4. Numerical Differentiation and Integration

4.1 Numerical Differentiation

The method of obtaining the derivative of a function using a numerical technique is known as numerical differentiation. The general method for deriving the numerical differentiation formulae is to differentiate the interpolating polynomial. There are essentially two situations where numerical differentiation is required. They are:

1. The function values are known but the function is unknown. Such functions are called tabulated function.
2. The function to be differentiated is continuous and therefore complicated and difficult to differentiate.

Since, analytical methods give exact answers; the numerical techniques provide only approximations to derivatives. Numerical differentiation methods are very sensitive to round off errors, in addition to the truncation error introduced by the methods of themselves. So, it is necessary to discuss the errors and ways to minimize them.

The given set of values of x and y, we shall derive formula to compute,

$\frac{dy}{dx}, \frac{d^2y}{dx^2}$, ... for any value of x in $[x_0, x_n]$.

Differentiating Tabulated Function

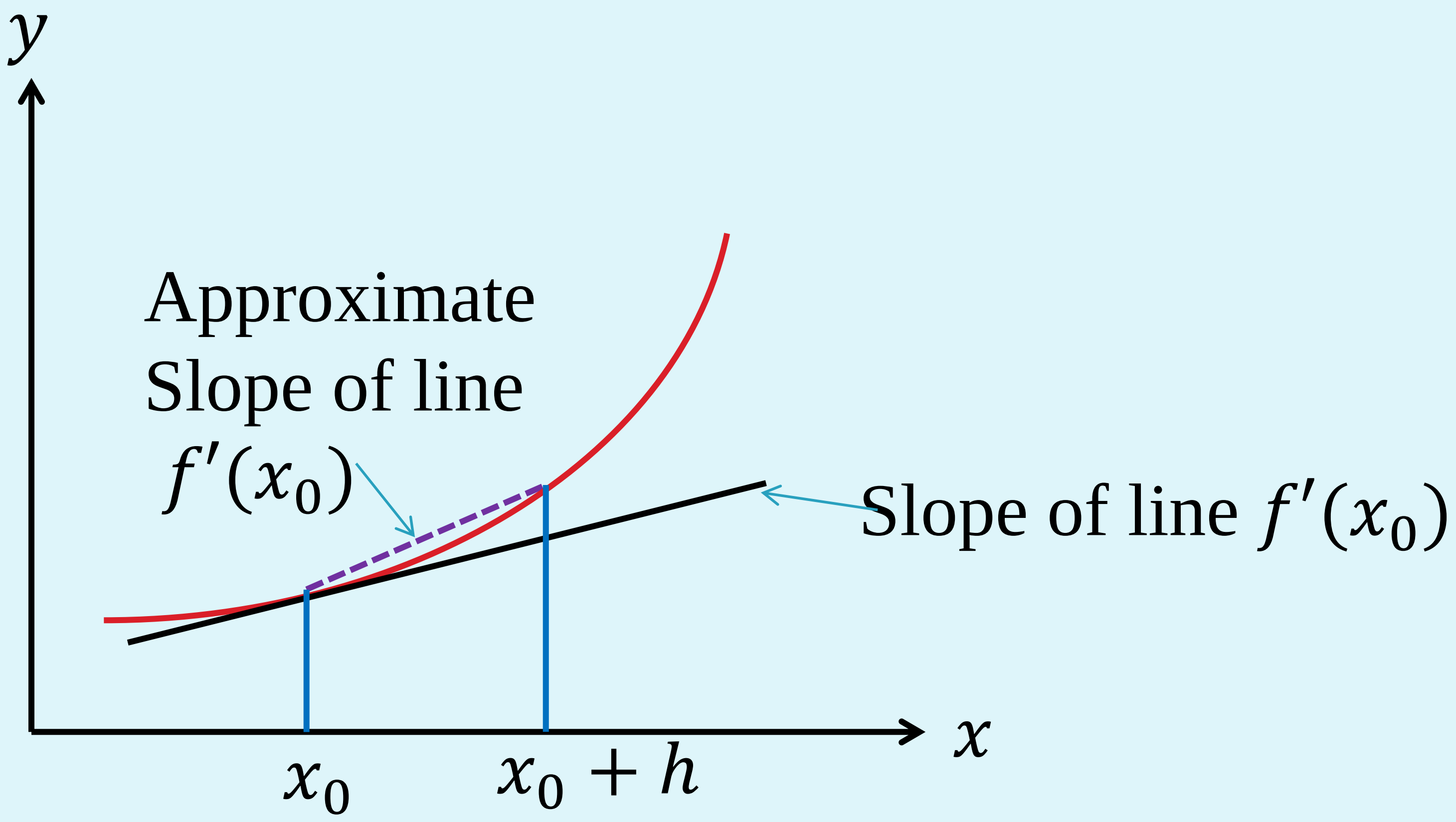
The given set of values of x and y, we shall derive formula to compute,

$\frac{dy}{dx}, \frac{d^2y}{dx^2}, \ldots$ for any value of x in $[x_0, x_n]$.

4.1.1 First Derivatives

Under this section we are mainly focused to approximate the slope of a curve $f(x)$ at a particular point $x = x_0$, in terms of $f(x_0)$ and the value of $f(x)$ at a nearby point where $x = (x_0 + h)$.

The shorter broken line in following figure may be thought of as giving a reasonable approximation to the required slope, if h is small enough.



Therefore,

$$\boldsymbol{f'(x_0) = slope\ of\ short\ broken\ line}$$
$$\boldsymbol{= \frac{f(x_0 + h) - f(x_0)}{h}}$$

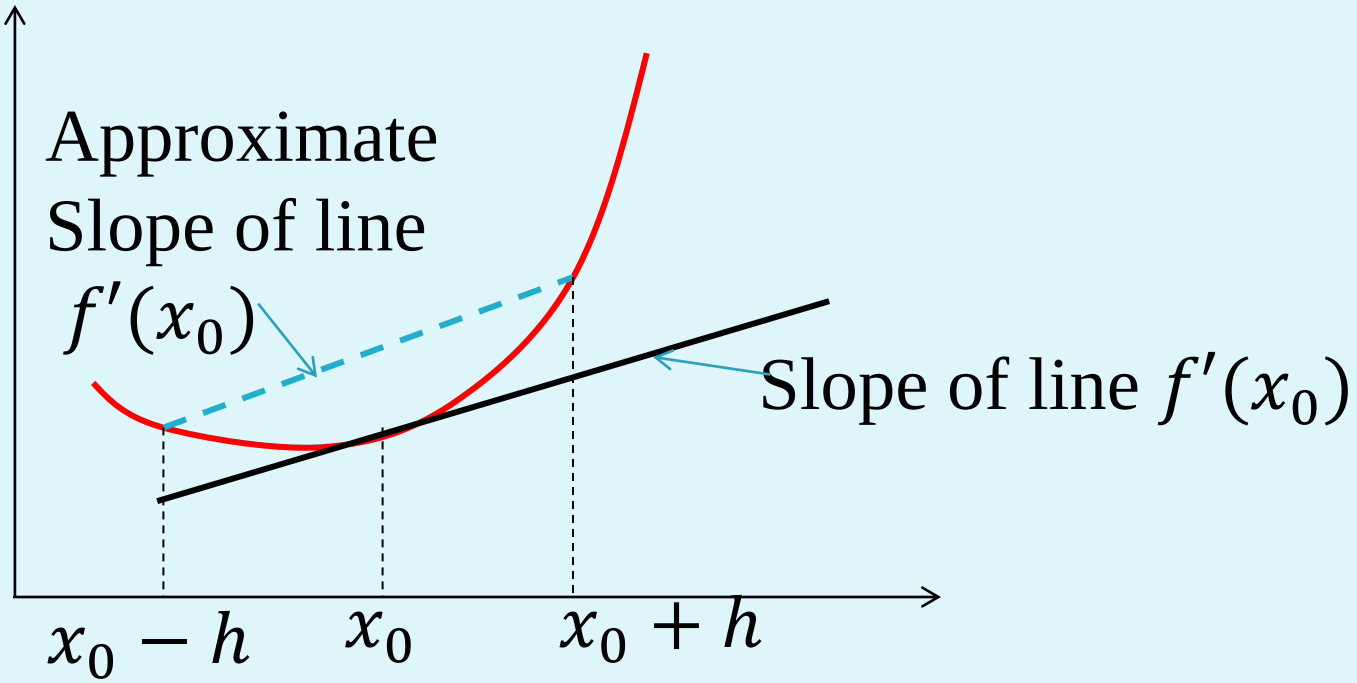
This is called **forward difference** approximation to the derivative of $f(x)$. A second version of this arises on considering a point to the left of x_0, rather than to the right as we did above.

In this case we obtain the approximation,

$$\boldsymbol{f'(x_0) = \frac{f(x_0) - f(x_0 - h)}{h}}$$

and this is called **backward difference** approximation to $f'(x_0)$.

A third method of approximating the first derivative of $f(x)$ can be seen bellow.



Here we approximate as,

$$\boldsymbol{f'(x_0) = slope\ of\ short\ broken\ line}$$
$$\boldsymbol{= \frac{f(x_0+h) - f(x_0-h)}{2h}}$$

and this is called **central difference** approximation to $f'(x_0)$.

First derivative approximation summary

Three approximations to the derivative $f'(x_0)$ are,

i. Forward difference, $\boldsymbol{f'(x_0) = \frac{f(x_0+h)-f(x_0)}{h}}$

ii. Backward difference, $\boldsymbol{f'(x_0) = \frac{f(x_0)-f(x_0-h)}{h}}$

iii. Central difference, $\boldsymbol{f'(x_0) = \frac{f(x_0+h)-f(x_0-h)}{2h}}$

Example 1:

Let $f(x) = e^x$ and using forward difference, backward difference and central differences approximate $f'(4)$ using $h = 0.1$ and 0.05.

Example 2:

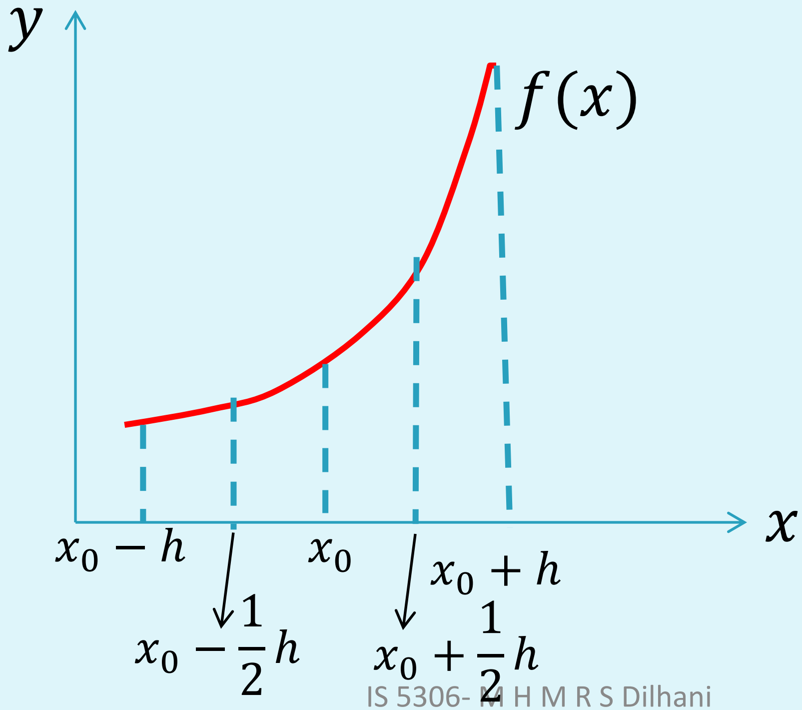
Use a forward difference, and the values of h shown, to approximate the derivative of $\cos(x)$ at $x = \pi/3$.

i. $h = 0.1$
ii. $h = 0.01$
iii. $h = 0.001$
iv. $h = 0.0001$

4.1.2 Second Derivatives

An approach which has been found to work well for second derivative involves applying the notation of a central difference three times. We begin with,

$$f''(x_0) = [f'(x_0 + 1/2\, h) - f'(x_0 - 1/2\, h)]/h$$



Next we approximate the two derivatives in the numerator of this expression using central differences as follows:

$$f'\left(x_0 + \frac{1}{2}h\right) = \frac{f(x_0+h) - f(x_0)}{h} \text{ and}$$

$$f'\left(x_0 - \frac{1}{2}h\right) = \frac{f(x_0) - f(x_0 - h)}{h}$$

Combining these three results gives,

$$f''(x_0)$$
$$= \frac{1}{h}\left\{\left(\frac{f(x_0+h) - f(x_0)}{h}\right) - \left(\frac{f(x_0) - f(x_0 - h)}{h}\right)\right\}$$

$$\therefore f''(x_0) = \frac{f(x_0+h) - 2f(x_0) + f(x_0 - h)}{h^2}$$

Example1:

The distance x of a runner from a fixed point is measures and the obtained data are as follows.

$t/(s)$	0.0	0.5	1.0	1.5	2.0
$x/(m)$	0.00	3.65	6.70	9.80	11.15

Use central difference to approximate the runner's acceleration at $t = 1.5\ s$.

Example 2:

The specific heat capacity(amount of heat required to change the temperature of a given body by given amount) is an important element in thermodynamics process. For a process in which the pressure is constant, the specific heat capacity, C_p, equals the slope of the relationship between specific enthalpy (a measure of the total energy of a thermodynamic system), H, and the temperature, T, as follows:

$$C_p = \frac{dH}{dT}$$

Specific enthalpy, H, data obtained with respect to the temperature T are shown bellow.

$\boldsymbol{T/(^0F)}$	**800**	**1000**	**1200**	**1400**	**1600**
$H/(Btu/lb)$	1205	1360	1485	1605	1725

By using central difference method, calculate the specific heat capacity (C_p) and the rate of change of C_p at the temperature, $1200\ ^0\boldsymbol{F}$.

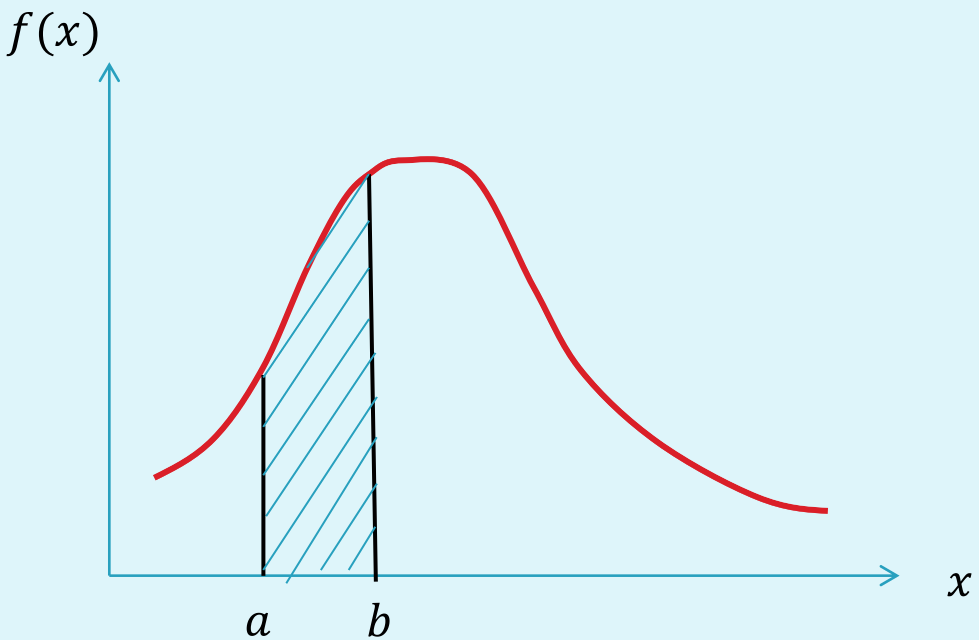
4.2 Numerical Integration

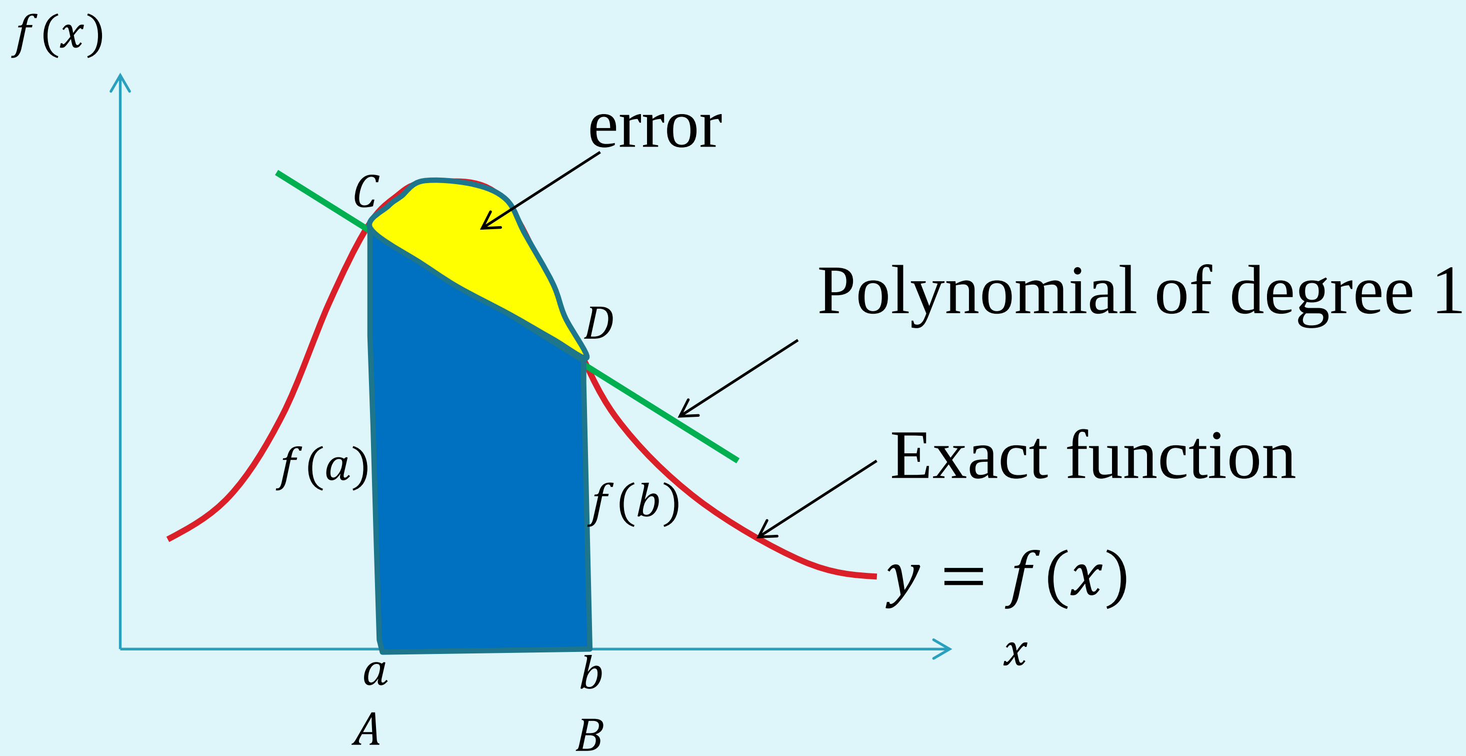
The general objective of numerical integration is to compute the value of the definite integral given below for a given set of data points $(x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$ of a function $y = f(x)$, where $f(x)$ is not know explicitly. So, $f(x)$ is determined using the interpolation technique and then integration is carried out.

Like in numerical differentiation, we need to evaluate the numerical integration in the following cases.

1. Functions do not possess closed form solutions. Example: $f(x) = C\int_0^x e^{-t^2}\,dt$
2. Closed form solutions exist but these solutions are complex and difficult to use for calculations.
3. Data for variables are available in the form of table, but no mathematical relationship between them is known, as is often the case with the experimental data.

A definite integral of the form: $\int_a^b f(x)dx$ can be treated as the area under curve $y = f(x)$, enclosed between the limits $x = a$ and $x = b$. Then the problem of integration is then simply reduced to the problem of finding the shaded area.





$$a = x_0, \qquad f(a) = y_0$$

$$b = x_1, \qquad f(b) = y_1;$$

$$h = (x_1 - x_0)$$

4.2.1 Trapezoidal Rule

Here we take, $I = \int_a^b f(x)dx \approx$ Area of the trapezian $ABCD$.

i.e

$$\boxed{I = \frac{1}{2}(b-a)[f(a) + f(b)]}$$

Another way to look at this is to see that we pass a 1st degree polynomial and take the area of the polynomial as the integral.

$P(x) = y_0 + s\Delta y_0 \leftarrow$ (1st degree Newton's forward difference polynomial)

where,

$$s = \left(\frac{x - x_0}{h}\right)$$

Then,

$$I = \int_a^b (y_0 + s\Delta y_0)dx$$

$$= \int_{x_0}^{x_1} (y_0 + s\Delta y_0)dx$$

$$= \int_0^1 (y_0 + s\Delta y_0)hds\,;$$

$$dx = hds; \quad when\, x = x_0,\ s = 0$$

$$x = x_1,\ \ s = 1$$

$$\therefore I = h\int_0^1 (y_0 + s\Delta y_0)ds$$

$$= h\left[y_0 s + \frac{s^2}{2}\Delta y_0\right]_{s=0}^{s=1}$$

$$= h\left[y_0 + \frac{1}{2}(y - y_0)\right];$$

Where, $\Delta y_i = y_{i+1} - y_i$

$$= \frac{h}{2}[y_0 + y_1]$$

$$\therefore I = \frac{(b-a)}{2}[f(a) + f(b)]$$

Example: Approximate the following integrals by using Trapezoidal rule.

i. $\int_0^{\pi/4} \sin x \, dx$

ii. $\int_0^{3} e^{-x^2/2} dx$

Error in Trapezoidal Rule

Here, we define e_T, the error in the simple trapezoidal rule to be the difference between the actual value of the integral and our approximation to it, i.e.

$$e_T = \int_a^b f(x)dx - \frac{1}{2}(b-a)(f(a) + f(b))$$

It is enough for our purposes here to obtain the anticipated error directly as:

$$e_T = -\frac{1}{12}(b-a)^3 f''(c)$$

where c is a number between a and b. The principal drawback with this expression is that we do not know what c is, but will find a way to work around that.

There are two factors which can influence e_T.

1) If $(b - a)$ is small then, clearly, e_T will also be small.

2) If f“ is small everywhere in $a < x < b$ then e_T will be small. That means, if f is a long way from being a straight line, then f“ will be large and hence so will the error, e_T.

Therefore, here we replace $f"(c)$ with a worst case value to obtain an “Upper bound” on e_T. This worst case value is the largest value that $f"(x)$ achieves for $a \le x \le b$. This leads to:

$$|e_T| \le \max_{a \le x \le b} |f"(x)| \frac{(b-a)^3}{12}$$

Example:

Estimate the error bound for the simple trapezoidal method approximations to,

a) $\int_0^{\pi/4} \sin x \, dx$

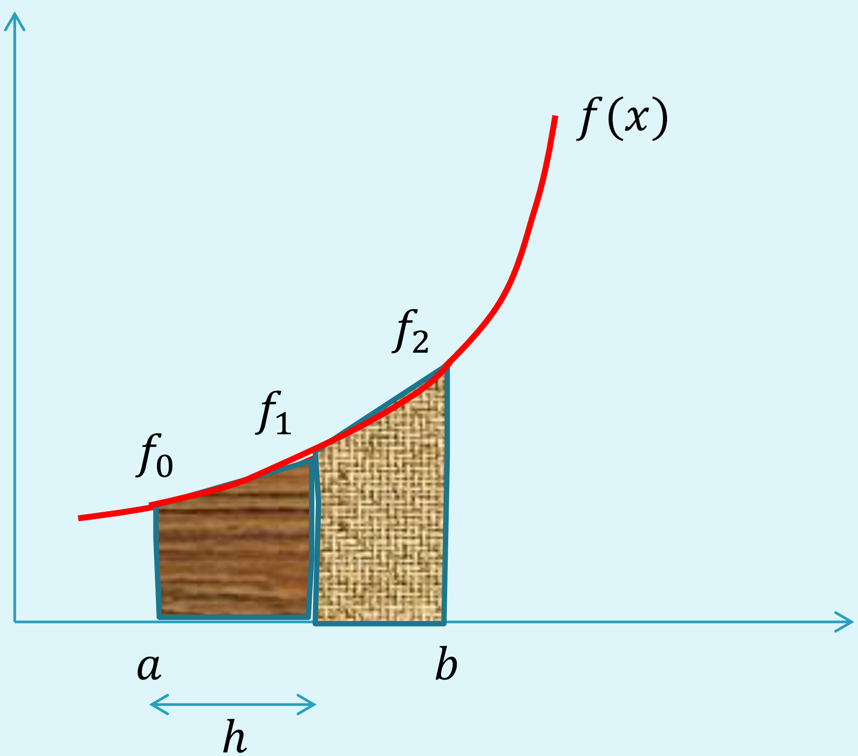
b) $\int_1^2 \ln x \, dx$

4.2.2 The Composite Trapezoidal Rule

The general idea is to split the interval [a,b] into a sequence of N small sub intervals of equal width $h = \frac{(b-a)}{N}$.

Then we apply the simple trapezoidal rule to each of the subintervals.

Lets consider the case of $N = 2$.

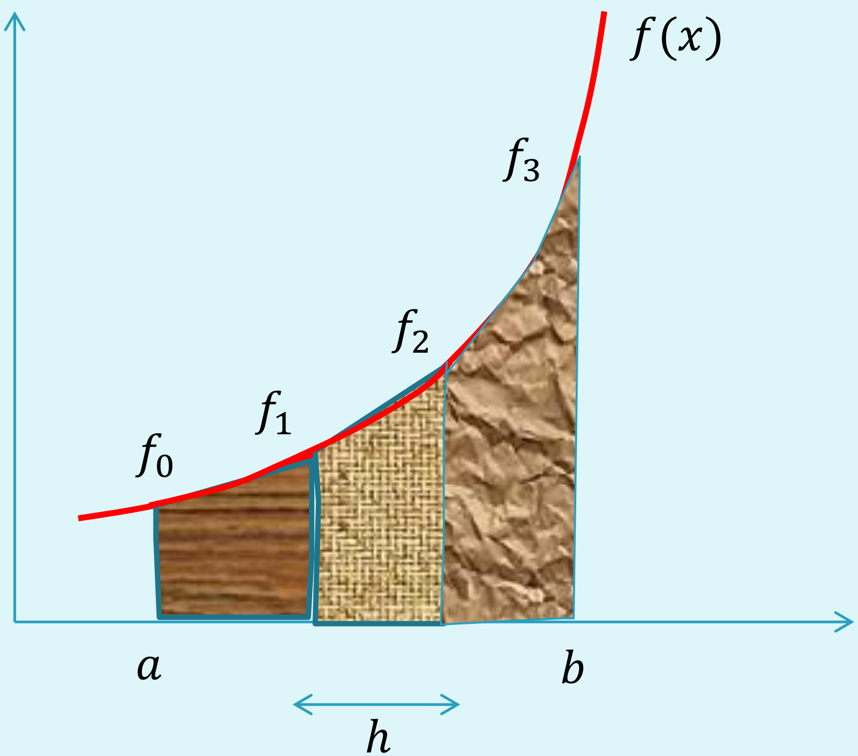


$$\therefore \int_a^b f(x)dx =$$
$$(Area\ of\ 1st\ trapezium) +$$
$$(Area\ of\ 2nd\ trapezium)$$

$$= \frac{1}{2}h(f_0 + f_1) + \frac{1}{2}h(f_1 + f_2)$$

$$= \frac{1}{2}h(f_0 + 2f_1 + f_2)$$

Lets consider the case of $N = 3$.



$$\therefore \int_a^b f(x)dx =$$

$$\frac{1}{2}h(f_0 + f_1) + \frac{1}{2}h(f_1 + f_2) +$$

$$\frac{h}{2}(f_2 + f_3)$$

$$= \frac{1}{2}h(f_0 + 2[f_1 + f_2] + f_3)$$

Therefore, we can write the composite trapezoidal rule for the case of N subintervals,

$$\boxed{\int_a^b f(x)dx \approx \frac{1}{2}h(f_0 + 2[f_1 + f_2 + \cdots + f_{N-1}] + f_N)}$$

Where,

$$h = \frac{b-a}{N},$$

$$f_0 = f(a), f_1 = f(a+h), \ldots, f_n = f(a+nh), \ldots,$$

$$f_N = f(a+Nh) = f(b)$$

Example: Approximate the area under the curve $y = \sqrt{x}$ on the interval $2 \leq x \leq 4$ using $n = 5$ subintervals. That is approximate the definite integral $\int_2^4 \sqrt{x}\, dx$ by the composite trapezoidal rule.

Error in composite Trapezoidal rule

We have to apply the error rule for each subinterval and let the total error is equal e_T^N, then we can write,

$$e_T^N \leq \max_{\substack{1st\ sub \\ interval}} |f''(x)| \frac{h^3}{12} + \max_{\substack{2nd\ sub \\ interval}} |f''(x)| \frac{h^3}{12} + \cdots + \max_{\substack{last\ sub \\ interval}} |f''(x)| \frac{h^3}{12}$$

$$= \frac{h^3}{12} \left(\boldsymbol{N} \max_{a \leq x \leq b} |f''(x)| \right)$$

Here, as the process of calculating $|f''|$ separately in each subinterval is time consuming, we replace it with the biggest error in the full interval.

Now, $Nh = (b - a)$

So,

$$\boxed{e_T^N \leq \left(\max_{a \leq x \leq b} |f''(x)|\right) \frac{(b-a)h^2}{12}}$$

Example:

The function f is known to have a second derivative with the property that

$$|f''(x)| \leq 14$$

for x between -1 and 4.

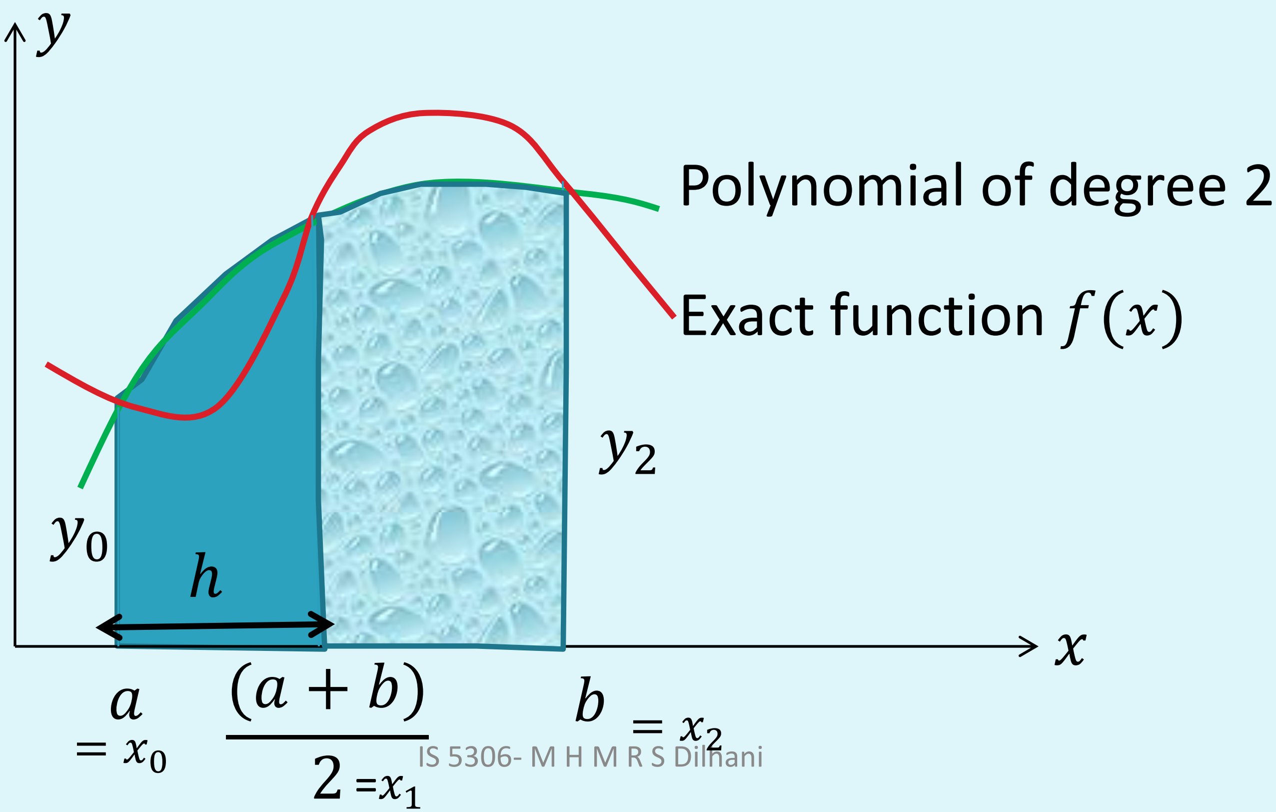
Determine, how many sub intervals are required so that the composite trapezoidal rule used to approximate

$$\int_{-1}^{4} f(x)dx$$

Can be guaranteed to have an error less than 0.00001.

4.2.3 Simpson's Rule

This method is based on passing a quadratic (polynomial of degree 2) through three equally spaced points, rather than passing a straight line through two points as we did for simple trapezoidal rule.



$$I = \int_a^b p(x)dx$$

$$I = \int_{x_0}^{x_2} (y_0 + s\Delta y_0 + \frac{s(s-1)}{2!}\Delta^2 y_0)dx$$

$$= \int_0^2 \left(y_0 + s\Delta y_0 + \frac{s(s-1)}{2}\Delta^2 y_0\right) hds$$

$$= h\left(y_0 s + \frac{s^2}{2}\Delta y_0 + \frac{(s^3/3 - s^2/2)}{2}\Delta^2 y_0\right)\Bigg|_{s=0}^{s=2}$$

$$= h\left[2y_0 + 2(y_1 - y_0) + \frac{1}{3}(y_0 - 2y_1 + y_2)\right]$$

$$= h\left[2y_0 + 2y_1 - 2y_0 + \frac{1}{3}(y_0 - 2y_1 + y_2)\right]$$

$$\boxed{I = \frac{h}{3}(y_0 + 4y_1 + y_2)}$$

Example 1:

Approximate $\int_0^{\pi/4} Sin\ x\ dx$ by using simpson's rule.

Example 2:

Find the approximations of $I_1 = \int_0^1 x^2 dx$ and $I_2 = \int_0^1 x^3 dx$ by using,

a. Trapezoidal rule
b. Simpson's rule
c. Directly

Error in Simpson's rule

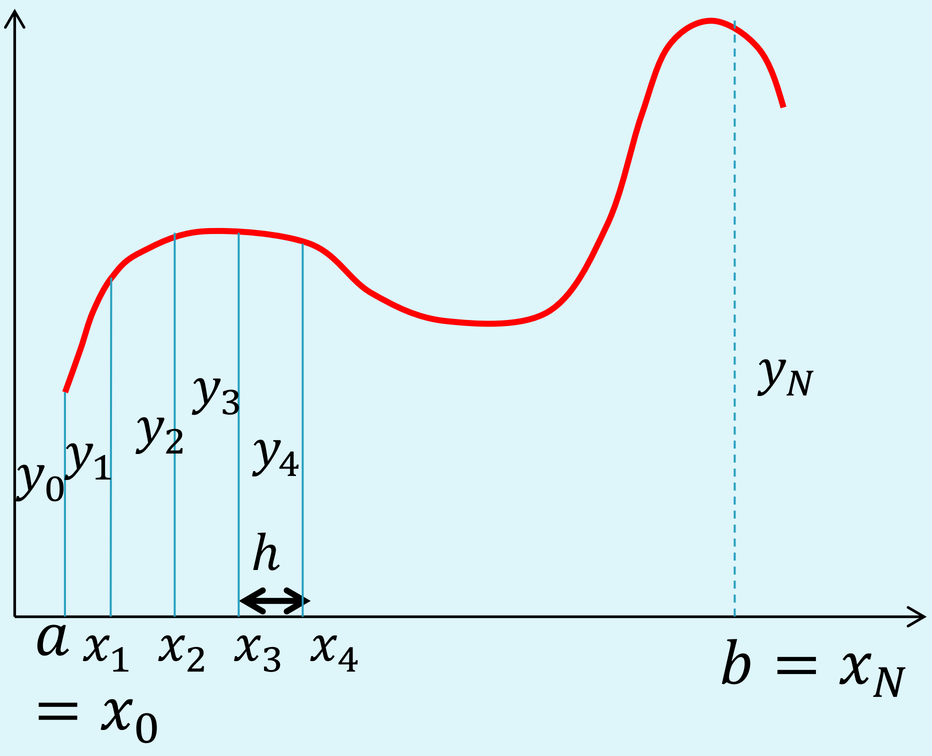
The error in Simpson's rule can be derived as

$$|e_s| \leq \max_{a\leq x\leq b} \frac{\left|f^{(4)}(x)\right|h^5}{90}$$

Example:

Estimate the error bound for the Simpson's rule approximations of $\int_0^{\pi/4} \sin x\, dx$

4.2.4 The composite Simpson's Rule



By dividing the interval $[a, b]$ into an even number of sub intervals of equal size, we can apply the simpson's rule for $[x_0, x_1, x_2], [x_2, x_3, x_4], [x_4, x_5, x_6]$ ... as so on.

Here, $h = x_{i+1} - x_i, \;\; h = \frac{(b-a)}{N}, where\ N$ is even interger

Then we get,

$$I = \int_a^b f(x)dx = \frac{h}{3}[y_0 + 4y_1 + y_2] + \frac{h}{3}[y_2 + 4y_3 + y_4] + \frac{h}{3}[y_4 + 4y_5 + y_6] + \cdots + \frac{h}{3}[y_{N-2} + 4y_{N-1} + y_N]$$

Then

$$\boxed{I = \int_a^b f(x)dx = \frac{h}{3}[y_0 + 4(y_1 + y_3 + y_5 + \cdots + y_{N-1}) + 2(y_2 + y_4 + y_6 + \cdots + y_{N-2}) + y_N]}$$

Example:

Using 6 subintervals in the composite Simpson's rule, approximate,

$$\int_{1}^{4} \ln x \, dx$$

Error in composite Simpson's Rule

The error in the N-subinterval composite Simpson's approximation to $\int_a^b f(x)dx$ is bounded above by,

$$|e_s^N| \leq \max_{a \leq x \leq b} \frac{|f^4(x)|(b-a)h^4}{180}$$

Example:

The function f is known to have a fourth derivative with the property that, $|f^4(x)| < 6$ for x between -1 and 5. Determine how many subintervals are required so that the composite Simpson's rule used to approximate,

$$\int_{-1}^{5} f(x)\,dx$$

incurs an error less than 0.001.

4.2.5 Gaussian Quadrature Method

We consider the numerical evaluation of the integral

$$I = \int_a^b f(x)dx$$

In the previous sections, we derived some integration formulae which require values of the function at equally-spaced points of the interval. Gauss derived a formula which uses the same number of function values but with different spacing and gives better accuracy.

Gauss' formula is expressed in the form

$$\int_{-1}^{1} F(t)dt = c_1F(t_1) + c_2F(t_2) + \cdots + c_nF(t_n) = \sum_{i=1}^{n} c_iF(t_i) \rightarrow (1)$$

where the c_i and t_i are called the weights and abscissae, respectively. An advantage of the formula is that the abscissae and weights are symmetrical with respect to the middle point of the interval.

In equation (1), there are altogether 2n arbitrary parameters and therefore the weights and abscissae can be determined such that the formula is exact when F(t) is a polynomial of degree not exceeding (2n-1).

Hence, we start with

$$F(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \cdots + a_{2n-1} t^{2n-1}$$

We then obtain from (1)

$$\int_{-1}^{1} F(t)dt = \int_{-1}^{1} (a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \cdots + a_{2n-1} t^{2n-1})\, dt$$

$$= 2a_0 + a_1 \times 0 + \frac{2}{3} a_2 + a_3 \times 0 + \frac{2}{5} a_4 + \cdots \quad \rightarrow (2)$$

Substituting these values on the right hand side of (1), we obtain

$$\int_{-1}^{1} F(t)dt = c_1(a_0 + a_1 t_1 + a_2 t_1^2 + \cdots + a_{2n-1} t_1^{2n-1})$$
$$+c_2(a_0 + a_1 t_2 + a_2 t_2^2 + \cdots + a_{2n-1} t_2^{2n-1})$$
$$+c_3(a_0 + a_1 t_3 + a_2 t_3^2 + \cdots + a_{2n-1} t_3^{2n-1}) + \cdots$$
$$+c_n(a_0 + a_1 t_n + a_2 t_n^2 + \cdots + a_{2n-1} t_n^{2n-1})$$

Which can be written as

$$\int_{-1}^{1} F(t)dt = a_0(c_1 + c_2 + \cdots + c_n)$$
$$+a_1(c_1t_1 + c_2t_2 + \cdots + c_nt_n)$$
$$+a_2(c_1t_1^2 + c_2t_2^2 + \cdots + c_nt_n^2) + \cdots$$
$$+a_{2n-1}(c_1t_1^{2n-1} + c_2t_2^{2n-1} + \cdots + c_nt_n^{2n-1} \rightarrow (3)$$

Now equation (2) and (3) are identical for all values of a_i and hence comparing the coefficient of a_i we obtain the 2n equations.

$$c_1 + c_2 + \cdots + c_n = 2$$
$$c_1 t_1 + c_2 t_2 + \cdots + c_n t_n = 0$$
$$c_1 t_1^2 + c_2 t_2^2 + \cdots + c_n t_n^2 = 2/3$$
$$\vdots$$
$$c_1 t_1^{2n-1} + c_2 t_2^{2n-1} + \cdots + c_n t_n^{2n-1} = 0$$

in 2n unknowns c_i and t_i ($i=1, 2, \ldots, n$).

As and illustration, we consider the case n=2. Then the formula is

$$\int_{-1}^{1} F(t)dt = c_1 F(t_1) + c_2 F(t_2)$$

$$\int_{-1}^{1} F(t)dt = \sum_{i=1}^{2} c_i F(t_i)$$

Since this formula is exact when $F(t)$ is a polynomial of degree not exceeding 3, we put successively $F(t) = 1, t, t^2$ and t^3. Then,

$$c_1 + c_2 = 2$$
$$c_1 t_1 + c_2 t_2 = 0$$
$$c_1 t_1^2 + c_2 t_2^2 = 2/3$$
$$c_1 t_1^3 + c_2 t_2^3 = 0$$

The solution of these equations is

$$c_1 = c_2 = 1$$
$$t_2 = -t_1 = \sqrt{\frac{1}{3}} = 0.5773$$

So,

$$\int_{-1}^{1} f(t)dt = f(-0.5773) + f(0.5773)$$

It is remarkable that adding these two values of the function gives the exact value for the integral of any cubic polynomial over the interval from -1 to 1.

Example 1:

Calculate the $\int_{-1}^{1}(2x^3 + 5)dx$, by using

i. Gaussian quadrature with n=2

ii. Directly.

Now suppose our limits of integration are from $a\ to\ b$ and not -1 to 1 for which we derived the formula. To use the tabulated Gaussian quadrature parameters, we must change the interval of integration to (-1,1) by a change of variable. We replace the given variable by another to which it is linearly related according to the following scheme:

If we let,

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

So that $dx = \frac{b-a}{2} dt$

Then,

$$\int_a^b f(x)dx$$

$$= \left(\frac{b-a}{2}\right) \int_{-1}^{1} f\left(\frac{(b-a)t+(b+a)}{2}\right) dt$$

Example 2:

Evaluate, $I = \int_0^{\pi/2} \sin x \, dx$, by using two-term Gaussian formula.

Higher point Gaussian Quadrature formulas.

$$\int_{-1}^{1} f(t)dt = c_1 f(t_1) + c_2 f(t_2) + c_3 f(t_3)$$

is called the three-point Gauss quadrature rule. The coefficients (weighting factors) $c_1, c_2 and\ c_3$ and functional arguments $t_1, t_2\ and\ t_3$ are calculated by assuming the formula gives exact expressions for integrating a 5th order polynomial,

$$\int_{-1}^{1}(a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5)dt$$

Also, the general n-point Gaussian quadrature rules would approximate the integral,

$$\int_{-1}^{1} f(t)dt = c_1 f(t_1) + c_2 f(t_2) + \cdots + c_n f(t_n)$$

$$\int_{-1}^{1} f(t)dt = \sum_{i=1}^{n} c_i f(t_i)$$

The coefficients and arguments are given in handbooks, for n-point Gauss-quadrature rule as follows.

Points	Weighting Factors	Function Arguments
2	$c_1 = 1.0000$	$t_1 = -0.5773$
	$c_2 = 1.0000$	$t_2 = 0.5773$
3	$c_1 = 0.5555$	$t_1 = -0.7746$
	$c_2 = 0.8888$	$t_2 = 0.0000$
	$c_3 = 0.5555$	$t_3 = 0.7746$
4	$c_1 = 0.3478$	$t_1 = -0.8611$
	$c_2 = 0.6521$	$t_2 = -0.3399$
	$c_3 = 0.6521$	$t_3 = 0.3399$
	$c_4 = 0.3478$	$t_4 = 0.8611$

Example 1: Evaluate $\mathrm{I} = \int_{0.2}^{1.5} e^{-x^2}\, dx$ using the 3 – term Gaussian formula.

Example 2: Human vision has the remarkable ability to infer 3D shapes from 2D images. In order to replicate some of these abilities on a computer, it is required to integrate the following vector field.

$$I = \int_0^{100} f(x)dx$$

Where,

$$f(x) = \begin{cases} 0 & 0 < x < 30 \\ -9.1688 * 10^{-6}x^3 + 2.7961 * 10^{-3}x^2 - 2.8487 * 10^{-1}x + 9.6778 & 30 \leq x \leq 172 \\ 0 & 172 < x < 200 \end{cases}$$

Use two point Gaussian quadrature rule to find the value of the integral.