CHAPTER 1

Errors in Numerical Calculations

1.1 INTRODUCTION

Numerical methods are very powerful and popular tools for solving a variety of engineering, mathematical and scientific problems using the four basic arithmetical operations. In this chapter we introduce Numerical techniques are used to solve problems involving higher order polynomials. They are used in solving transcendental equations. The numerical methods are also used in solving equations involving several variables. The techniques employed in numerical analysis are times approximate. Therefore the results (i.e., outcomes) obtained by numerical methods have some errors.

Let X_E denote an exact number and α be a number that differs slightly from X and is used in place of X in calculations, then α is called **an approximate number**.

If α is less than X then it is called a minor approximation of X , and if α is greater than X, then it is called a major approximation of X

Definition: Let x be an exact number and α be the approximate number of x, then the difference between x and α is called the **error of $\boldsymbol{\alpha}$**.

It is denoted by E and is given by

$$E = X - \alpha \qquad \text{.....(1.1)}$$

If X > α , then the error is positive, and if $X < \alpha$, then the error is negative.

1.2 ABSOLUTE ERROR

The absolute error E_A of an approximate number α is the absolute value of the difference between the corresponding exact number x and the numberα.

$$\therefore E_A = |X_E - X_A|$$

It is also denoted by ΔX

1.3 LIMITING ABSOLUTE ERROR

Definition: The limiting absolute error of an approximate number is any number that is not less than the absolute error of that number.

Thus if Δ_α is the limiting absolute error of an approximate number α which takes the place of the exact numberX_E, then

$$|X_E - \alpha| \leq \Delta_\alpha$$

The exact number X lies within the range

$$\alpha - \Delta_\alpha \leq X_E \leq \alpha + \Delta_\alpha$$

We can write X $= \alpha \pm \Delta_\alpha$

Note: The absolute error does not suffice to describe the accuracy of a measurement or a computation. An essential point in the accuracy of the measurements is the absolute error related to unit length. It is called the relative error.

1.4 RELATIVE ERROR

The relative error E_R of an approximate number α is the ratio of the absolute error E_A of the number to the modulus of the corresponding exact number x.

From the definition we have

Relative error $E_R = \frac{E_A}{X}$

We can also write $E_R = \frac{E_A}{X} = \delta$X

1.5 THE LIMITING RELATIVE ERROR

The limiting relative error of a given approximate number α, is any number not less than the relative error of that number. It is denoted by δ_α .

By definition we have $\delta \leq \delta_\alpha$

That is $\frac{\Delta}{|X|} \leq \delta_\alpha$, when $\Delta \leq$ |x| δ_α

Thus for limiting relative absolute error of a number α we can take Δ_α= x| δ_α|, from which, knowing the relative error δ_α we can obtain for the exact number. Since the exact number lies between α (1$-\delta_\alpha$) and α (1+δ_α)

We can write X =$(1 \pm \delta_\alpha)$

If α is an approximate number taking the place of an exact number X, and Δ_α is the limiting absolute error of α taking

$$X > 0, \alpha > 0 \text{ and } \Delta_\alpha < \alpha, \text{ we get}$$

$$\delta = \frac{\Delta}{X} \leq \frac{\Delta\alpha}{\alpha - \Delta_\alpha} \Rightarrow \delta_\alpha = \frac{\Delta\alpha}{\alpha - \Delta_\alpha}$$

For the limiting relative error of the number α.

Similarly we can show that $\Delta_\alpha = \frac{\delta_\alpha}{\alpha - \delta_\alpha}$

Note: If Δ_α is very much less than α, and δ_α is very much less than 1, we can take

$$\delta_\alpha \approx \frac{\Delta\alpha}{\alpha} \text{ and } \Delta_\alpha \approx \alpha\delta_\alpha$$

1.6 PERCENTAGE ERROR

The percentage error E_P is defined by

$$E_P = E_R \times 100$$

1.7 SOURCES OF ERRORS

The errors in mathematical solution of problems are of five types.

1. Errors involved in the statement of the problems
2. Errors stemming from the presence of infinite processes in mathematical analysis
3. Errors due to numerical parameters whose value can only be determined approximately
4. Errors associated with the system of numeration
5. Errors due to operations involving approximate numbers

In this section we discuss two types of errors namely truncation error and computational errors.

The errors which are inherent in the numerical methods employed for finding numerical solutions are known as **truncation errors**

The truncation error arises due to the replacement of an infinite process such as summation or integration by a finite one. These errors are caused by using approximate formulae in computation

Trigonometric functions are computed of by summing series.

$S = \sum_{r=0}^{\infty} a_r x^r$ is replaced by the finite sum $\sum_{r=0}^{n} a_r x^r$

For example consider $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$

Is summed t n terms. Suppose we wish to calculate $e^{\frac{1}{3}}$. We might begin by creating an error by specifying $e^{0.3333}$

So that the propagated error = $e^{0.3333} - e^{\frac{1}{3}} = - 0.0000465196$

Then we might truncate the series after 5 term, leading to the truncation error

$$= -\left(\frac{(0.3333)^5}{5!} + \frac{(0.3333)^6}{6!} + \ldots\right) = -0.0000362750$$

Finally we might sum with the rounded values:

$$1 + 0.3333 + 0.0555 + 0.0062 + 0.00005 = 1.3955$$

Where the propagated error from the rounding's is -0.00002963

The error is called **inherent error**

When performing computations with approximate numbers, we naturally carry the errors of original detain to the final result. In this respect errors of operation are inherent

1.8 SIGNIFICANT DIGITS, THE NUMBER OF CORRECT DIGITS

If α is a positive number it can be represented as a terminating or non-terminating decimal as follows.

$$\alpha = \alpha_m 10^m + \alpha_{m-1} 10^{m-1} + \alpha_{m-2} 10^{m-2} + \ldots + \alpha_{m-n+1} 10^{m-n+1} + \ldots \qquad \ldots..(1.2)$$

where α_i are digits of the number α [i.e α_i = 0,1,2 ,3, . . ,9]

$\alpha_m \neq 0$ is called the leading digit m is the highest power of ten in the expansion. It is an integer

For example consider the number 5214. 73. It can be written as follows:

5214. 73 = $5 \bullet 10^3 + 2 \bullet 10^2 + 1 \bullet 10^2 + 4 \bullet 10 + 7 \bullet 10^{-1} + 3 \bullet 10^{-2} + \ldots$

1.9 SIGNIFICANT DIGITS

A **significant digit** of an approximate number is any non-zero digit, in its decimal representation or any zero lying between significant digits or used as placeholder, to indicate a retained place. All the other zeros of the approximate number that serve

Only to fix the position of the decimal point are not be considered as significant digits.

For example consider the number 0.007040. The first three zeros are not significant digits, since they serve only to fix the position of decimal point and indicate the place values of the other digits. The other two zeros are significant digits since the first lies between the digits 7 and 4 and the second shows that we retain the decimal place 10^{-6} in the approximate number. If the last digits of 0.007040, then the number must be written as 0.00704. From this point of view the numbers 0.007040 and 0.00704 are not the same because

The former has four significant digits and the latter has only three. When writing large numbers, the zeros on the right can serve both to indicate the significant digits and to fix the place values of other digits. This can lead to misunderstanding when the numbers are written in the ordinary way.

1.10 CORRECT DIGITS

In this section we now introduce the notion of correct digits of an approximate numbers.

Definition: If the first n significant digits of an approximate number are correct if the absolute error of the number does not one half unit in the nth place counting from left to right

If α is an approximate number given by (1.2) which takes the place of an exact number X we know that

$$| X - \alpha | \leq \frac{1}{2} 10^{m-n+1}$$

Then by definition the first n digits

$\alpha_m, \alpha_{m-1}, \alpha_{m-2}, \ldots, \alpha_{m-n+1}$ of this number are correct for example consider the exact number X = 25.97. Then with respect to X, the number α = 26 .00 is an approximation correct to three digits, since

$$| X - \alpha | = |25.97 - 26.00| = 0.03 \leq \frac{1}{2} 10^{m-n+1}$$

$$\Rightarrow \mathbf{|25.97 - 26.00| = 0.03 \leq \frac{1}{2} \times (0.1)}$$

1.11 GENERAL ERROR FORMULA

In this section we derive a general formula for determining error committed in using certain functions.

Let $\quad u = f(x_1 , x_2 , \ldots , x_n)$

Be a differentiable function in the variables $x_1 , x_2 , \ldots ,$ and x_n

Then we get $u + \Delta u = f(x_1 + \Delta x_1 , x_2 + \Delta x_2 , \ldots , x_n + \Delta x_n)$

$$\Rightarrow \Delta u = u + \Delta u - u$$

$$= f(x_1 + \Delta x_1 , x_2 + \Delta x_2 , \ldots , x_n + \Delta x_n) - f(x_1 , x_2 , \ldots , x_n)$$

Expanding the right handed side by Taylor's series we get

Δu = f(x_1 , x_2 , . . . , x_n) +$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i$ + terms involving $(\Delta x_i)^2$ and other higher order terms which are negligible

$- \text{ f}(x_1 , x_2 , \ldots , x_n)$

Neglecting squares and higher powers Δx_i we have

$$\Delta u \approx \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i = \frac{\partial f}{\partial x_1} \Delta x_1 + \frac{\partial f}{\partial x_2} \Delta x_2 + \ldots + \frac{\partial f}{\partial x_n} \Delta x_n .. \quad \text{.....(1.3)}$$

The above equation is the equation for the absolute error of u.

Dividing (1.2) by u, we get

$$\frac{\Delta u}{u} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\Delta x_i}{u} = \frac{\partial f}{\partial x_1} \frac{\Delta x_1}{u} + \frac{\partial f}{\partial x_2} \frac{\Delta x_2}{u} + \ldots + \frac{\partial f}{\partial x_n} \frac{\Delta x_n}{u} \quad \text{.....(1.4)}$$

which is the formula for finding the Relative Error hence we have

$$E_R = \frac{\Delta u}{u} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\Delta x_i}{u} = \frac{\partial f}{\partial x_1} \frac{\Delta x_1}{u} + \frac{\partial f}{\partial x_2} \frac{\Delta x_2}{u} + \ldots + \frac{\partial f}{\partial x_n} \frac{\Delta x_n}{u} \quad \text{.....(1.5)}$$

Remarks: Let $| \Delta x_i |$, $(\mathrm{I} = 1,2, \ldots , \mathrm{n})$ be absolute errors of the arguments of the function .Then the absolute error of the function is

$$|\Delta \text{ u}| = |\text{f}(x_1 + \Delta x_1 , x_2 + \Delta x_2 , \ldots , x_n + \Delta x_n) - \text{f}(x_1 , x_2 , \ldots , x_n)|$$

Expanding by Taylor's theorem and proceeding as mentioned above we get

$$|\Delta \text{ u}| = |\text{df}(x_1 , x_2 , \ldots , x_n)| = \left|\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i \right|$$

$$\leq \sum_{i=1}^{n} \left|\frac{\partial f}{\partial x_i}\right| |\Delta x_i|$$

Thus $\quad |\Delta \text{ u}| \leq \sum_{i=1}^{n} \left|\frac{\partial f}{\partial x_i}\right| |\Delta x_i|$

Theorem 1: If a positive approximate number a has n correct digits in the narrow sense, the relative error δ of this number does not exceed $(\frac{1}{10})^{n-1}$ divided by the first significant digit of the given number, or $\delta \leq \frac{1}{\alpha_m} (\frac{1}{10})^{n-1}$

Cor 1: If for the limiting relative error of the number $\alpha_m$we can take

$$\delta_\alpha = \frac{1}{\alpha_m} (\frac{1}{10})^{n-1}$$

where α_m is the first significant digit of the number α_m.

Cor 2: α has more than two correct i.e. n $\geq$ 2, then for all practical purposes the following formula If the number holds

$$\delta_\alpha = \frac{1}{2\alpha_m} (\frac{1}{10})^{n-1}$$

1.12 ERROR OF A SUM

Theorem: The absolute error of an algebraic sum of several approximate numbers does not exceed the sum of the absolute errors of the numbers

Proof: Let u_1,u_2, . . . , u_n denote the n numbers. Let u denote the algebraic sum of these numbers.

We have $u = \pm u_1 \pm u_2 \pm . . . \pm u_n$

$\Rightarrow$ $|u| = |u_1| + |u_2| + . . + |u_n|$

Hence, we have

$\Rightarrow$ $|\Delta u| \leq |\Delta u_1| + |\Delta u_2| + \quad . . . + |\Delta u_n|$

Cor 3: For the limiting point absolute error of an algebraic we can take the sum of the limiting absolute errors of terms

$$\Delta_u = \Delta_{u_1} + \Delta_{u_2} + . . . + \Delta_{u_n}$$

From the above Inequality is follows that the limiting absolute error of the sum cannot be less that the least accurate term, which is to say the term having the maximum absolute error.

1.13 RULES FOR THE ADDITION OF APPROXIMATE NUMBERS

(i) Find the numbers with the least member number of decimal places and leave them unchanged

(ii) Round off the remaining numbers, retaining one or two more decimal places than those with the smallest number of decimals

(iii) Add the numbers, taking into account, taking into account all retail decimals

(iv) Round off the result, reducing it by one decimal

The rounding error of the sum does not exceed

$$\Delta_{max} \leq n \bullet \frac{1}{2} \bullet 10^m$$

Theorem 2: If the terms one and the same sign , the same sign ,the relative error of their sum does not exist the maximum limiting error of any of the terms

i.e. $\delta_u \leq \max(\delta_{u_1}, \delta_{u_2}, . . . , \delta_{u_n})$

1.14 ERROR OF DIFFERENCE

Let u denote the difference between approximate numbers u_1 and u_2

Then we have u = $u_1 - u_2$

The limiting absolute error of the difference is

$$\Delta_u = \Delta_{u_1} + \Delta_{u_2}$$

Hence the limiting absolute error of difference is equal to the sum of the limiting absolute error of the difference is diminuend

$$\delta_u = \frac{\Delta_{u_1} + \Delta_{u_2}}{E}$$

Where E is the exact value of the absolute magnitude of the difference between the numbers $u_1 and\ u_2$

Sol and the numbers with examples

Example 1: Find the percentage error in computing

y = $3x^2$– 6x at x =1, if the error in x is 0.05

Solution: We have y = $3x^2$– 6x, Δx = 0.05

Differentiating we get

$$\frac{dy}{dx} = 6x - 6 \Rightarrow dy = (6x - 6) \quad dx \Rightarrow$$

$$E_A = (6x - 6) \quad \Delta x = 0, \text{ at } x = 1$$

$$E_R = \frac{E_A}{X} = 0 \qquad \text{at } X = 1$$

Example 2: The height of a tower was estimated to be 50 m

Using Theodolite But the height was 45. Calculate the absolute error, relative error and percentage error involved in the measurement

Solution: We have Actual height = $X_E = 45$ m

Estimated height = $X_A = 50$ m

Absolute error = E_A = | $X_E - X_A$ | = |50 −45|

= 5 m

Relative Error = $E_R = \frac{E_A}{X} = \frac{5}{45}$ = 0.111

Percentage error = $E_R \times 100$ = 11.1 %

Example 3: If U = 10 $x^2y^2z^3$ and errors involved in x, y, z are 0.01, 0.02, 0.03 respectively are x = 1, y = 2, z = 3. Calculate the absolute error, relative error, and percentage relative error involved in evaluating u.

Solution: It is given that u = 10 $x^2y^2z^3$

X = 1, y = 2, z = 3 and ΔX = 0.0 1, Δy = 0.0 2, Δz = 0.0 3

Exact value = u = 10 $1^2 2^2 3^3$ = 1080

We have absolute error

$$\Delta u = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \frac{\partial u}{\partial z}\Delta z$$

$$= 20\ x\,y^2z^3\Delta x + 20\ x^2\ yz^3\Delta y + 30x^2y^2z^2\Delta z$$

$$= 20\,(1)\,(2^2)\,(3^3)\,(0.01) + 20\,(1^2)\,(2)\,(3^3)\,(0.02)$$

$$+ 30\,(1^2)\,(2^2)\,(3^2)\,(0.03) = 140.4 = 75.6$$

The Relative Error = $E_R = \frac{E_A}{X} = \frac{75.6}{1080}$ = 0.07

Percentage Error = 100 E_R = 100 × 0.07 = 7.13 = 7%

Example 4: What is the limiting relative error if n = 3 & a_m = 3.

Solution: From the given data we have n = 3, $\alpha_m = 3$

Therefore we get

Using Cor 2

$$\delta_\alpha = \frac{1}{2\alpha_m}\left(\frac{1}{10}\right)^{n-1} = \frac{1}{2.3}\left(\frac{1}{10}\right)^{3-1} = \frac{1}{6}\left(\frac{1}{10}\right)^{2} = \frac{1}{6}\%$$

Example 5: Young's modulus is determined from the deflection of a rod, a and b are the dimensions of the cross section

$$E = \frac{1}{4}\cdot\frac{l^3\,p}{a^3bs}$$

where l = length of the rod, a and b are the dimensions of the cross section, s is the bending deflection, and p is the load. Compute the limiting relative error in a determination of young's modulus E if p = 20 kg, δ_p = 0.1 %, a = 3mm, b = 44mm

δ_b= 1 %, l = 50 cm, δ_l = 1%, s = 2.5 cm, δ_s = 1%,

S is the bending deflection and p is the load. Compute the limiting relative error in a determination of Young's modulus E. If p =20 kgs,

Solution: It is given that

$$E = \frac{1}{4}\cdot\frac{l^3\,p}{a^3bs}$$

Taking logarithms on both sides, we get

Ln E = 3 ln + ln p – 3 ln a - ln b – ln s – ln 4

Replacing increments by differentials, we get Relative error

$$= E_R = \frac{\Delta E}{E} = 3\frac{\Delta l}{l} + \frac{\Delta p}{p} - 3\frac{\Delta a}{a} - \frac{\Delta b}{b} - \frac{\Delta s}{s}$$

The relative error = 3 .0 × 0.01 + 0.00 1 + 3 .0 × 0.01 + 0.01 + 0.01 = 0.081

Therefore the 8 % Error

EXERCISE

1. Define the terms Absolute error and Relative Error
2. Briefly explain "Round off rule "
3. Define percentage error
4. Round off the following to three decimals
   (i) 2.3645 [***Ans:*** 2.364]
   (ii) 4.3455 [***Ans:*** 4.346]
5. Round off the following numbers to 4 significant digits
   (i) 63.38257 [***Ans:*** 63.38]
   (ii) 0.009231542 [***Ans:*** 0.009232]
   (iii) 0.2537514 0 [***Ans:*** 0.2538]
6. If the number N is correct upto 3 significant digits, then what will be the maximum relative error
7. Round off each of the following numbers to three significant figures
   (i) 58.56258 [Ans: 58.6]
   (ii) 0.0039417 [Ans: 0.00394]
8. Find the percentage error in approximating in computation of x – y for x = 12.05 and y = 8.02 having absolute errors

   Δx = 0.0005, Δy = 0.001

   [***Hint:*** Apply, Relative Error = $\frac{\Delta x - \Delta y}{x - y}$ = 0.001]
9. If π = 3.14 instead of $\frac{22}{7}$, find the relative error and percentage error.

   [***Ans:*** 0.00093, 0.093]
10. Round-off the number 4.5126 to four significant figures and find the percentage error. [: – 0.0088]
11. Calculate the value of e^x at 0.75 [***Ans:*** 2.12]

12. Find the limiting absolute and relative errors of the volume of a sphere V = $\frac{1}{6}\pi d^3$ if the diameter d = 3.7 $\pm$ 0.05 cm and π = 3.14 [***Ans:*** 4%]

13. Given u = xy +yz + zx, find the estimate of relative percentage error in the evaluation of u for x = 2.104

    Y= 1.935 and z = 0.845, which are the Approximate values correct to the last digit [***Ans:*** 0.062]

14. Find the sum of following approximate numbers, correct

    To the last digits o.348, 0.1834, 345.4, 235.2 11.75, 0.0849, 0.0002435 and 0.0214 [***Ans:*** 602.2]

15. The length x and the width and y of a plate is measured accurate up to 1 cm as x = 5.43 m and y 3.82 m. Area Find the area of the plate and indicate its error [$\boldsymbol{Ans}$: 0.925 m^2]

16. $Given\ f\ (x\ \ y, z\ )\ =\ \frac{5x\ y^2}{z^2}$: Find the relative maximum error in the absolute in the evaluation of f (x, y, z) at x = y = z = 1 is 5 and if Δ x = 0.1, Δy = 0.1, Δz = 0.1 are the of x, y, z absolute errors [***Ans:*** Hint: we have]

    $$\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \frac{\partial f}{\partial z}\Delta z \Rightarrow$$

    $$|\Delta f_{max}| = \left|\frac{\partial f}{\partial x}\Delta x\right| + \left|\frac{\partial f}{\partial y}\Delta y\right| + \left|\frac{\partial f}{\partial z}\Delta z\right|$$

    $$= \left|\frac{5\ y^2}{z^2}\Delta x\right| + \left|\frac{10x\ y}{z^2}\Delta y\right| + \left|-\frac{5x\ y^2}{z^3}\Delta z\right|$$

    Substituting the given values we get $|\Delta f_{max}|$ = 2.5

    Hence $[E_R\ ]_{max}$ = 0.5]