UTM
UNIVERSITI TEKNOLOGI MALAYSIA

CHAPTER 5

Introduction to Numerical Methods
(Error & Accuracy)

Introduction

- Computational Mathematics has been used since more than 3700 years ago by Babylon tribe to get solution of quadratic equation and square root of n integer.
- Computational Mathematics involves calculation of numerical solution for problems that can be expressed in mathematical form.
- This kind of problems that involved various fields usually can only be solved using **numerical method**.
- Nevertheless, Computational Mathematics involves more complex calculation to get more accurate answer compared to original mathematical method that use easier method, and also give more exact solution.
- Problem solving for original mathematical is performed analytically and limit only to certain problems only.

Introduction

Problem Solving in Computational Mathematics

Problems in mathematic such as complex integration, calculation certain value that is not provided in the data series in a table and linear systems require a solution using **numerical method**.

Introduction

Example 1

Integration of $\int_1^3 2x^3 dx$ is easy to evaluate but

integration of $\int_1^3 \frac{2x^3}{1+\sin x} dx$ is difficult to

evaluate and require numerical method.

Introduction

Example 2

An experiment is conducted to determine velocity of an object in a certain interval. The experimental results is given in Table 1.

Table 1

Time, t (s)	0.0	1.0	2.0	3.0
Velocity, v (m/s)	0	0.5	1.0	1.5

To find the velocity at $t = 2.7$ is difficult because the data not shown in the table.

Therefore, numerical method is required to solve the problem.

Introduction

Example 3

- The following linear system can be solved using linear system solution technique.

$$
\begin{aligned}
2x + y &= 4 \\
4x + 2y + 2z &= 16 \\
2x + 3y + z &= 4
\end{aligned}
$$

- Most mathematical problem in real world involves more that three equation. Number of equation to be solved may reach more than hundred thousand. The best solution to solve the problem is using numerical method with the assistance of computer.

- Numerical method is a field that involved applied mathematical in representing and solving mathematical problem. This problem use step of works that involve operation of numbers. It is important to know the exact mathematical solution. High accuracy of answer can be obtained through advanced calculation.

Accuracy

In numerical method, if the correct calculation method is chosen to solve the problem, the **accuracy** of answer is affected by two factors namely mistake and error.

- **Mistake**

Two source of mistake are from human carelessness and technical error by machine or computer.

- **Error**

Error occurs due to approximation process. The error is because of limitation in the power of human and machine. Error by human can be overcome but not the error by machine.

Stability & Convergence

- **STABILITY AND CONVERGENCE**
  - STABILITY in numerical analysis refers to the trend of error change iterative scheme. It is related to the concept of convergence.

    It is stable if initial errors or small errors at any time remain small when iteration progresses. It is unstable if initial errors or small errors at any time get larger and larger, or eventually get unbounded.
  - CONVERGENCE: There are two different meanings of convergence in numerical analysis:

  a. If the discretized interval is getting finer and finer after dicretizing the continuous problems, the solution is convergent to the true solution.

  b. For an iterative scheme, convergence means the iteration will get closer to the true solution when it progresses.

Numerical Errors

- **NUMERICAL ERRORS**

  When we get into the real world from an ideal world and finite to infinite, errors arise.

  - SOURCES OF ERRORS:
    - Mathematical problems involving quantities of infinite precision.
    - Numerical methods bridge the precision gap by putting errors under firm control.
    - Computer can only handle quantities of finite precision.

Types of Errors

– TYPES OF ERRORS:

- Truncation error (finite speed and time) - An example:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!}\right) + \sum_{n=4}^{\infty} \frac{x^n}{n!}$$

$$= p_3(x) + \sum_{n=4}^{\infty} \frac{x^n}{n!}$$

- Round-off error (finite word length): All computing devices represent numbers with some imprecision, except for integers.

- Human errors: (a) Mathematical equation/model. (b) Computing tools/machines. (c) Error in original data. (d) Propagated error.

Types of Errors

Data Error

Data error is because numbers used in the experiment or calculation is measured at certain accuracy or decimal points. For example, if the volume of a sphere is 6.24 ml3 , so the result contains error. This is because volume of sphere is true only at 2 decimal points.

Error in Method/Rounding

Numerical method involves many infinite series in calculation to derive the approximation of the exact value. Only part of the expression in the infinite series is utilised and this lead to the error in method.

Rounding Error

A floating number that contains many decimal points or infinite decimal points can be rounded to certain decimal point but it involves error. Rounding error is the difference between the rounded value and actual value. For example, if 1.67777777….. is rounded to 4 decimal points will give value 1.6778.

Measure of Errors

– MEASURE OF ERRORS:

Let α be a scalar to be computed and let $\bar{\alpha}$ be its approximation.

Then, we define

- Absolute error = | true value – approximated value |.

$$\varepsilon = \left| \alpha - \bar{\alpha} \right|$$

- Relative error = $\left| \dfrac{\text{true value} - \text{approximated value}}{\text{true value}} \right|$

$$\varepsilon_r = \left| \frac{\alpha - \bar{\alpha}}{\alpha} \right|$$

Error, Modulus Error and Relative Error

Let, **N** is an exact value and **n** is an **approximate value**. Error is a difference between exact value and approximate value and is represented as ε (error):

$$\varepsilon = N - n$$

Sign for error value is not important. **Modulus** is used to show that error is always positive. This value is called modulus error or **absolute error**, $|\varepsilon|$.

$|\varepsilon|$ = | exact value – approximate value |

$|\varepsilon| = | N - n |$

Error, Modulus Error and Relative Error

Ratio for error is called **relative error** and is expressed as:

$$\varepsilon_{rel} = \frac{|\text{exact value} - \text{approximate value}|}{|\text{exact value}|}$$, or $$\varepsilon_{rel} = \frac{|N-n|}{|N|}$$

Because of its exact nature, in general N is known, therefore

$$\varepsilon_{rel} = \frac{|N-n|}{|n|}$$

For calculation purpose, relative error will give more accurate result compared to the error itself.

Error, Modulus Error and Relative Error

Example 4

Let the true value of π be **3.1415926535898** and its approximation be **3.14** as usual. Compute the absolute error and relative error of such an approximation.

Solution:

The absolute error:

$$\varepsilon = \left|\pi - \bar{\pi}\right| = \left|3.1415926535898 - 3.14\right| = 0.0015926535898$$

which implies that the approximation is accurate up to 2 decimal places.

The relative error:

$$\varepsilon_r = \left|\frac{\pi - \bar{\pi}}{\pi}\right| = \frac{0.0015926535898}{3.1415926535898} = 0.000506957382897$$

which implies that the approximation has a accuracy of 3 significant figures.

Error, Modulus Error and Relative Error

Example 5

Find the error value and relative error for the following:

a) N = 3.141592 and n = 3.14

|e| = | 3.141592 – 3.14 | = 0.001592

$$\varepsilon_{rel} = \frac{|3.141592 - 3.14|}{|3.141592|} = 0.0057$$

b) N =1 000 000 and n=999 996

|e| = | 1 000 000 – 999 996 | = 4 ; $\varepsilon_{rel} = \dfrac{|1\,000\,000 - 999\,996|}{|1\,000\,000|} = 0.000004$

c) N = 0.000012 and n = 0.000009

|e| = | 0.000012 – 0.000009 | = 0.000003 ; $\varepsilon_{rel} = \dfrac{|0.000012 - 0.000009|}{|0.000012|} = 0.25$

Error, Modulus Error and Relative Error

Example 5 (cont'd)

- Example 5(a), no big difference between error value and relative error. Therefore, either error value or relative error can be used to determine the accuracy of approximate value, n.
- Example 5(b), N value is in magnitude 10^6 and the resulted error is big but the relative error is small. This show that approximate value(n) may be good approximate to the exact value, N.
- Example 5(c), with magnitude 10^{-6}, resulted error is the smallest between among three examples but the resulted relative error is the biggest among three examples. Percentage of relative error in 5(c) is 25% and this show that approximate value (n) is not good approximate for N.

We can see that when $|N|$ move far from 1 (bigger or smaller), value of relative error is better approximate to the accuracy of approximate value.

Maximum Modulus for Rounding Error

If a number is rounded to *n* decimal points, so its error satisfies the inequality:

$$\left|\varepsilon_{rel}\right| \le \frac{1}{2} \times \frac{1}{10^n} \quad \text{, or} \quad \left|\varepsilon_{rel}\right| \le \frac{1}{2} \times 10^{-n}$$

Maximum Modulus for Rounding Error

Example 6

Given a number rounded to 2 decimal points, the relative error is:

$$|\varepsilon_{rel}| \leq \frac{1}{2} \times 10^{-2} \quad = \quad |\varepsilon_{rel}| \leq 0.005$$

Effect of Rounding Error in Addition and Subtraction

Given n_1, n_2 are approximate value of two quantities, N_1, N_2. Error are ε_1, ε_2 where $\varepsilon_1 = N_1 - n_1$ and $\varepsilon_2 = N_2 - n_2$

If N_3 is total for exact value, so

$$N_3 = N_1 + N_2$$

Total of approximate value, is given by

$$n_3 = n_1 + n_2$$

and the error is

$$\begin{aligned} \varepsilon_3 &= N_3 - n_3 \\ &= (N_1 + N_2) - (n_1 + n_2) \\ &= (N_1 - n_1) + (N_2 - n_2) \\ &= \varepsilon_1 + \varepsilon_2 \end{aligned}$$

Effect of Rounding Error in Addition and Subtraction

For any two numbers, a and b, the inequality is given by

$$|a \pm b| \leq |a| + |b|$$

Then, we obtain

$$|e_3| \leq |e_1| + |e_2|$$

If e_3 is exact difference between error e_1, e_2, so we can show that

$$e_3 = e_1 - e_2$$

We can conclude that modulus error from addition or subtraction of two values is less or equal with total modulus for each value.

Effect of Rounding Error in Addition and Subtraction

Example 7

Find the most accurate value of the following operation:

$$3.69 + 5.432 - 2.37 - 3.5214$$

where each number has been rounded to their decimal points.

Solution:

Exact calculation gives result as follow:

$$3.69 + 5.432 - 2.37 - 3.5214 = 3.2306$$

Modulus error is given by,

$$|\varepsilon_3| \leq 0.005 + 0.0005 + 0.005 + 0.00005 = 0.01055$$

Therefore, the result is between

$3.2306 + 0.01055$ and $3.2306 - 0.01055$

3.24115 and 3.22005

Both of these values are same if rounded to 1 decimal point. Therefore, the best approximate value is 3.2 (1 decimal point).

Exercise #1

The volatge in a high-voltage transmission line is stated to be 2.4 MV while the actual voltage may range from 2.1 MV to 2.7 MV. What is the maximum absolute and relative error of voltage?

Exercise #2

Find the most accurate value for the following operation, where each number has been rounded to a certain decimal point,

$$3.314 + 2.0008 + 4.15 - 1.07$$

Exercise #3

Given approximate values for $\bar{x} = 6.171$ and $\bar{y} = 3.70$ Estimate the value $x + y$, given that the exact value, $x = 6.171256$ and $y = 3.70123$. Then, based on this addition operation, find the following:

i) Error

ii) Absolute error

iii) Relative error