6. APPROXIMATION AND ERRORS

In the practice of numerical method it is important to be aware that computed solutions are not exact mathematical solutions. Perfect accuracy in most computational processes is impossible. We must make certain approximations, and this introduced **errors**.

The **error** in a computed quantity is defined as

Error = true value – approximate value

$= x_T - x_A$

The **relative error** is a measure of the error in relation to the size of the true being sought:

$$\text{Relative error} = \left|\frac{\text{error}}{\text{true value}}\right|$$

$$= \left|\frac{x_T - x_A}{x_T}\right|$$

$$\text{True percent relative error}, \varepsilon_t = \frac{\text{true value} - \text{approximat e value}}{\text{true value}} \times 100\%$$

$$\text{Approximat e percent}, \varepsilon_a = \frac{\text{present approximat ion} - \text{previous approximat ion}}{\text{present approximat ion}} \times 100\%$$

relative error

The **significant figures** of a number are those that can be used with confidence. They correspond to the number of certain digits plus one estimate digit.

Example 1:
How many significant figues in the following numbers?

a.	0.3	1 significant figure
b.	0.03	1 significant figure
c.	0.030	2 significant figures
d.	300	may be one, two or three significant figures, depending on whether the zeros are known with confidence.
e.	3×10^2	1 significant figure
f.	3.0×10^2	2 significant figures
g.	3.00×10^2	3 significant figures

The concept of significant figures will have relevance to the definition of accuracy and precision:

Precision refers to how closely individual measured or computed values agree with each other. Precision is governed by the number of digits being carried in the numerical calculations

Accuracy refers to how closely a number agrees with the true value of the number it is representing. Accuracy is governed by the errors in the numerical approximation.

6.1 Approximation and Round-off errors

Round-off error originate from the fact that computers retain only a finite number of decimal places during a calculation. Therefore the results of its arithmetic operations are only approximations to the true results. For example in Matlab, all numbers are rounded to the nearest eighth decimalI in single precision) or to the nearest sixteenth decimal (in double precision). In addition, rounding errors become important when the step size h is comparable with the precision of the computations. This is because computers use a base-2 representation, they can not precisely represent certain exact base-10 numbers.The discrepancy introduced by this omission of significant figures is called **round-off error**.

Example 2:

Numbers such as π, e, or $\sqrt{3}$ can not be expressed by a fixed number of decimal places. Therefore they can not be represented exactly by the computer.

Consider the number π. It is irrational, i.e. it has infinitely many digits after the period:

π = 3.1415926535897932384626433832795.....

The round-off error computer representation of the number π depends on how many digits are left out.

Number of digits	Approximation for π	Absolute error	Percent Relative errror
1	3.100	0.041593	1.3239%
2	3.140	0.001593	0.0507%
3	3.142	0.000407	0.0130%

6.2 Taylor Series and Truncation Errors

Definition- Truncation errors are those that result from using an approximation in place of an exact mathematical procedure.

Approximation of the derivative

$$\frac{df}{dt} \cong \frac{\Delta f}{\Delta t} = \frac{f(t_{i+1}) - f(t_i)}{t_{i+1} - t_i}$$

One of the most important methods used in numerical methods to approximate mathematical functions is: Taylor series

The Taylor series provides a means to predict a function at one point in terms of the function value and its derivative at another point.

Therefore, any smooth function can be approximated as a polynomial.

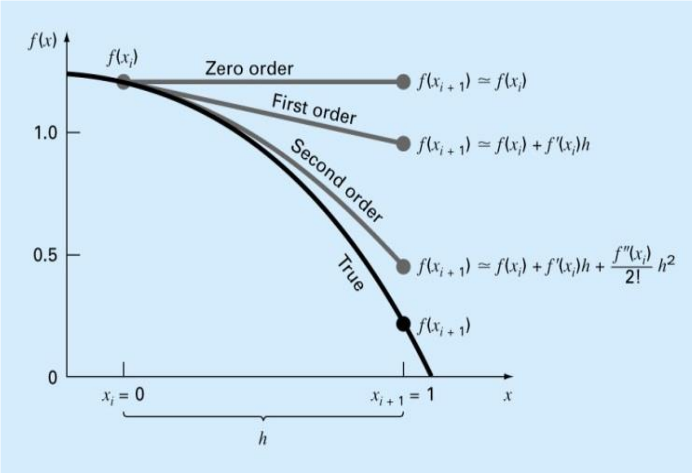

Figure 1: Taylor series

The expression of Taylor series for any function is:

$$f(x_{i+1}) = f(x_i) + f'(x_i)(x_{i+1} - x_i) + \frac{f''(x_i)}{2!}(x_{i+1} - x_i)^2 + \frac{f'''(x_i)}{3!}(x_{i+1} - x_i)^3 + \ldots + \frac{f^n(x_i)}{n!}(x_{i+1} - x_i)^n + R_n$$

A remainder term is included to account for high order terms neglected (from (n+1) to infinity):

$$R_n = \frac{f^{n+1}(\xi)}{(n+1)!}(x_{i+1} - x_i)^{n+1}$$

This represents the remainder for the n^{th}-order approximation. ξ is a value that lies somewhere between x_i and x_{i+1}.

It is very convenient to write the Taylor series under a compact form by replacing (x_{i+1}- x_i) by the step (h):

$$f(x_{i+1}) = f(x_i) + f'(x_i)h + \frac{f''(x_i)}{2!}h^2 + \frac{f'''(x_i)}{3!}h^3 + ... + \frac{f^n(x_i)}{n!}h^n + R_n$$

and

$$R_n = \frac{f^{n+1}(\xi)}{(n+1)!}h^{n+1}$$

Although, theoretically an infinite number of terms are needed to yield to an exact estimate, in practice only few terms are sufficient.

The number of terms needed is dependent on the application, the precision needed and it is determined using the remainder term of the expansion.

However, the determination of the remainder term R_n is not straightforward, since we have:

To know ξ

To differentiate the function f (n+1) times, and for that we need to know the function f !

The only term that we can control in this expression is (h). Therefore, it is very convenient to express R_n as:

$$R_n = 0(h^{n+1})$$

Which must be interpreted as a truncation error of order n+1. This means that the error is proportional to h^{n+1}.

Therefore, if h is sufficiently small, an accurate estimate for $f(x_{i+1})$ can be reached using only few terms.

The use of Taylor series exists in so many aspects of numerical methods ns. For example, you must have come across expressions such as

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \qquad (1)$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \qquad (2)$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \qquad (3)$$

All the above expressions are actually a special case of Taylor series called the Maclaurin series. Why are these applications of Taylor's theorem important for numerical methods? Expressions such as given in Equations (1), (2) and (3) give you a way to find the approximate values of these functions by using the basic arithmetic operations of addition, subtraction, division, and multiplication.

Example 3

Find the value of $e^{0.25}$ using the first five terms of the Maclaurin series.

Solution

The first five terms of the Maclaurin series for e^x is

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}$$

$$e^{0.25} \approx 1 + 0.25 + \frac{0.25^2}{2!} + \frac{0.25^3}{3!} + \frac{0.25^4}{4!} = 1.2840$$

The exact value of $e^{0.25}$ up to 5 significant digits is also 1.2840.

But the above discussion and example do not answer our question of what a Taylor series is.

Here it is, for a function $f(x)$

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \cdots \qquad (4)$$

provided all derivatives of $f(x)$ exist and are continuous between x and $x+h$.

Example 4:

Take $f(x) = \sin(x)$, we all know the value of $\sin\left(\frac{\pi}{2}\right) = 1$. We also know the $f'(x) = \cos(x)$ and $\cos\left(\frac{\pi}{2}\right) = 0$. Similarly $f''(x) = -\sin(x)$ and $\sin\left(\frac{\pi}{2}\right) = 1$. In a way, we know the value of $\sin(x)$ and all its derivatives at $x = \frac{\pi}{2}$. We do not need to use any calculators, just plain differential calculus and trigonometry would do. Can you use Taylor series and this information to find the value of $\sin(2)$?

Solution

$$x = \frac{\pi}{2}\text{, so } x + h = 2$$

$$\begin{aligned} h &= 2 - x \\ &= 2 - \frac{\pi}{2} = 0.42920 \end{aligned}$$

So

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + f'''(x)\frac{h^3}{3!} + f''''(x)\frac{h^4}{4!} + \cdots$$

$$x = \frac{\pi}{2} \therefore h = 0.42920$$

$$f(x) = \sin(x),\ f\left(\frac{\pi}{2}\right) = \sin\left(\frac{\pi}{2}\right) = 1$$

$$f'(x) = \cos(x),\ f'\left(\frac{\pi}{2}\right) = 0$$

$$f''(x) = -\sin(x),\ f''\left(\frac{\pi}{2}\right) = -1$$

$$f'''(x) = -\cos(x),\ f'''\left(\frac{\pi}{2}\right) = 0$$

$$f''''(x) = \sin(x),\ f''''\left(\frac{\pi}{2}\right) = 1$$

Hence

$$f\left(\frac{\pi}{2}+h\right) = f\left(\frac{\pi}{2}\right) + f'\left(\frac{\pi}{2}\right)h + f''\left(\frac{\pi}{2}\right)\frac{h^2}{2!} + f'''\left(\frac{\pi}{2}\right)\frac{h^3}{3!} + f''''\left(\frac{\pi}{2}\right)\frac{h^4}{4!} + \cdots$$

$$f\left(\frac{\pi}{2}+0.42920\right) = 1 + 0(0.42920) - 1\frac{(0.42920)^2}{2!} + 0\frac{(0.42920)^3}{3!} + 1\frac{(0.42920)^4}{4!} + \cdots$$

$$\begin{aligned} &= 1 + 0 - 0.092106 + 0 + 0.00141393 + \cdots \\ &\cong 0.90931 \end{aligned}$$

6.2.1 Using Taylor series to estimate truncation errors

As you have noticed, the Taylor series has infinite terms. Only in special cases such as a finite polynomial does it have a finite number of terms. So whenever you are using a Taylor series to calculate the value of a function, it is being calculated approximately.

The Taylor polynomial of order n of a function $f(x)$ with $(n+1)$ continuous derivatives in the domain $[x, x+h]$ is given by

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + \cdots + f^{(n)}(x)\frac{h^n}{n!} + R_n(x)$$

where the remainder is given by

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c).$$

where

$$x < c < x+h$$

that is, c is some point in the domain $(x, x+h)$.

Example 5:

The Taylor series for e^x at point $x = 0$ is given by

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

a) What is the truncation (true) error in the representation of e^1 if only four terms of the series are used?

b) Use the remainder theorem to find the bounds of the truncation error.

Solution

a) If only four terms of the series are used, then

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$$

$$e^1 \approx 1 + 1 + \frac{1^2}{2!} + \frac{1^3}{3!}$$

$$= 2.66667$$

The truncation (true) error would be the unused terms of the Taylor series, which then are

$$\begin{aligned} E_t &= \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots \\ &= \frac{1^4}{4!} + \frac{1^5}{5!} + \cdots \\ &\cong 0.0516152 \end{aligned}$$

b) But is there any way to know the bounds of this error other than calculating it directly? Yes,

$$f(x+h) = f(x) + f'(x)h + \cdots + f^{(n)}(x)\frac{h^n}{n!} + R_n(x)$$

where

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c),\ x < c < x+h \text{, and}$$

c is some point in the domain $(x, x+h)$. So in this case, if we are using four terms of the Taylor series, the remainder is given by $(x = 0, n = 3)$

$$\begin{aligned} R_3(x) &= \frac{(0-1)^{3+1}}{(3+1)!} f^{(3+1)}(c) \\ &= \frac{1}{4!} f^{(4)}(c) \\ &= \frac{e^c}{24} \end{aligned}$$

Since

$$\begin{aligned} &x < c < x+h \\ &0 < c < 0+1 \\ &0 < c < 1 \end{aligned}$$

The error is bound between

$$\begin{aligned} &\frac{e^0}{24} < R_3(1) < \frac{e^1}{24} \\ &\frac{1}{24} < R_3(1) < \frac{e}{24} \\ &0.041667 < R_3(1) < 0.113261 \end{aligned}$$

So the bound of the error is less than 0.113261 which does concur with the calculated error of 0.0516152 .

6.3 Total numerical error

The total error is the sum of the truncation and round-off errors, however:

Round-off ↓	↑ the number of figures
Round-off ↑	With subtractive cancellation and the increase in the number of computations.
Truncation errors ↓	↓ h [step], however this leads to subtractive cancellation or to the increase in computations.

Therefore, we are facing a dilemma: decreasing one component of the total errors increases the other.
However, with actual computers, the round-off errors can be minimized and therefore we will be able to decrease the truncation error by reducing (h).

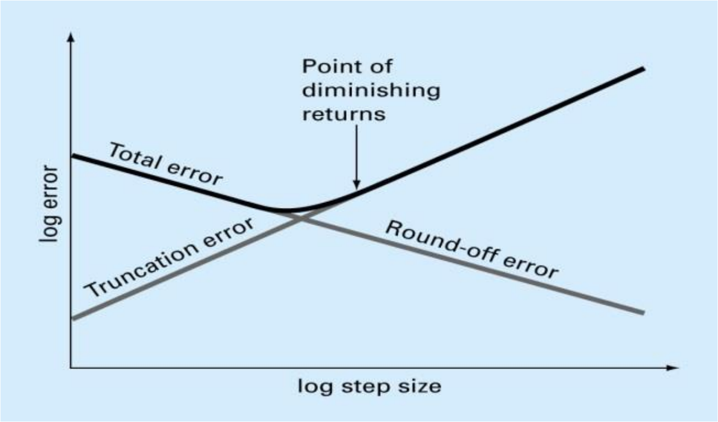

Figure 3: Variation of total error as a function of the step size.

<u>Some advice to control numerical errors</u>

i) Avoid subtracting two nearly equal numbers.
ii) Avoid subtractive cancellation by reformulating your problem.
iii) When you add or subtract numbers sort them and start with the smallest one.

Exersice 6:

1. Perform the following computations exactly
   a) using three-digit chopping arithmetic
   b) using three-digit rounding arithmetic
   c) compute the absolute error in part ii and iii
   d) compute the relative error in part ii and iii for
      i) 133 +0.921
      ii) $\left(\frac{2}{9}\right)\cdot\left(\frac{9}{7}\right)$

2. Let

$$f(x)=\frac{x\cos x-\sin x}{x-\sin x}$$

Use four-digit rounding arithmetic to evaluate $f(0.1)$

Replace each trigonometric function with its Maclaurin polynomial, order $n = 2$, and repeat part (a)

The actual value is $f(0.1)= -1.9989998$. Find the relative error for the values obtained in parts (a) and (b).

3. The Taylor series for e^x at point $x=0$ is given by

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\cdots$$

As you can see in the previous example that by taking more terms, the error bounds decrease and hence you have a better estimate of e^1. How many terms it would require to get an approximation of e^1 within a magnitude of true error of less than 10^{-6} ?