Numerical Analysis course of B.Tech (ECE)

Approximation in Numerical Computation

Approximations and errors are integral part of human life and they can not be avoided. Errors comes in varity of forms and sizes, some may be avoidable but some are not. To get a final solution of a problem we have to minimize the error.

Actually in numerical analysis we deviate from conventional mathematics and try to analyse the numerical properties of given data and compute the solution with some sacrifice of the exact solution, so that cost of computation can be lowered and helpful in computing with computer. So instead of exact solution we will go for approximate solution here due to lower costs of computation.

Errors

Basically errors are of two types Inherent errors and Numerical errors.

Inherent errors are such errors which is inherent in the given data. For example, when we are going to measure the length of line segment then it can not be exact or same all the time.

Another type of inherent error is the error due to data conversion process. For example if we want to convert 0.1_{10} to binary, it will come as $0.0001100110011001 \cdots_2$ never ending

so $0.1_{10} = 0.00011001_2 \longrightarrow$ this is error due to data conversion.

Numerical errors

This are the errors occur when we are going to implement numerical methods.

They are,

Rounding off errors, Truncation errors, chopping off errors, Significant errors

Rounding of errors

Error due to rounding off a number called rounding of errors.

For example.

Say. 1.983027 and we are going to round this number to five decimal places then answer is

1.98303, $\sqrt{3} = 1.732 \ldots\ldots$ upto infinite

$\pi = 3.141519 \ldots\ldots$

if these numbers we use in computers we have to round off or chop off the numbers, then error will be introduced.

Chopping off errors.

If error occurs due to drop of significant digits for some needs then this error may be called as chopping off errors.

For example 2.7351238

if we need a number with maximum 6 decimal places then for representing this number we have to write

either 2.735123, or 2.735124

Cases: when last digit is simple choped — when last digit is adjusted with rounding of facility.

Rounding off numbers:

1.983027 rounding it upto $\overset{n}{5}$th decimal place

then see the $\overset{n+1}{(5+1)} = \overset{(n+1)th}{6th}$ place,

if it is greater than 5, then '+1' is added in $\overset{(n)}{5th}$ place

if " " lesser " 5 " ~~it is~~ $\overset{n+1}{(5+1)}$th place value is simply discarded.

if it is equal to 5 then the ~~5~~ nth place here $\overset{n}{5}$th place is seen, if it is even then the value of $\overset{n+1}{(5+1)}$th place is simply discarded,

if $\overset{(n)}{5}$th place value is odd then '+1' is added in $\overset{(3)}{5}$th place value

1.983024 can be rounded ~~to~~ to 5th decimal places then 1.98302 will be the rounded value.

if 1.9830$\underline{25}$ then rounded value is = 1.98302.

~~if 1.983025~~

if 1.9830$\underline{35}$ then rounded value is = 1.98304.

Truncation error:-

This is the error comes when we are going to calculate a value, by ignoring higher order terms from a infinite series.

For example

$$e^x = 1 + x_1 + x^2/2! + \frac{x^3}{3!} + \cdots$$

then if ~~...~~ we ignore the third or higher order terms in calculation of e^x then errors obviously introduced, within the calculation of e^x for a given 'x'. This is called ~~truncation and errors~~ truncation error.

Significant Error:-

This is defined as the error introduced due to loss of significant digits..

For example;

$$\left.\begin{aligned} \text{say } a &= 2.45762 \\ \text{and } b &= 2.45764 \end{aligned}\right\} \cdots (1)$$

$$\left.\begin{aligned} \text{therefore } a &= 2.4576 \text{ upto four decimal places} \\ b &= 2.4576 \text{ " " " } \end{aligned}\right\} (II)$$

From (1). $a - b = 2.45764 - 2.45762$

$= 0.00002$, then rounded upto 4 decimal places = 0

now, if ~~we~~ we are going to evaluate the expression $100b - 110a$

$= 100 \cdot 2.45764 - 110 \cdot 2.45762$

before rounding off $= |245.764 - 270.3382| \cdot 24\ldots$

After a, b, rounded upto four decimal places.
we get.

$$100b - 110a$$
$$= 100(b-a) - 10a$$
$$= -100(a-b) - 10a.$$
$$= -100.0 - 10a.$$
$$= 0 - 10a$$
$$= -10a. = -10.2\cdot4576$$
$$= |-24\cdot5760| = 24\cdot5760$$

So we can see the difference of error.

$$24\cdot5760 - 24\cdot5742$$
$$= \cdot0018.$$

Here this rounding off causes little bit effect, but whenever such expression we are going to evaluate this may cause severe error in the output data.

Absolute Error.

It is denoted as the difference between approximate value and corresponding true value. Approximate value may be called as a estimated value.

True value is denoted as V_T,
Absolute error is " " E_a
Approximate or Estimated value is denoted by V_A

Then $|E_a| = |V_T - V_A|$

Relative error:

is defined as. $|E_r| = \dfrac{|V_T - V_A|}{|V_T|}$

Percentage error: is defined as. $\dfrac{|V_T - V_A|}{|V_T|} \times 100\%$.

Ill conditioned and well conditioned system

A computation process, in which the cumulative error of all input is greatly magnified is said to be numerically unstable.

A system is such that, if we do small change in any input parameter causes massive change in output is called as a inherently unstable system and corresponding problem is called ill conditioned system. Otherwise the system is well conditioned system.

For example,

$$2x_1 + x_2 = 25$$
$$2.001x_1 + x_2 = 25.01,$$

From these, we can get the value

as $x_1 = 10$, $x_2 = 5$,

If we change the 2nd eqn as

$$2x_1 + x_2 = 25$$
$$2.0005x_1 + x_2 = 25.01$$

then $x_1 = 20$, $x_2 = -15$,

So we can see the little change in input coefficient results the huge change in output, so it is a ill-conditioned system.

Error Propagation

This means error in the succeeding steps of a process due to an occurrence of an earlier error.

This propagated error is very very important. If errors are magnified continuously then it can overshadow the true value and can destroy its validity — we say that corresponding method is unstable. If the method continues, and errors will be decreasing state then then method is [illegible]

~~Stable methods~~

General formula for estimation of errors

Let $u = f(u_1, u_2, u_3, \ldots u_n)$ be a differentiable function of independent variable u_i (where $i = 1, 2, 3, \ldots n$) and corresponding absolute errors are $|\Delta u_i|$ (where $i = 1, 2, 3, \ldots n$) ~~respectively~~ and relative errors $\left|\frac{\Delta u}{u}\right|$ ~~of the~~ respectively.

$$|\Delta u| = \Big| f(u_1 + \Delta u_1, u_2 + \Delta u_2, \ldots, u_n + \Delta u_n) - f(u_1, u_2 \ldots u_n) \Big|$$

$$= \Big| f(u_1, u_2, \ldots u_n) + \left[\Delta u_1 \frac{\partial}{\partial u_1} + \Delta u_2 \frac{\partial}{\partial u_2} + \cdots + \Delta u_n \frac{\partial}{\partial u_n} \right] f + \cdots + \cdots - f(u_1, u_2, \ldots u_n) \Big|$$

By neglecting higher order terms.

$$|\Delta u| = \left| \Delta u_1 \frac{\partial f}{\partial u_1} + \Delta u_2 \frac{\partial f}{\partial u_2} + \Delta u_3 \frac{\partial f}{\partial u_3} + \cdots + \Delta u_n \frac{\partial f}{\partial u_n} \right|$$

$$\leq |\Delta u_1| \left|\frac{\partial f}{\partial u_1}\right| + |\Delta u_2| \left|\frac{\partial f}{\partial u_2}\right| + |\Delta u_3| \left|\frac{\partial f}{\partial u_3}\right| + \cdots + |\Delta u_n| \left|\frac{\partial f}{\partial u_n}\right|$$

$$\therefore |\Delta u| \leq \sum_{i=1}^{n} \left|\frac{\partial f}{\partial u_i}\right| |\Delta u_i|.$$

Then the relative error E_r of the function is given by

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

Application of Error formula on Fundamental Operation of Arithmatic.

① **Addition.**

Let u be sum of n approximate numbers $u_1, u_2, u_3, \ldots u_n$.

$\therefore \; u = u_1 + u_2 + u_3 + \cdots + u_n$

$$|\Delta u| = |(u_1+\Delta u_1) + (u_2+\Delta u_2) + \cdots + (u_n+\Delta u_n) - u_1 - u_2 - \cdots - u_n|$$

$$= |\Delta u_1 + \Delta u_2 + \Delta u_3 + \cdots + \Delta u_n|$$

$$\le |\Delta u_1| + |\Delta u_2| + |\Delta u_3| + \cdots + |\Delta u_n|$$

(b) Subtraction

Let u be the difference between two approximate numbers u_1, u_2, then

$$u = u_1 - u_2$$

$$\therefore |\Delta u| = \cancel{|u_1 - u_2|}\ |(u_1+\Delta u_1) - (u_2+\Delta u_2) - (u_1-u_2)|$$

$$= |\Delta u_1 - \Delta u_2|$$

$$\le |\Delta u_1| + |\Delta u_2|$$

~~(c) $u = u_1 u_2 u_3 \cdots u_n$, $\ln u = \sum \ln u_i$, by dividing differentiating both side we get~~

(c) $\therefore u = u_1 u_2 u_3 \cdots u_n$

$\therefore \ln u = \ln u_1 + \ln u_2 + \ln u_3 + \cdots + \ln u_n$

by differentiating both side

$$\frac{1}{u}\Delta u = \frac{\Delta u_1}{u_1} + \frac{\Delta u_2}{u_2} + \cdots + \frac{\Delta u_n}{u_n}$$

$$\therefore E_r = \sum \frac{\Delta u_i}{u_i}$$

and $E_r \le \sum \left|\frac{\Delta u_i}{u_i}\right|$

Q. Write down the approximate representation of $\frac{2}{3}$ correct to four significant figures and then find:

(a) Absolute error, (b) relative error

(c) percentage error.

Ans

$\frac{2}{3} \to$ upto four significant ~~error~~ digit

$= 0.666\ldots$

$= 0.6667.$

So
(a) Absolute error $= \left|\frac{2}{3} - 0.6667\right|$

$= \left|\frac{2 - 2.0001}{3}\right|$

$= \left|\frac{.0001}{3}\right|$

$= .000033$

(b) relative error $= \frac{.000033}{2/3}$

$= \frac{.0000990}{2}$

$= .000495 \cong .00005$

(c) $\therefore$ Relative percentage error.

$.00005 \times 100\% = .005\%$

Q. Find the relative error of computation of $x-y$ for ~~xx~~ $x = 12.05$ and $y = 8.02$, having absolute errors $\Delta x = .005$, and $\Delta y = .001$,

$\therefore$ ~~Ans~~ relative error

$E_r = \left|\frac{\Delta x}{x}\right| - \left|\frac{\Delta y}{y}\right|$

$= \frac{0.005}{12.05} - \frac{.001}{8.02}$

$= 0.000415 - 0.000125$

$= 0.00029.$

Q. If $y = 4x^6 - 5x$ find percentage error at $x=1$ where error in $x = 0.04$

$\therefore$ $y = 4x^6 - 5x$

$\therefore$ $\frac{dy}{dx} = 4.6x^5 - 5 = 24x^5 - 5$

$\therefore$ $\Delta y = \frac{dy}{dx} \cdot \Delta x$

$= (24x^5 - 5) \cdot .04$

$= (24 \cdot 1^5 - 5) \cdot .04$

$= (24 - 5) \times \frac{4}{100} = \frac{76}{100}$

$\therefore$ Percentage error $= \frac{76}{100} \times 100 = 76\%$.

Q. if $f(x) = 4\cos x - 6x$, find the relative percentage error in $f(x)$ for $x = 0$, if the error in $x = 0.005$

Ans.

Now, $f(x) = 4\cos x - 6x$

$\therefore$ $\frac{d}{dx} f(x) = -4\sin x - 6$

$\therefore$ $df(x) = -(4\sin x + 6) \cdot dx$

$\therefore$ $\Delta f(x) = -(4\sin x + 6) \cdot \Delta x$

$\therefore$ relative error $= \left| \frac{\Delta f(x)}{f(x)} \right|$

$= \left| \frac{-(4\sin x + 6) \cdot \Delta x}{(4\cos x - 6x)} \right|$

$= \left| \frac{-(4\sin 0 + 6) \cdot 0.005}{4\cos 0^\circ - 6.0} \right| = \left| \frac{-6 \times 0.005}{4.1 - 0} \right|$

$= \left| -\frac{.03}{4} \right|$

$\therefore$ relative percentage error $= \frac{\delta S}{S} \times 100$

$= 1.75\,\%$.

Problems