<u>Numerical methods</u> =>Numerical methods are extremely powerful problem solving tools. they are capable of handling large system of equations, nonlinear entries and complicated geometries.

Numerical Methods is use to reduce higher mathematics to basic arithmetic operation

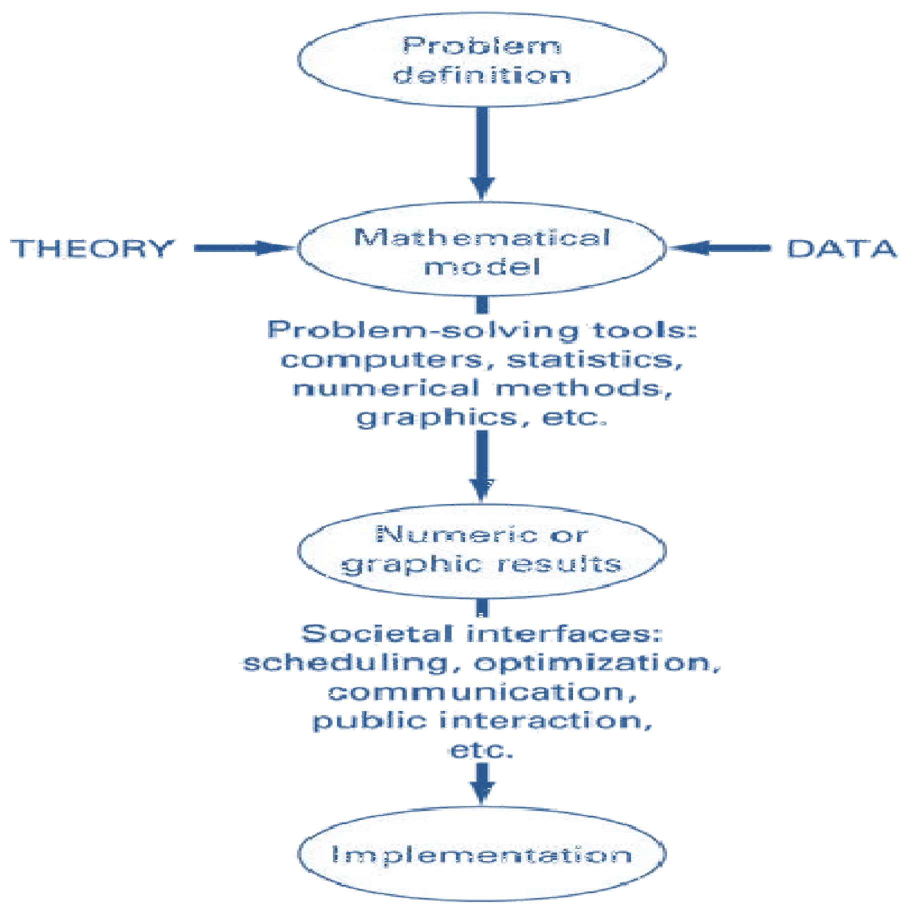
Numerical Methods are use to solve such complicated calculations.

Example: Integration

Analytical solution: Computer as usual as in calculus

Numerical method: Use trapezoidal rule or Simpson's rules

Problem Solving Process



Simple Mathematical model

Equations that expresses the essential features of a physical systems

Represented as a functional relationship in the form of

Dependent Variables = f (independentvariables, parameters, forcing function,)

Where,

Dependent Variables - Reflects the behavior or state of the system

Independent Variables - Dimensions, such as time and space

Parameters - Reflective of the system's properties or composition

Forcing Function - External influence acting upon it

Newton's 2nd law of Motion

States that "*the time rate change of momentum of a body is equal to the resulting force acting on it.*"

•The model is formulated as

F = ma (eqn 1)

F=net force acting on the body (N)

m=mass of the object (kg)

a=its acceleration (m/s2) Equation 1 can be written as:

•a = F / m (eqn 2)

•simple algebraic equation that can be solved analytically

To determine the terminal velocity of a free-falling body near the earth's surface using Newton 2nd law.

•Express acceleration as the time rate of change of the velocity (dv/dt) and substituting into eq. (2) we get

dv/dt = F/m (eqn. 3)

F = m *(dv/dt)*

F '+ve' : accelerate

F '-ve' : decelerate

F = 0 (constant velocity)

Express the net force in term of measurable variables and parameters, in which the net force is composed of 2 opposing forces:

The downward pull of gravity FD and the upward force

of air resistance Fu:

F = FD + Fu (eqn. 4)

If downward force is '+ve', 2nd law can be used to formulate the force due to gravity, as

FD = mg (eqn. 5) (where g = 9.8 m/s2)

The air resistance that acts in an upward direction;

Fu = - cv (eqn. 6)

c = drag coefficient (kg/s) Fu

The net force is the difference between the downward (FD) and upward (FU).

By combining eqn 5 and 6

F= mg- cv (put the value of F in eqn 3)

dv/dt = (mg – cv)/m (eqn. 7)

dv/dt = g – (c/m)v (eqn. 8)

Eq. (7) is a differential equation. The exact solution of eq. (8) cannot obtained by simple algebraic manipulation, which needs calculus to obtain an exact or analytical solution.

If v = 0 at t=0, calculus can be used to solve eq. (8) for

$$\mathbf{v(t) = (gm/c)\,[1-e^{-(c/m)t}]} \quad \text{(eqn 9)}$$

<u>Numerical computing=></u>Numerical computing is an approach for solving complex Mathematical problems using only Simple Arithmetic operations.This approach involves formulation of mathematical models of the physical situations that can be solved with arithmetic operations. It also requires development analysis and use of Algorithms.

Great Mathematician like Gauss ,Newton, Lagrange and many others developed numerical techniques which are widely used. The digital computers has enhanced the speed and accuracy of Numerical computing techniques.

<u>Process of Numerical Computing=</u>>The Process of Numerical Computation can be divided in to following phases:-

1. Formulation of mathematical model.
2. Construction of an appropriate numerical method.

3. Implementation of the method to obtain a solution.
4. Validation of the solution.

In first phase we have various types of mathematical model equations that can be used for formulating mathematical model of the physical processes before selecting a particular method.

In the second phase we select appropriate method to solve given problem. This selection may base on the following parameters. Such as type of equation, type of computer available, accuracy required, speed of execution and programming and maintenance efforts required . The modeling is process of translating a physical problem into mathematical problem.

The third phase of the numerical computing process is the implementation process of the method selected, which is based on the following three tasks

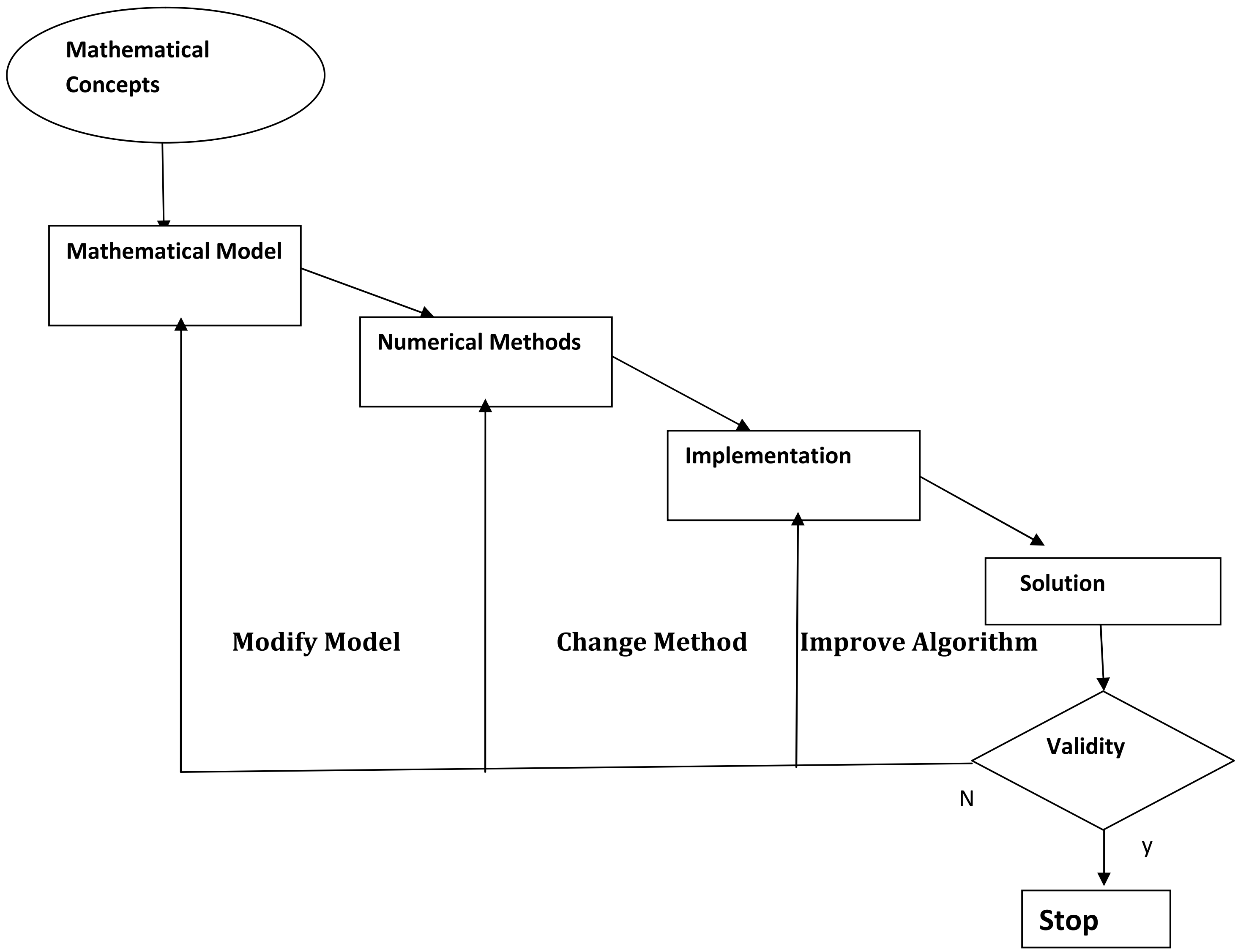
a. Design of the algorithm
b. Writing an algorithm
c. Executing it on a computer to obtain the result.

Once we are able to obtain results , the next step is the validation of the process. Validation means the desired limit and accuracy. If it is not valid then we must go back to check

a. Mathematical model
b. Numerical methods selected
c. Computational algorithm used implement the method

Once the modification is introduced the cycle begins again.

This whole process is given in the figure



Characteristics of Numerical Computing

Numerical Methods show certain computational characteristics during their implementation. It is important to consider these characteristics while choosing a particular method for implementation.

The characteristics are-:

1. **Accuracy** – Accuracy describes how close an approximation to an actual value. Every method of numerical computing introduces errors. They may be either due to using an approximation in place of an exact mathematical procedure or due to in exact representation and manipulation of numbers in computer
2. **Rate of Convergence** – Many Numerical methods based on the idea of an interactive process. This process a sequence of approximations with the hope that the process will converge to required solution.
   There are several techniques for acceleration the rate of convergence of certain methods.

3. **Numerical Stability** – Another problem introduced by some numerical computing methods is that of numerical instability. In some cases, the errors tends to grow exponentially, with disastrous computational results. A computing process that exhibits such exponential error growth is said to be numerically unstable. We must choose method that are not fast but also stable.

4. **Efficiency -**One more consideration in choosing a numerical method for solution of mathematical model is efficiency. The method that requires less of computing time and less of programming effort and yet achieves the desired accuracy is always preferred.

Numerical Data:- Numerical computing involves two types of data, namely discrete data and Continuous data.

i) Discrete Data:- Data that are obtaining by counting are called discrete data.

For example:- the total number of balls in box, total number of people participating in event and digital clock.

ii) Continuous Data:- Continuous data are obtained through measurement are called continuous data.

For example:- Continuous data are the speed of a vehicle as given by speedometer or temperature of patient is measured by thermometer.

Anolog computing

Analog Computing is based on inputs that vary continuously, such as current, voltage or temperature. The earliest computers were analog and functioned on the basis of electrical voltages. Calculations were performed by adding, subtracting, multiplying and dividing voltages, Analog computers are fast,but their accuracy is limited by precision with which the physical quantities can be read.

Example of application of analog computers is a machine used in a parcel office to convert the weight of a parcel in to the cost of parcel needed for mailing.

The basic requirement of analog computers is the writing down of differential equations describing the physical system of interest for a particular application.

Digital Computing:-Digital computer is a computing device that operates on inputs which are discrete in nature. The input data are numbers that may represent numerical, letters or other special symbols . just as a digital clock directly counts the seconds and minutes in an hour, a digital computers counts discrete data values to compute results.

Digital computers are more accurate than analog computers. Analog computers may be accurate to within 0.1% of correct value.

Digital computers are widely used for many different applications and are often called general purpose computers

<u>Significant digits</u> :- There are two kinds of numbers, exact and approximate numbers. Examples of exact numbers are 1,2,3…….1/2,3/2.

Approximate numbers are those that represents the numbers to a certain degree of accuracy. Thus the approximate value of π (PI) is 3.1416 or if we desire a better approximation, it is 3.14159265 .but we cannot write the exact value ofπ (PI).

The digit that are used to express to a number are called significant digit or significant figures. Thus the numbers 3.146, 0.66667 and 4.0687 have 5 significant digits.

The concept of significant digits has been introduced primarily to indicate the accuracy of a numerical value.

Notion of significant digits

1)All Non-zero numbers are significant.

2) All zeros occurring between non-zero digits are significant digits.

Example:- 3.50 , 65.0 and 0.230 have 3 significant digits each.

3) Zeros between the decimal point and preceding a non-zero digit are nonsignificant.

Example:- 0.0001234 have 4 significant digit.

4) when a decimal point is not written, trailing zero are not considered to be significant digit.

Example: 1) 4500 may be written as have 2 significant digit. And 500.0 have 4 significant digit.

5)Scientific notation, all digits in the coefficient are significant. The exponent is not considered in terms of significant figure.

Example 1**:** 2.100×10^3 will have with certainty 4 significant figures.
Example 2**:** 3.50×10^3 will have with certainty 3 significant figures.
Example 3**:** 3.54×10^5 has three significant digits.

Example 4**:** 1.6340×10^{-6} has five significant digits.

The concept of accuracy and precision are closely related to significant digits.

1) Accuracy refers to the number of significant digits in a value.
   For ex=> The number 57.396 is accurate to five significant digit.
2) Precision refers to the number of decimal positions that is the order of magnitude of last digit in a value.
   For ex=> The number 57.396 has a Precision of 0.001 or 10^{-3}

Problems=> Find the accuracy and precision of following number
1)12.345

The accuracy of 12.345 is 5 significant digit

The Precision of 12.345 is 10^{-3}

2)750.5

The accuracy of 750.5 is 4 significant digit

The Precision of 750.5 is 10^{-1}

Errors=> Errors come in a variety of forms and sizes some are avoidable and some are not. Errors arises, from the use of approximations to represent exact mathematical operations and quantities. The approximation can be formulated as,

True value = approximation+ error

E_t = true value- Approximation

(where E_t is used to exact value of the error)

Approximation Error :- The Approximation error in a some data is the difference between an exact value and some approximation to it. An Approximation error can occur because:-

1) The measurement of data is not exact due to instrument.
   Ex: the accurate reading of paper is 4.5 cm but since the ruler does not use decimal you round it 5 cm
2) Approximation are used instead of real data(PI=3.14 instead of PI=3.1415…..)

Inherent Error:- Inherent Errors are those that are present in the model(also known as Input Error) Inherent error contain two components, Data errors and Conversion errors.

Data errors arises when the data for a problem are obtained by some experimental means and of limited accuracy and Precision.

Conversion Error arises due to the limitations of the computer to store the data exactly. We know that the floating point representation retain only a specified number of digits.

In Computations , inherent errors can be better data, by correcting obvious errors in the data and by using computing aids of higher precision.

<u>Numerical Errors</u> => Numerical errors are introduced during the process of implementation of a Numerical Method.

They come in two forms, Round off error and Truncation error

The total numerical error is the summation of these two errors.

This error can be reduced by devising suitable techniques for implementing the solution.

1**)<u>Round off error =></u>** Round off error occurs because of the computing devices inability to deal with certain numbers. Such numbers need to be round off to some near approximation, which is depend on the word size used to represent numbers of the device.

Number such as π , e or $\sqrt{7}$ cannot be expressed by a fixed number of significant figures. There for they cannot be represented exactly by the computer.

Computer use base-2 representation, they cannot represent exact base-10 numbers. The discrepancy introduced by this omission of significant figures is called round-off error.

a)**<u>Chopping:</u>**=> In chopping, the extra digit are dropped. this is called truncating the Number.

Suppose we are using a computer with a fixed word length of four digits. Then a number like 42.7823 will be stored as 42.78 and the digit 23 will be dropped.

The general form of True value x is,

$$\text{True } x = (f_x + g_x \text{ X } 10^{-d})\ 10^{E}$$

$$x = f_x \text{X } 10^{E} + g_x \text{ X } 10^{E-d}$$

True x = approximate + error

Where ,f_x is the mantissa , d is length of the mantissa permitted and E is the exponent. In chopping, g_x is ignored entirely and therefore,

$$\text{Error} = g_x \text{ X } 10^{E-d} \quad , \quad 0 \le g_x < 1$$

Ex 1) Round off the Number 42.7823 up to four significant digit

$$x = 0.427823 \text{ X } 10^{2}$$

$$x = (0.4278 + 0.000023) \text{ X } 10^{2}$$

$$x = (0.4278 + 0.23 \text{X} 10^{-4}) \text{ X } 10^{2}$$

$$x = (0.4278 \text{ X } 10^{2} + 0.23 \text{X} 10^{2-4})$$

by formula here, $x = (f_x \text{ X } 10^{E}) + (g_x \text{ X } 10^{E-d})$

we know that error= $g_x \text{ X } 10^{E-d}$

in our problem $g_x \text{ X } 10^{E-d} = 0.23 \text{X} 10^{2-4}$

so we drop the error

True $x = 0.4278 \text{ X } 10^{2}$

x=42.78

b) <u>Symmetric Round-off=></u>In the symmetric round-off digit is "rounded up" by 1,if the first discarded digit is larger or equal to 5.

For ex:- 42.7893 would become 42.79

True $x = f_x X\ 10^E + g_x\ X\ 10^{E-d}$

Where $g_x<0.5$, entire g_x is truncated

Approximate $x= (f_x\ X\ 10^E)$

But, when $g_x \geq 0.5$ then last digit in the mantissa is increased by 1

<u>Approximate $x= (f_x\ X\ 10^E) + 10^{E-d}$</u>

<u>Error = $(g_x-1)\ X\ 10^{E-d}$</u>

Where $g_x \geq 0.5$

Absolute error $\leq 0.5\ X\ 10^{E-d}$

Example:- Round off the Number 1.3526 up to four significant digit.

$x= 1.3526$

$x= 0.13526\ X\ 10^1$

$x= (0.1352+0.00006)\ X\ 10^1$

$x= (0.1352 + 0.6\ X\ 10^{-4})\ X\ 10^1$

$x= (0.1352\ X\ 10^1) + (0.6\ X\ 10^{1-4})$

$x = (0.1352\ X\ 10^1) + (0.6\ X\ 10^{-3})$

$g_x= 0.6 \geq 0.5$ (error is greater than 0.5) so symmetric round-off is as follows

Approximate $x= 0.1352\ X\ 10^1 + 1\ X\ 10^{-3}$

x= 1.352 + .001

x=1.353

Error value = (0.6 -1) X 10^{-3}

= (-0.4) X10^{-3}

= -0.0004

Truncation Error:= Truncation Error arises from using an approximation in place of an exact mathematical procedure. we use some finite number of terms to estimate the sum of an infinite series.

For ex:- S= $\sum_{i=0}^{\infty} a_i x^i$ is replaced by the finite sum $\sum_{i=0}^{n} a_i x^i$ the series has been truncated.

Consider the following infinite series,

Sin x = $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$

When we calculate the sine of an angle using the series, we can not use all the terms in the series, for computation. We usually terminate the process , after a certain term is calculated. The terms truncated introduce an error which is called truncation error.

Truncation error can be reduced by increasing the number of points at which function is integrated. We often use library trigonometric function, hyperbolic function and so on. In all these cases, a series is used to evaluated these functions. It is important to know the truncation errors introduced by these functions. It is important to know the truncation errors introduced by these library function.

Prob:- 1)Estimate the maximum truncation error in evaluating the expression

$$x^3 – 25\ x^2 + 3.1\ x – 1.5 \quad \text{at } x=1.25$$

solution:- in the given problem we have four terms in polynomial we truncate one by one term and calculated truncation error in each part

case 1: if we truncate only one term than truncation error as -1.5

case 2: if we truncate only two term from the expression then we have truncation error as

$$3.1\ x -1.5 \quad \text{(put } x=1.25)$$

$$3.1(1.25) – 1.5$$

$$2.375$$

Truncation error in this case as 2.375

Case 3: if we truncate only three term from the expression then we have truncation error as

$$- 2.5\ x^2 + 3.1\ x -1.5 \quad \text{(put } x=1.25)$$

$$-2.5\ (1.25)^2 + 3.1\ (1.25) – 1.5$$

$$-2.5\ (1.5625) +3.875 – 1.5$$

$$-3.90625+2.375$$

$$1.53125$$

Case 4 :- if we truncate all term from the expression then we have truncation error as

$x^3 - 25\,x^2 + 3.1\,x - 1.5$ at x=1.25

$(1.25)^3 - 25(1.25)^2 + 3.1\,(1.25) - 1.5$

1.953125 - (1.5312)

0.421925

Hence maximum truncation error in the second case as 2.375, so we truncate two term from polynomial

Prob:-2) Evaluate sin 1 by using truncates series

$x = \ x - \frac{x^3}{3!} + \frac{x^5}{5!}$for sin x

Solution:- $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!}$

$\text{Sin } 1 = 1 - \frac{1^3}{3!} + \frac{1^5}{5!}$

$= 1 - \frac{1}{6} + \frac{1}{120}$

Sin 1= 0.84166

Actual value of sin 1= 0.84147

Error value= 0.84147- 0.84166

= - 0.00019

Blunders :- Blunders are errors that are caused due to human imperfection. As the name indicates, such errors may cause a very serious disaster in the result. Since there errors are due to human mistakes, it should be possible to avoid them to a large extent by acquiring a sound knowledge of the problem as well as the numerical process.

Human error may occurs at any stage of numerical processing, some common types of error are,

1) Lack of understanding of problems.
2) Wrong assumptions.
3) Over look of some basic assumptions required for formulating the model.
4) Selecting a wrong numerical method for solving mathematical model.
5) Selecting a wrong algorithm for implementing the numerical model.

Modeling Error:=> Mathematical Models are the basis for numerical solutions. They formulated to represent physical process using certain parameters involved in situations.

In many situations, it is impossible to include all of the real problem and therefore, certain simplifying assumptions are made.

For ex:- while developing a model for calculating the force acting on a falling body, we not b able to estimate the air resistance coefficient or drag coefficient properly or determine the direction and magnitude of wind force, acting on the body, to simplify the model, we may assume that the force due to air resistance is linearly proportional to the velocity of the falling body or we may assume that there is no wind force acting on the body. All such simplifications certainly reset in errors in the output from such models.

Absolute and Relative error:- Let us suppose that the ture value of a data item is denoted by x_t and its approximate value is denoted by x_a . then, they are related as follows:

True Value x_t= Approximate value (x_a) + Error

$x_t = x_a + \text{Error}$

$$\text{Error} = x_t - x_a$$

The error may be negative or positive depending on the values of x_t and x_a . In error analysis, magnitude of errors is important not the sign. Therefore ,we normally consider ,what is known as absolute error which is denoted by,

$e_a = | x_t - x_a |$

Relative error which is nothing but the “ normalized” absolute error. The Relative error defined as follows:

$e_r = \frac{absolute\ Error}{|\ true\ value|}$

$= | x_t - x_a | / | x_t|$

$e_r = |1 - x_a/x_t|$

more often, the quantity that is known to us is x_a and there for, we can modify the above relation as follows:

$e_r = x_t - x_a / x_a$

$= |1 - x_t/x_a|$

The functional form of e_r can also be expressed as the percent relative error as ,

Persent $e_r = e_r$ X 100

Prob1) A civil engineer has measured the height of a floor building as 2950 cm working height of each beam as 35 cm while the true value are 2945 cm and 30 cm ,respectively compare their absolute and relative errors.

Sol^n :- Absolute error is measuring in the height of building is,

$e_{a1} = 2950 – 2945= 5$ cm

The relative error is,

$e_{r1=}$ 5/ 2945 = 0.0017 = 0.17%

Absolute error is measuring in the height of beam is,

$e_{a2} = 35 – 30 = 5$ cm

The relative error is,

$e_{r1=}$ 5/ 30 = 0.17 = 17%

Although the absolute errors are the same and the relative error are differ by 100 times, it shows that there is something wrong in measurement of the height of the beam.

Prob 2) Suppose that you have the task of measuring the length of a bridge and a bolt come up with 9999 and 9 cm respectively. If the true values are 10,000 and 10 cm respectively compute

a) The true value and
b) The true percent relative error for each case.

Sol^n :- The error for measuring the bridge is,

$e_t = 10{,}000 - 9999 = 1$ cm

and for the bolt is,

$e_t = 10 - 9 = 1$ cm

The percent relative error for the bridge is,
$e_r = 1/1000 \ X\ 100\ \%= 0.01\%$

The percent relative error for the bolt is,
$e_r = 1/10 \ X\ 100\ \% = 10\ \%$

Error Propagation :- Every numerical computation involves a series of computations. It consists of basic arithmetic operations. Therefore individual round off is not much important as the final error on the result. Actually an error at once point in the process propagates and it affects final total error.

Consider addition of two numbers, say x and y and x_t and y_t represents exact value

$x_t + y_t = x_a + e_x + y_a + e_y$

$= (x_a+y_a) + (e_x+e_y)$

there fore,

Total error = $e_{x+y} = e_x + e_y$

Similarly for subtracting,

$e_{x-y} = e_x - e_y$

Note that e_{x+y} does not means that error will increase in all cases. It depends on the sign of individual errors

Since we do not normally know sign of errors, we can only estimate error bounds. That is

Total Error $e_{x\pm y} \leq | e_x| + |e_y|$

" The magnitude of the absolute error of a addition(or subtraction) is equal to or less than the sum of magnitude of the absolute error of the operands"