Numerical Analysis

By

Assistant Professor Nabaa Naeem Mahdi

Chapter One (Theory and method of processing error)

Introduction

Approximate values can be used to solve mathematical or engineering problems instead of their exact real values, and this is what we notice in our practical life.

The approximate solution methods are called algorithms, named after the scientist Al-Khwarizmi, which are defined as a set of procedures for implementing specific mathematical operations that lead to solving the given problems in order to provide the required solutions and accuracy. The solutions obtained are known as numerical solutions.

Numerical analysis can be defined as the methods and theories for finding numerical solutions, and its importance is summarized in obtaining numerical and approximate results for functions that cannot be solved quickly and with high accuracy, and these numerical and approximate results may be accompanied by errors that can be measured.

Another definition Numerical Analysis is a field of mathematics that concerned with the study of approximate solutions of mathematical problems, where it is difficult or impossible to find the exact solutions for these problems.

For instance, we can't find the exact value of the following integral, by using known integration methods $\int_0^1 e^{x^2}dx$

While we can find an approximate value for this integral, using numerical methods.

The other example, if we aim to find the solution of a linear system $Ax = b$

where $A \in R^{n\times n}$, when n is too large , it so difficult to find the exact solutions for this system by hand using known methods, so in this case, it is easier to think about how to find the approximate solution using a suitable algorithm and computer programs.

The importance of Numerical Analysis

To interpret any real phenomena, we need to formulate it, in a mathematical form. To give a realistic meaning for these phenomena we have to choose complicated mathematical models, but the problem is, it so difficult to find explicit formulas to find the exact solutions for these complicated models. Therefore, it might be better to find the numerical solutions for complicated form rather than finding the exact solutions for easier forms that can't describe these phenomena in realistic way.

The Nature of Numerical analysis

Since for any numerical Algorithm (the steps of the numerical method), we have lots of mathematical calculations, we need to choose a suitable computer language such as Matlab ,C^{++}, FORTRAN and write the algorithm processes in programing steps.

In fact, the accuracy of numerical solutions, for any problem, is controlled by three criteria:

1- The type of algorithm,

2- The type of computer language and programs,

3- The advancement of the computers which are used.

MATLAB can be used to solve mathematical problems. This program is a language that was developed specifically to serve practical and engineering applications to give the answer quickly and with the least possible effort. This program appeared in FORTRAN by (Cleve Barry Moler), a mathematician and computer programmer specializing in numerical analysis. Then (John N. Little), the president of (math works) company, participated in writing the MATLAB program in (C) instead of FORTRAN. The origin of the word MATLAB is two words (Matrix Laboratory)

Types and sources of Errors

When we compute the numerical solutions of mathematical model that we use to describe a real phenomenon, we get some errors; therefore we should study the types and sources of these errors.

Error: It is the difference between the real value and the approximate value that represents it. The real value is represented by the symbol x, the approximate value by the symbol $\hat{x}$, and the error by the symbol ε.

$$\varepsilon_x = x - \hat{x}$$

Source errors

The approximate solution to a problem is often due to the accumulation of several types of errors. The sources of error can be classified into the following types:

1-Formulation error:

Analyzing a particular problem in a mathematical way is often formulated in a simplified model that describes the basic problem with a small number of variables. That is, it may neglect some factors or influences if we see that they simplify the model and enable us to solve it. Thus, the results we obtain from this model contain errors called formulation errors. For example, when measuring the speed of a body moving on the surface of the Earth, we often neglect air resistance, so the speedometer is inaccurate and contains errors.

2- Inherent error:

Using data that does not have an exact value in mathematical formulas also leads to inaccurate results, for example:

$$\pi = 3.142857 , \quad e = 2.71828 , \quad \sqrt{2} = 1.41421$$

Talking about the computer applies to other measuring devices, as measuring devices used in scientific laboratories such as (the balance, the clock, the voltmeter, the thermometer) all of these devices are subject to error no matter how accurate they are, because they are of human manufacture, and the materials from which they are made are affected by the surrounding weather conditions such as pressure, temperature and humidity, and the errors resulting from them are called machine errors.

3-Rounding and chopping errors:

Many numbers contain infinite decimal places. When rounding these decimal places and using them, it leads to an error in the results.

Chopping errors: Let us assume that the length of the storage unit in the computer is four ranks and we entered the numbers (0.00034, 0.34195), then the computer will store them in the following form: (0.0003, 0.3419)

Rounding errors: this type of errors can be got, because of the rounding of numbers in computer programming languages. Through it, we choose the value of the rank after the fourth. If it is greater than or equal to (0.5), (1) is added to the fourth rank and what comes after that is cancelled to round the previous numbers as follows: (0.0003, 0.3420)

Or: 4.99…9 round to 5, and 4.0001 round to 4.

4- Truncation errors:

Many mathematical functions are defined in the form of an infinite series. Since calculating this series is impossible, a specific number of terms of the series must be determined. With this determination, there is an error in the result. For example, the following function:

$$f(x) = 1 - \cos(x)$$

The truncated series of the function about the point x=0 is:

$$f(x) = \frac{x^2}{21} - \frac{x^4}{41} + \frac{x^6}{61} + \ldots\ldots$$

5- Accumulated error:

It is the error that occurs from subsequent steps and calculations of the process based on the error that occurred in previous steps. If the error increases more and more as the process continues, then it is said that the formula used for the solution is not convergent (unstable). This is what happens in iterative formulas, noting that the error in stable formulas decreases as the calculation process continues.

Types of Errors

Let $\hat{x}$ is the approximate value of x, there are three methods can be used to measure the errors:-

1- **Absolute Error: -** It is the absolute value of the exact value minus the approximate value according to the formula below:

$$e_x = |x - \hat{x}|$$

2- **Relative error: -** It is the ratio of the absolute error to the exact value it is calculated based on the following formula:

$$R_x = \frac{e_x}{|x|} = \frac{|x - \hat{x}|}{|x|}$$

3- **Percentage error: -** It is the relative error multiplied by 100%, it is calculated based on the following formula:

$PE_x = R_x \times 100\%$

Example:- Let $\hat{x} = 3.14$, $x = 3.141592$. Find the Absolute and Relative errors.

Solution

$$e_x = |x - \hat{x}| = |3.241592 - 3.14| = 0.001592$$

$$R_x = \frac{e_x}{|x|} = \frac{|x - \hat{x}|}{|x|} = \frac{0.001592}{3.141592} = 0.000507$$

Remark: Clearly, $R_x < e_x$

Example:- Let $\hat{x} = 3.26$, $x = 3.257$ find the absolute and relative errors.

<u>**Errors in mathematical operations**</u>

Dealing with approximate values leads to errors in the results of the four mathematical operations. Therefore, it is necessary to know the absolute and relative error for each mathematical operation, where $\hat{x}, \hat{y}$ they are the approximate values of the real values x, y.

(a) Addition process: If they represent approximate values $\hat{x}, \hat{y}$ for x, y, then the absolute error of the sum is:

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

Proof

$$e_{x+y} = (x+y) - (\hat{x} + \hat{y})$$
$$e_{x+y} = x + y - \hat{x} - \hat{y}$$
$$e_{x+y} = (x - \hat{x}) + (y - \hat{y})$$
$$\therefore e_{x+y} = e_x + e_y$$

The calculation of the relative error depends on the following formula:

$$R_{x+y} = \frac{e_x + e_y}{x + y}$$

(b) Subtraction process: When subtracting two approximate numbers $\hat{x}, \hat{y}$, the absolute error of the subtraction operation is obtained.

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

Proof

$$e_{x-y} = (x-y)-(\hat{x}-\hat{y})$$
$$e_{x-y} = x-y-\hat{x}+\hat{y}$$
$$e_{x-y} = (x-\hat{x})-(y-\hat{y})$$
$$\therefore e_{x-y} = e_x - e_y$$

The relative error is calculated according to the following formula:

$$R_{x-y} = \frac{e_x - e_y}{x-y} \quad ,or \quad R_{x-y} = \frac{xR_x - yR_y}{x-y}$$

(c) Multiplication process: When multiplying two approximate numbers $\hat{x}, \hat{y}$, the absolute error of the multiplication operation is produced.

$$e_{x\times y} = xe_y + ye_x$$

Proof

$$e_{x\times y} = (xy)-(\hat{x}\hat{y})$$
$$e_x = x-\hat{x} \Rightarrow \hat{x} = x-e_x \quad ,e_y = y-\hat{y} \Rightarrow \hat{y} = y-e_y$$
$$e_{x\times y} = xy-\left[(x-e_x)(y-e_y)\right]$$
$$e_{x\times y} = xy-xy+xe_y+ye_x-e_xe_y$$
$$e_xe_y = 0$$
$$\therefore e_{x\times y} = xe_y + ye_x$$

The relative error is calculated according to the following formula:

$$R_{x\times y} = R_y + R_x$$

(d) Division process: When dividing the two approximate numbers $\hat{x}, \hat{y}$, the absolute error of the division operation is produced.

$$e_{\frac{x}{y}} = \frac{x}{y}\left(\frac{e_x}{x} - \frac{e_y}{y}\right)$$

Proof

$$e_{\frac{x}{y}} = \frac{x}{y} - \frac{\hat{x}}{\hat{y}}$$

$$\frac{\hat{x}}{\hat{y}} = \frac{x-e_x}{y-e_y} = \frac{x-e_x}{y\left(1-\frac{e_y}{y}\right)} = \frac{x-e_x}{y} \times \frac{1}{1-\frac{e_y}{y}} = \frac{x-e_x}{y\left(1-\frac{e_y}{y}\right)} = \frac{x-e_x}{y} \times \frac{1}{\frac{y-e_y}{y}}$$

$$\frac{\hat{x}}{\hat{y}} = \frac{x-e_x}{y}\left(1+\frac{e_y}{y}+\left(\frac{e_y}{y}\right)^2+\ldots\ldots\right)$$

$$\frac{\hat{x}}{\hat{y}} = \frac{x-e_x}{y}\left(1+\frac{e_y}{y}\right)$$

$$\frac{\hat{x}}{\hat{y}} = \left(\frac{x}{y}-\frac{e_x}{y}\right)\left(1+\frac{e_y}{y}\right)$$

$$\frac{\hat{x}}{\hat{y}} = \frac{x}{y}+\frac{xe_y}{y^2}-\frac{e_x}{y}-\frac{e_xe_y}{y^2}, \quad \frac{e_xe_y}{y^2}=0$$

$$e_{\frac{x}{y}} = \frac{x}{y}-\left(\frac{x}{y}+\frac{xe_y}{y^2}-\frac{e_x}{y}\right)$$

$$e_{\frac{x}{y}} = \frac{x}{y}-\frac{x}{y}-\frac{xe_y}{y^2}+\frac{e_x}{y}$$

$$e_{\frac{x}{y}} = \frac{e_x}{y}-\frac{xe_y}{y^2}$$

$$e_{\frac{x}{y}} = \frac{x}{y}\left(\frac{e_x}{x}-\frac{e_y}{y}\right)$$

The relative error of the division operation is equal to

$$R_{\frac{x}{y}} = R_x - R_y$$