

Lesson 2

1. What is round off error?

All calculating devices perform finite precision arithmetic i.e. they cannot handle more than a specified number of digits. Hence, if a computer can handle numbers up to 's' digits only, then the product of two such numbers, which in reality have $2s$ or $2s-1$ digits, can only be stored and hence used in subsequent calculations up to the first s digits only. The effect of such rounding off may accumulate after extensive calculations. In an unstable algorithm this can rapidly lead to erroneous results.

2. What is truncation error?

Truncation errors occur when a limiting process is truncated before it has reached its limiting value. A typical example may be a series which converges after 'n' terms, is truncated i.e. cut-off after the first $n/2$ terms. Another example of truncation error occurs when a nonlinear equation is linearized, e.g. if the nonlinear function $f(x)$ of x is linearized about the value x_0

$$f(x) = f(x_0) + \left. \frac{df}{dx} \right|_{x_0} \underbrace{(x - x_0)}_{\text{stepsize}=h} + \underbrace{\text{quadratic terms}}_{\text{truncation error}}$$

3. How can the truncation error be reduced?

The truncation error can be reduced and one can achieve arbitrarily high accuracy by choosing the step size h to be sufficiently small. However reducing h means increased number of function evaluations and higher computational expense. In order to avoid having to take very small step sizes 'h', one can try approximating $y(x)$ by a higher order polynomial instead of a step wise linear function. In case of linearization of a non-linear function, the truncation error is of the order h^2 ($o(h^2)$)

4. What is the trapezoidal rule?

The trapezoidal rule is an algorithm for integrating a function numerically. For instance if the function $y=f(x)$ is to be integrated between $x = a$ and $x = b$ i.e. we are interested in

evaluating $I = \int_a^b f(x) dx$ numerically. Using the trapezoidal rule, the area between the

curve $y=f(x)$ and the x axis is approximated with the sum $I(h)$ of the areas of a series of trapezoids, each with a constant base width of h .

5. What is p and h refinement?

Using a higher order polynomial to approximate an unknown function is known as ' p ' refinement while using a relatively lower order polynomial (e.g. a linear function) to

approximate an unknown function while reducing the step size to improve accuracy is known as 'h' refinement.

6. What is Richardson extrapolation?

Richardson extrapolation is based on the idea that it is possible to reduce the error in the approximation of a function $y(x)$ by a linear function $y^L(x)$ if $y^L(x)$ is evaluated for different step sizes $h, 2h$ etc. and the results are combined. If for instance we are evaluating the integral $I = \int_a^b f(x) dx$ using trapezoidal rule, then by evaluating I for a step size of h , denoted as $I(h)$, as well as evaluating I for a step size of $2h$, denoted as $I(2h)$, and combining the results as $I(h) + \frac{1}{3}[I(h) - I(2h)]$ it is possible to reduce the truncation error to levels much less than the errors arising from $I(h)$ and $I(2h)$ individually.

7. What is absolute error and relative error?

Suppose a is the true value of a function and \tilde{a} is a numerical approximation, then $\tilde{a} - a$ is the absolute error and $\frac{\tilde{a} - a}{a}$ is the relative error.

8. Given a numerical solution .0001234 and knowing that the error in the solution is less than 5.E-5, what are (a) the number of digits (b) number of decimal places (c) number of significant digits in the solution?

The number of digits is 4. The number of decimal places is 7. The number of significant digits is 2.

9. What are the rules of rounding and why is rounding better than chopping?

The rules of rounding are:

- (a) If the digit to the right of the t^{th} decimal place is **less than** $.5 \times 10^{-t}$, then the t^{th} decimal place is left unchanged
- (b) If the digit to the right of the t^{th} decimal place is **greater than** $.5 \times 10^{-t}$, then the t^{th} decimal place is raised by 1.
- (c) If the digit to the right of the t^{th} decimal place is **exactly equal** to $.5 \times 10^{-t}$, then the t^{th} decimal place is raised by 1 if it is odd, and left unchanged if it is even.

Rounding is better than chopping because the errors due to rounding will tend to cancel out and not accumulate since the probability of the $(t+1)^{\text{th}}$ decimal being greater than the t^{th} decimal is equal to the probability of the $(t+1)^{\text{th}}$ decimal being less than the t^{th} decimal, and the errors arising from rules (a) and (b) have the opposite sign. Similarly since the probability of the t^{th} decimal place being odd or even is equal (each probability being equal to half), if rule (c) is followed the resulting error will be positive or negative equally often. These errors too will thus **tend to cancel off and not accumulate**.

10. What are the error bounds on addition and subtraction?

If \tilde{x}_1 is the numerical approximation to x_1 with absolute error ε_1 and x_2 is the numerical approximation to \tilde{x}_2 with absolute error ε_2 then the bounds on the error in $\tilde{x}_1 - \tilde{x}_2$ and $\tilde{x}_1 + \tilde{x}_2$ are given by:

$$\|(x_1 - x_2) - (\tilde{x}_1 - \tilde{x}_2)\| \leq \varepsilon_1 + \varepsilon_2$$

$$\|(x_1 + x_2) - (\tilde{x}_1 + \tilde{x}_2)\| \leq \varepsilon_1 + \varepsilon_2$$

11. What are the error bounds on multiplication and division?

If \tilde{x}_1 is the numerical approximation to x_1 with relative error r_1 and x_2 is the numerical approximation to \tilde{x}_2 with absolute error r_2 then the bounds on the error in $\tilde{x}_1 \tilde{x}_2$ and $\tilde{x}_1 / \tilde{x}_2$ are given by:

$$\|(x_1 \cdot x_2) - (\tilde{x}_1 \cdot \tilde{x}_2)\| \leq r_1 + r_2$$

$$r_1 \ll 1, r_2 \ll 1$$

$$\|(x_1 / x_2) - (\tilde{x}_1 / \tilde{x}_2)\| \leq r_1 - r_2$$

12. What are the uses of error analysis?

In numerical analysis we are frequently required to add, subtract, multiply or divide variables for which only numerical approximations to their true values are known. If bounds on the error in the numerical approximations are known for individual variables, error analysis enables calculation of bounds on the results of mathematical operations that combine several approximately known variables.

13. What is error due to cancellation of terms and how can one avoid such errors?

Numerical calculations involving subtractions between two numbers, where the difference between the numbers is considerably less than either of the numbers, are prone to large errors. This type of error is known as error due to cancellation of terms. One can avoid such errors by rewriting formulas to avoid subtraction between nearly equal terms.

14. What is the general formula for error propagation?

The general formula for error propagation allows estimation of a bound on the error in the value of y where y is a function of more than one independent variable i.e. $y(x_1, x_2, \dots, x_n)$ if the errors in the independent variables are known. It attempts to find bounds on the error in y , say Δy as a function of the errors in the independent variables $\Delta x_1 = \tilde{x}_1 - x_1, \Delta x_2 = \tilde{x}_2 - x_2, \dots, \Delta x_n = \tilde{x}_n - x_n$. $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ are approximate values of the independent variables x_1, x_2, \dots, x_n .