| 2018–19 Onwards | Elective III : NUMERICAL ANALYSIS | III | 18MST34E |
|---|---|---|---|

**Objective :** To improve the mathematical skills among the PG students

**UNIT – I**
Errors in Numerical Calculations – Introduction – Error and their Computations – Relative Error – A general error formula – Error in series approximation.

**UNIT – II**
Solution of Algebraic and Transcendental equations – Bisection Method – Method of False position – Iteration method – Newton-Raphson method – Horner's method.

**UNIT – III**
Solution of Linear system of equations – Gauss - elimination method – Gauss-Jordon method – Iterative methods – Gauss - Jacobi and Gauss – Seidal methods – Inverse of a matrix by Gauss – Elimination method

**UNIT – IV**
Initial value problems for Ordinary Differential Equations – Introduction – Taylor series method – Euler's method – Modified Euler's method – Runge-Kutta methods – Predictor – Corrector methods – Adam's and Milne's method.

**UNIT – V**
Numerical solution for Partial Differential Equations – Introduction – Finite Difference approximations to derivatives – Laplace's equation – Parabolic equations .

**Text Books:**


1. Dr. B.S. Grewal : " Numerical Methods in Engineering & Science", Khanna Publishers, New Delhi,Fifth Edition, 2000.
2. S.S. Sastry : "Introductory methods of Numerical Analysis", , PHI Learning Pvt Ltd, New Delhi, Fifth Edition, 2013.


**Reference Books:**


1. Dr.M.K. Venkataraman : "Numerical Methods in Science and Engineering", the National Publishing Company, Chennai, Fifth Edition 2001.
2. P.Kandasamy, V.Thilagavathy, K.Gunavathi : "Numerical Methods",S.Chand& Company Ltd, New Delhi, 2016.

# UNIT-I

## APPROXIMATIONS AND ERRORS IN COMPUTATION

### Introduction:

The limitations of analytical methods in practical applications have led scientists and engineers to evolve numerical methods. We know that exact methods often fail in drawing plausible infrences from a given set of tabulated data or in finding roots of transcendental equations or in solving non-linear differential equations. There are many more such situations where analytical methods are unable to produce desirable results. Even if analytical solutions are ~~valu~~ available, these are not amenable to direct numerical interpretation. The aim of numerical analysis is therefore, to provide constructive methods for obtaining answers to such problems in a numerical form.

With the advent of high speed computers and increasing demand for numerical solution to various problems, numerical techniques have become indispensible tools in the hands of engineers and scientists.

The input information is rarely exact since it comes from some measurement or the other and the method also introduces further error. As such, the error in the final result may be due to error in the initial data or in the method or both. Our effort will be to minimize these errors, so as to get best possible results.

We therefore begin by explaining various kinds of approximations and error which may occur in a problem and derive some results on error propogation in numerical Calculations.

## Accuracy of Numbers:

### 1. Approximate numbers:

There are two types of numbers exact and approximate. Exact numbers are 2, 4, 9, 13, $7/2$, 6.45, .... etc., But there are numbers such as $\frac{4}{3}$ (= 1.33333 .....), $\sqrt{2}$ (= 1.414213 ......) and $\pi$ (= 3.141592 .....) which cannot be expressed by a finite number of digits. These may be approximated by numbers 1.3333, 1.4142 and 3.1416 respectively. Such numbers which represent the given numbers to a Certain degree of accuracy are called approximate numbers.

### 2. Significant figures:

The digits used to express a number are called Significant digits (figures). Thus each of the numbers 7845, 3.589, 0.4758 contains four significant figures while the numbers 0.00386, 0.000587 and 0.0000296 contain only three Significant figures since zeroes only help to fix the position of the decimal point. Similarly the numbers 45000 and 7300.00 have two Significant figures only.

### 3. Rounding off:

There are numbers with large number of digits. (e.g) $22/7 = 3.142857143$.

It is desirable to limit such numbers to a manageble number of digits such as 3·14 or 3·143. This process of dropping unwanted digits is called rounding off.

4. Rule: to round off a number to n significant figures:

i) Discard all digits to the right of the $n^{th}$ digit.

ii) If this discarded number is,

a) less than half a unit in the $n^{th}$ place, leave the nth digit unchanged;

b) greater than half a unit in the $n^{th}$ place, increase the $n^{th}$ digit by unity;

c) Exactly half a unit in the $n^{th}$ place, increase the $n^{th}$ digit by unity if its odd otherwise leave it unchanged.

For instance, the following numbers rounded off to three significant figures are:

7·893 to 7·89
12·865 to 12·9
6·4356 to 6·44
3·567 to 3·57
84767 to 84800
5·8254 to 5·82

Also the numbers 6·284359, 9·864651, 12·464762 rounded off to four places of decimal at 6·2844, 9·8646, 12·4648 respectively.

The numbers thus rounded off to n significant figures (or n decimal places) are said to be correct to 'n' significant figures (or n decimal places).

ERRORS:

In any numerical computation, we come across the following type of errors:

1. Inherent Errors:

Errors which are already present in the Statement of a problem before its Solution, are called inherent errors. Such errors arise either due to the given data being approximate or due to the limitations of mathematical tables.. Calculations on the digital computer Inherent errors can be minimized by taking better data or by using high precision computing aids.

2. Rounding Errors:

Rounding error asise from the process of rounding off the numbers during the computation. Such errors are unavoidable in most of the Calculations due to the limitations of the computing aids. Rounding errors can, however, be reduced:

i) by changing the calculation procedure so as to avoid subtraction of nearly equal numbers or divison by a Small number;

(or)

ii) by retaining atleast one more significant figure at each step than that given in the data and rounding off at last step.

## 3. Truncation Errors :

Truncation errors are caused by using approximate results or on replacing an infinite process by a finite one. If we are using a decimal computer having a fixed word length of 4 digits, rounding off of 13.658 gives 13.66 whereass truncation gives 13.65.

For Example, if

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots\ldots \infty = X \text{ (say)}.$$

is replaced by,

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} = X' \text{ (say)},$$ then the truncation

error is $X - X'$.

Truncation error is a type of algorithm error.

## 4. Absolute, Relative and percentage errors:

If $X$ is the true value of a quantity and $X'$ is its approximate value, then $|X - X'|$. (i,e) $|\text{Error}|$ is called the absolute error $E_a$.

The relative error is defined b,

$$E_r = \left|\frac{X - X'}{X}\right|. \quad \text{(i,e)} \quad \frac{|\text{Error}|}{|\text{True value}|}.$$

and the percentage error is,

$$E_p = 100\, E_r = 100 \left|\frac{X - X'}{X}\right|$$

If $\bar{X}$ be Such a number that $|X - X'| \le \bar{X}$, then $\bar{X}$ is an upper limit on the magnitude of absolute error and measures the absolute accuracy.

Notes:

1. The relative and percentage errors are independent of the units used while absolute error is expressed in terms of these units.

2. If a number is correct to n decimal places then the error $= \frac{1}{2} 10^{-n}$.

For Example, if the number is 3.1416 correct to 4 decimal places, then the error $= \frac{1}{2} 10^{-4} = 0.00005$.

## Useful Rules for Estimating Errors:

To estimate the errors which creep in when the numbers in a calculation are ~~truncated~~ truncated or rounded off to a certain number of digits. The following rules are useful.

If the approximate value of a number $x$ having $n$ decimal digits is $x'$, then

1. Absolute error due to truncation to $k$ digits

$$= |x - x'| < 10^{n-k}$$

2. Absolute error due to rounding off to $k$ digits

$$= |x - x'| < \frac{1}{2} 10^{n-k}$$

3. Relative error due to truncation to $k$ digits.

$$= \left| \frac{x - x'}{x} \right| < 10^{1-k}$$

4. Relative error due to rounding off to $k$ digits

$$= \left| \frac{x - x'}{x} \right| < \frac{1}{2} 10^{1-k}$$

Notes:

1. If a number is correct to $n$ significant digits, then the maximum relative error $\le \frac{1}{2} 10^{-n}$.

If a number is correct to $d$ decimal places, then the absolute error $\le \frac{1}{2} 10^{-d}$.

2. If the first significant figure of a number is $k$ and the number is correct to $n$ significant figures, then the relative error $< 1/(k \times 10^{n-1})$.

Let us verify this result by finding the relative error in the number 864.32 correct to five significant figures.

Here, $k = 8$, $n = 5$ and

absolute error $> 0.01 \times \frac{1}{2} = 0.005$

$$\therefore \text{Relative error} \le \frac{0.005}{864.32} = \frac{5}{864320} = \frac{1}{2 \times 86432} < \frac{1}{2 \times 80000}$$

$$= \frac{1}{2 \times 8 \times 10^4} < \frac{1}{8 \times 10^4}$$

$$\text{(i,e)} \; \frac{1}{k \times 10^{n-1}}$$

$\therefore$ Hence the result is verified.

---

Problem:

Round off the numbers 865250 and 37.46235 to four significant figures and compute $E_a, E_r, E_p$ in each case.

Solution:

i) Number rounded off the to four significant figures = 865200

$$E_a = |X - X_1| = |865250 - 865200| = 50$$

$$E_r = \left|\frac{X - X_1}{X}\right| = \frac{50}{865250} = 6.71 \times 10^{-5}$$

$$E_p = E_r \times 100 = 6.71 \times 10^{-3}$$

ii) Number rounded off to four significant figures $= 37.46$

$$E_a = |x - x_1| = |37.46235 - 37.46000|$$
$$= 0.00235$$

$$E_r = \left|\frac{x - x_1}{x}\right| = \frac{0.00235}{37.46235} = 6.27 \times 10^{-5}$$

$$E_p = E_r \times 100 = 6.27 \times 10^{-3}.$$

---

## Error Propagation:

A number of computational steps are carried out for the solution of a problem. It is necessary to understand the way the error propagates with progressive computation.

If the approximate values of two numbers $x$ and $y$ be $x'$ and $y'$ respectively, then the absolute error

$$E_{ax} = x - x' \text{ and } E_{ay} = y - y'$$

1. Absolute error in addition operation:

$$x + y = (x' + E_{ax}) + (y' + E_{ay})$$
$$= x' + y' + E_{ax} + E_{ay}$$

$$\therefore |(x+y) - (x'+y')| = |E_{ax} + E_{ay}| \leq |E_{ax}| + |E_{ay}|$$

Thus the absolute error in taking $(x' + y')$ as an approximation to $(x+y)$ is less than or equal to the sum of the absolute errors in taking $x'$ as an approximation to $x$ and $y'$ as an approximation to $y$.

2. Absolute error in subtraction operation:

$$X-Y = (X'+E_{ax}) - (Y'+E_{ay})$$
$$= (X'-Y') + (E_{ax}-E_{ay})$$

$$\therefore |(X-Y)-(X'-Y')| = |E_{ax}-E_{ay}| \leq |E_{ax}| + |E_{ay}|.$$

Thus the absolute error in taking $(X'-Y')$ as an approximation to $(X-Y)$ is less than or equal to the sum of the absolute errors in taking $X'$ as an approximation to $X$ and $Y'$ as an approximation to $Y$.

3. Absolute error in multiplication operation.

To find the absolute error $E_a$ in the product of two numbers $X$ and $Y$, we write,

$$E_a = (X+E_{ax})(Y+E_{ay}) - XY$$

Where $E_{ax}$ and $E_{ay}$ are the absolute errors in $X$ and $Y$ respectively. Then,

$$E_a = XE_{ay} + YE_{ax} + E_{ax}E_{ay}$$

Assuming $E_{ax}$ and $E_{ay}$ are reasonably small so that $E_{ax}E_{xy}$ can be ignored.

Thus, $E_a = XE_{ay} + YE_{ax}$ approximately.

4. Absolute Error in division operation:

Similarly the absolute error $E_a$ in the quotient of two numbers $X$ and $Y$, we write,

$$E_a = \frac{X+E_{ax}}{Y+E_{ay}} - \frac{X}{Y} = \frac{YE_{ax}-XE_{ay}}{Y(Y+E_{ay})}$$

$$= \frac{YE_{ax}-XE_{ay}}{Y^2(1+E_{ay}/Y)}$$

$$E_a = \frac{YE_{ax} - XE_{ay}}{Y^2}$$, assuming $E_{ay}/Y$ to be small.

$$= \frac{X}{Y}\left(\frac{E_{ax}}{X} - \frac{E_{ay}}{Y}\right)$$

Problem:

Find the absolute error and relative error in $\sqrt{6}+\sqrt{7}+\sqrt{8}$ correct to 4 significant digits.

Solution:

We have $\sqrt{6} = 2.449$,

$\sqrt{7} = 2.646$

$\sqrt{8} = 2.828$

$S = \sqrt{6}+\sqrt{7}+\sqrt{8} = 7.923$.

Then the absolute error $E_a$ in $S$, is

$$E_a = 0.0005 + 0.0007 + 0.0004 = 0.0016$$

This shows that $S$ is correct to 3 significant digits only. Therefore, we take $S = 7.92$.

Then the relative error $E_r$ is,

$$E_r = \frac{0.0016}{7.92} = 0.0002.$$

Error in the approximation of a function

Let $y = f(x_1, x_2)$ be a function of two variables, $x_1, x_2$. If $\delta x_1, \delta x_2$ be the errors in $x_1, x_2$, then the error $\delta y$ in $y$ is given by,

$$y + \delta y = f(x_1 + \delta x_1 + x_2 + \delta x_2)$$

Expanding the right hand side by Taylor's series, we get,

$$y+\delta y = f(x_1, x_2) + \left(\frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2\right)$$

+ terms involving higher powers of $\delta x_1$ and $\delta x_2 \rightarrow (1)$

If the errors $\delta x_1, \delta x_2$ be so small that their squares and higher powers can be neglected, than (1) gives,

$$\delta y = \frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2 \text{ approximately.}$$

$\therefore$ Hence, $\delta y = \frac{\partial y}{\partial x_1} + \frac{\partial y}{\partial x_2}$

$$\delta y = \frac{\partial y}{\partial x_1}\delta x_1 + \frac{\partial y}{\partial x_2}\delta x_2$$

In general, the error $\delta y$ in the function $y = f(x_1, x_2, \ldots x_n)$ Corresponding to the errors $\delta x_i$ in $x_i$ $(i = 1, 2, \ldots n)$ is given by,

$$\delta y = \frac{\partial y}{\partial x_1}\delta x_1 + \frac{\partial y}{\partial x_2}\delta x_2 + \ldots\ldots + \frac{\partial y}{\partial x_n}\delta x_n$$

and the relative error in $y$ is,

$$E_r = \frac{\delta y}{y} = \frac{\partial y}{\partial x_1}\frac{\delta x_1}{y} + \frac{\partial y}{\partial x_2}\frac{\delta x_2}{y} + \ldots + \frac{\partial y}{\partial x_n}\frac{\delta x_n}{y}.$$

---

problem:

If $u = 4x^2y^3/z^4$ and errors in $x, y, z$ be $0.001$, compute the relative maximum error in $u$ when $x = y = z = 1$.

Solution:

Since, $\frac{\partial u}{\partial x} = \frac{8xy^3}{z^4}$, $\frac{\partial u}{\partial y} = \frac{12x^2y^2}{z^4}$,

$$\frac{\partial u}{\partial z} = -\frac{16x^2y^3}{z^5}$$

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

$= \frac{8xy^3}{z^4}\delta x + \frac{12x^2y^2}{z^4}\delta y - \frac{16x^2y^3}{z^5}\delta z$

Since the errors $\delta x, \delta y, \delta z$ may be positive or negative, we take the absolute value of the terms on the right side, giving

$$(\delta u)_{max} = \left|\frac{8xy^3}{z^4}\delta x\right| + \left|\frac{12x^2y^2}{z^4}\delta y\right| + \left|\frac{16x^2y^3}{z^5}\delta z\right|$$

$= 8(0.001) + 12(0.001) + 16(0.001)$

$= 0.036$

Hence the maximum relative error $= (\delta u)_{max}/u$ ~~= 0.036~~

$= \frac{0.036}{4}$

$= 0.009$

---

Errors in a Series approximation:

We know that the Taylor's series for $f(x)$ at $x=a$ with a remainder after n terms is,

$$f(x) = f(a + \overline{x-a}) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \dots\dots + \frac{(x-a)^{n-1}}{(n-1)!}f^{n-1}(a) + R_n(x)$$

where, $R_n(x) = \frac{(x-a)^n}{n!}f^n(\theta),\ a < \theta < x$

If the series is convergent, $R_n(x) \to 0$ as $n \to \infty$ and hence if $f(x)$ is approximated by the first n terms of this Series, the maximum error will be given by the remainder term $R_n(x)$. On the other hand, if the accuracy required in a Series approximation is preassigned, then we can find n, the number of terms which would yield the desired accuracy.

Problem:

Find the number of terms of the exponential Series such that their Sum gives the value of $e^x$ correct to six decimal places at $x=1$.

Solution:

We have,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^{n-1}}{(n-1)!} + R_n(x) \longrightarrow (1)$$

where, $R_n(x) = \frac{x^n}{n!} e^{\theta}, \; 0 < \theta < x.$

$\therefore$ Maximum absolute error (at $\theta = x$) $= \frac{x^n}{n!} e^x$ and

The Maximum relative error $= \frac{x^n}{n!}$

Hence $(E_r)_{max}$ at $x=1$ is $\frac{1}{n!}$

For a Six decimal accuracy at $x=1$, we have

$$\frac{1}{n!} < \frac{1}{2} 10^{-6} \;;\; (i.e)\; n! > 2 \times 10^6$$

which gives $n = 10$.

Thus we need 10 terms of the series (i) in order that its sum is correct to 6 decimal places.

ORDER OF APPROXIMATION:

We often replace a function $f(h)$ with its approximation $\phi(h)$ and the error bound is known to be $\mu(h^n)$, n being a positive integer so that,

$$|f(h)-\phi(h)| \le \mu |h^n| \text{, for sufficiently small } h.$$

Then we say that $\phi(h)$ approximates $f(h)$ with order of approximation $O(h^n)$ and write,

$$f(h)=\phi(h)+O(h^n)$$

For instance,

$$\frac{1}{1-h}=1+h+h^2+h^3+h^4+h^5+\cdots\cdots$$

is written as,

$$\frac{1}{1-h}=1+h+h^2+h^3+O(h^4) \longrightarrow ①$$

to the 4th order of approximation,

Similarly, $\cos(h)=1-\frac{h^2}{2!}+\frac{h^4}{4!}-\frac{h^6}{6!}+\frac{h^8}{8!}-\cdots$

to the 6th order of approximation. becomes,

$$\cos(h)=1-\frac{h^2}{2!}+\frac{h^4}{4!}+O(h^6) \longrightarrow ②$$

The sum of ① and ② gives,

$$(1-h)^{-1}+\cos(h)=2+h+\frac{h^2}{2!}+h^3+O(h^4)+\frac{h^4}{4!}O(h^6) \longrightarrow ③$$

Since,

$$O(h^4)+\frac{h^4}{4!}=O(h^4) \text{ and}$$

$$O(h^4)+O(h^6)=O(h^4)$$

**Source** :

Text Books:

1. Dr . B.S. Grewal: " Numerical Methods in Engineering & Science", Khanna Publishers, New Delhi, Fifth Edition, 2000.