x^2}\, dx \approx \frac{0 \cdot 25}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + y_4)$$

$$\approx 0 \cdot 7854 \text{ (4 d.p.)}$$

3.Find the value of simpson rule for the given table

Estimate $\int_0^{2.5} f(x)\, dx$ from the data using Simpson's rule

x	0	0.5	1	1.5	2	2.5
$f(x)$	1.5	2	2	1.6364	1.25	0.9565

$h = 1/2$ Simpson's 3/8 rule Simpson's 1/3 rule

Because the number of panels is odd, we compute the integral over the first three intervals by Simpson's 3/8 rule, and use the 1/3 rule for the last two intervals:

$$I_S = \frac{3(0.5)}{8}[f(0) + 3f(0.5) + 3f(1) + f(1.5)]$$

$$+ \frac{0.5}{3}[f(1.5) + 4f(2) + f(2.5)]$$

$$= 2.8381 + 1.2655 = 4.1036$$

References

1. Eugene Barkov, 2007 Numerical methods, Addison Wesley, London
2. L.A. Pipes and L.R. Harvilli,2009 Applied Mathematics for Engineering and Physicists, Fiber Khanna publishers, New Delhi.
3. J. Millikan & Chalcis Tata 2001, Numerical methods for Engineering, McGraw Hill- New Delhi.
4. A.K. Ghattack, T.C. Goyal, Numerical methods 2005, Prentice, Hall Ltd

Source

1. https://scihub.wikicn.top/