Math 4543: Numerical Methods

Lecture 2 — Measuring Errors & Sources of Error

Syed Rifat Raiyan
Lecturer
Department of Computer Science & Engineering
Islamic University of Technology, Dhaka, Bangladesh
Email: rifatraiyan@iut-dhaka.edu

Lecture Plan

The agenda for today

- Different types of errors
- How to measure them?
- Discuss about their interpretability
- Concept of Significant Digits
- Relationship between Relative Approximate Error and Significant Digits
- Different sources of errors

Errors

How do we deal with them

In any numerical analysis, errors will arise during the calculations.

To be able to deal with the issue of errors, we —

- **Identify** where the error is coming from
- **Quantify** the error
- **Minimize** the error as per our needs.

Let's focus on *identifying* and *quantifying* the errors in this lecture.

True Error

What is True Error?

True error denoted by E_t is the difference between the true value (also called the exact value) and the approximate value.

$$E_t = \text{True value} - \text{Approximate value}$$

The ***True value*** is also referred to as the ***Exact value*** or the ***Ground Truth value***.

True Error

An example

The derivative of a function $f(x)$ at a particular value of x can be approximately calculated by

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

For $f(x) = 7e^{0.5x}$ and $h = 0.3$, find

a) the approximate value of $f'(2)$

b) the true value of $f'(2)$

c) the true error for part (a)

True Error

An example

Solution

a)

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

For $x = 2$ and $h = 0.3$,

$$\begin{aligned} f'(2) &\approx \frac{f(2+0.3) - f(2)}{0.3} \\ &= \frac{f(2.3) - f(2)}{0.3} \\ &= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3} \\ &= \frac{22.107 - 19.028}{0.3} \\ &= 10.263 \end{aligned}$$

True Error

An example

b) The exact value of $f'(2)$ can be calculated by using our knowledge of differential calculus.

$$f(x) = 7e^{0.5x}$$

$$\begin{aligned} f'(x) &= 7 \times 0.5 \times e^{0.5x} \\ &= 3.5e^{0.5x} \end{aligned}$$

So the true value of $f'(2)$ is

$$\begin{aligned} f'(2) &= 3.5e^{0.5(2)} \\ &= 9.5140 \end{aligned}$$

c) True error is calculated as

$$\begin{aligned} E_t &= \text{True value} - \text{Approximate value} \\ &= 9.5140 - 10.263 \\ &= -0.749 \end{aligned}$$

True Error

But is it properly interpretable?

The magnitude of true error does not show how bad the error is. A true error of $E_t = -0.749$ may seem to be small, but if the function given in Example 1 were $f(x) = 7 \times 10^{-6} e^{0.5x}$, the true error in calculating $f'(2)$ with $h = 0.3$, would be $E_t = -0.749 \times 10^{-6}$. This value of true error is smaller, even when the two problems are similar in that they use the same value of the function argument, $x = 2$, and the step size, $h = 0.3$. This recognition brings us to the definition of the relative true error.

A low True Error ***doesn't*** necessarily translate to a good approximation.

We need some **scale invariance** for proper interpretation!

Relative True Error

What is Relative True Error?

Relative true error is denoted by ϵ_t and is defined as the ratio between the true error and the true value.

$$\epsilon_t = \frac{\text{True Error}}{\text{True Value}}$$

Now, we can interpret it as — what *portion* of the True value is different from the Approximate value?

Relative True Error

Working with the same example

The derivative of a function $f(x)$ at a particular value of x can be approximately calculated by $f'(x) \approx \dfrac{f(x+h) - f(x)}{h}$

For $f(x) = 7e^{0.5x}$ and $h = 0.3$, find the relative true error in finding $f'(2)$.

Relative true error is calculated as

$$\begin{aligned}\epsilon_t &= \frac{\text{True Error}}{\text{True Value}} \\ &= \frac{-0.749}{9.5140} \\ &= -0.078726\end{aligned}$$

Relative true errors are also presented as percentages. For this example,

$$\begin{aligned}\epsilon_t &= -0.078726 \times 100\% \\ &= -7.8726\%\end{aligned}$$

Absolute relative true errors may also need to be calculated. In such cases,

$$\begin{aligned}|\epsilon_t| &= |-0.078726| \\ &= 0.078726 \\ &= 7.8726\%\end{aligned}$$

Now, try to find out the Relative True Error for $f(x) = 7 \times 10^{-6} e^{0.5x}$ using the same parameters $x = 2$ and the step size, $h = 0.3$.

Spoiler: It should also be -0.078726

Now we can interpret the error properly!

Approximate Error

What is Approximate Error?

Suppose, we **don't have** the luxury of knowing **the true value**. When we are solving a problem numerically, we will **only** have access to **approximate values**.

We need to know how to quantify error for such cases, and such an error is appropriately called the ***approximate error***.

Approximate error is denoted by E_a, and is defined as the difference between the present approximation and the previous approximation.

$$E_a = \text{Present Approximation} - \text{Previous Approximation}$$

Approximate Error

An example

The derivative of a function $f(x)$ at a particular value of x can be approximately calculated by

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

For $f(x) = 7e^{0.5x}$ and at $x = 2$, find the following

a) $f'(2)$ using $h = 0.3$

b) $f'(2)$ using $h = 0.15$

c) approximate error for the value of $f'(2)$ for part (b)

Approximate Error

An example

Solution

The approximate expression for the derivative of a function is

$$f'(x) \approx \frac{f(x+h)-f(x)}{h}$$

a) For $x = 2$ and $h = 0.3$,

$$\begin{aligned} f'(2) &\approx \frac{f(2+0.3)-f(2)}{0.3} \\ &= \frac{f(2.3)-f(2)}{0.3} \\ &= \frac{7e^{0.5(2.3)}-7e^{0.5(2)}}{0.3} \\ &= \frac{22.107-19.028}{0.3} \\ &= 10.263 \end{aligned}$$

b) Repeating the procedure of part (a) with $h = 0.15$

$$f'(x) \approx \frac{f(x+h)-f(x)}{h}$$

For $x = 2$ and $h = 0.15$,

$$\begin{aligned} f'(2) &\approx \frac{f(2+0.15)-f(2)}{0.15} \\ &= \frac{f(2.15)-f(2)}{0.15} \\ &= \frac{7e^{0.5(2.15)}-7e^{0.5(2)}}{0.15} \\ &= \frac{20.510-19.028}{0.15} \\ &= 9.8800 \end{aligned}$$

c) So the approximate error, E_a is

$$\begin{aligned} E_a &= \text{Present Approximation} - \text{Previous Approximation} \\ &= 9.8800 - 10.263 \\ &= -0.38300 \end{aligned}$$

Approximate Error

Facing the same interpretability problem?

The magnitude of approximate error does not reflect how bad the error is. An approximate error of $E_a = -0.38300$ in the above example may seem to be small, but repeating the above example for $f(x) = 7 \times 10^{-6} e^{0.5x}$, the approximate error in calculating $f'(2)$ with $h = 0.15$ would be $E_a = -0.38300 \times 10^{-6}$. This value of approximate error is much smaller, even when the two problems are similar in that they use the same value of the function argument, $x = 2$, and step sizes of $h = 0.15$ and $h = 0.3$. This recognition brings us to the definition of the relative approximate error in the next lesson.

A low Approximate Error ***doesn't*** necessarily translate to a good approximation.

We need some **scale invariance** for proper interpretation in this case as well!

Relative Approximate Error

What is Relative Approximate Error?

As observed in the previous lesson, approximate errors generally do not reflect how bad or acceptable an error is. We hence define the relative approximate error, denoted by ϵ_a as the ratio between the approximate error and the present approximation.

$$\epsilon_a = \frac{\text{Approximate Error}}{\text{Present Approximation}}$$

where

$$\text{Approximate Error}, E_a = \text{Current Approximation} - \text{Previous Approximation}$$

Now, we can interpret it as — what *portion* of the new approx. value is different from the previous approx. value?

Relative Approximate Error

Working with the same example

The derivative of a function $f(x)$ at a particular value of x can be approximately calculated by

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

For $f(x) = 7e^{0.5x}$, find the relative approximate error in calculating $f'(2)$ using values from $h = 0.3$ and $h = 0.15$.

Solution

From Example 1, the following values are already calculated

$$f'(2) \approx 10.263 \text{ using } h = 0.3 \text{ and}$$

$$f'(2) \approx 9.8800 \text{ using } h = 0.15.$$

Hence the approximate error

$$\begin{aligned} E_a &= \text{Present Approximation} - \text{Previous Approximation} \\ &= 9.8800 - 10.263 \\ &= -0.38300 \end{aligned}$$

The relative approximate error is calculated as

$$\begin{aligned} \epsilon_a &= \frac{\text{Approximate Error}}{\text{Present Approximation}} \\ &= \frac{-0.38300}{9.8800} \\ &= -0.038765 \end{aligned}$$

Relative approximate errors are also can be presented as percentages. For this example,

$$\begin{aligned} \epsilon_a &= -0.038765 \times 100\% \\ &= -3.8765\% \end{aligned}$$

Absolute relative approximate errors may also need to be calculated. In this example

$$\begin{aligned} |\epsilon_a| &= |-0.038765| \\ &= 0.038765 = 3.8765\% \end{aligned}$$

Now, try to find out the Relative Approx. Error for $f(x) = 7 \times 10^{-6} e^{0.5x}$ using the same step sizes, $h = 0.3$ and $h = 0.15$.

Significant Digits

What are Significant Digits?

Significant figures (also known as the significant digits, precision or resolution) of a number in positional notation are digits in the number that are **reliable** and **necessary** to indicate the quantity of something. These are the digits that are known reliably in a reported number.

Thermal Expansion Coefficient of Steel = $6.47\ \mu\text{in}/\text{in}/^\circ\text{F}$

$= 0.00000647\ \text{in}/\text{in}/^\circ\text{F}$

If we were writing the value up to 2 decimal places, it would be written as $0.00\ \text{in}/\text{in}/^\circ\text{F}$.

Hence, the emphasis in engineering is to write the numbers mostly in the *scientific format* as it represents the numbers with the proper number of significant digits.

Significant Digits

The rules that govern Significant Digits

- Non-zero digits are **always significant**.
- Zeros between non-zero digits are **always significant**.
- Leading zeros are **never significant**.
- Trailing zeros are **only significant** if the number contains a decimal point (or if willingly considered to be significant).

2.789 has four significant digits as all nonzero digits are significant.

0.0439 has three significant digits as zeros to the left of the first nonzero number are not significant.

4.590 has four significant digits as all zeros to the right of a decimal point are significant.

4008 has four significant digits as any zero between nonzero digits is significant.

4208.07 has six significant digits as any zero between nonzero digits is significant.

4008.0 has five significant digits as all zeros to the right of a decimal point are significant.

4000.0 has five significant digits as all zeros to the right of a decimal point are significant.

Significant Digits

The rules that govern Significant Digits

15000 may represent 2, 3, 4, or 5 significant digits. Such vagueness can be addressed by using the scientific (also called floating point) format as given below.

1.5×10^4 has two significant digits.

1.50×10^4 has three significant digits.

1.5000×10^4 has five significant digits.

An exact number has an infinite number of significant digits. If you have 5 books in your school bag, then the number can be written with an infinite number of trailing zeros as $\underbrace{5.000\ldots}_{\text{Arbitrary precision}}$

Mini Quiz

Practise identifying Significant Digits

Count the ***minimum*** number of significant digits and say which digits are significant.

81

26.2

0.007

5200.38

380.0

78800

78800.

Mini Quiz

Practise identifying Significant Digits (solutions)

Count the *minimum* number of significant digits and say which digits are significant.

81 — 2

26.2 — 3

0.007 — 1

5200.38 — 6

380.0 — 4

78800 — 3

78800. — 5

Significant Digits × | $\in_a$ |

How do they relate?

In a numerical method that uses iterative methods, a user can calculate relative approximate error ϵ_a at the end of each iteration. The user may pre-specify a minimum acceptable tolerance called the pre-specified tolerance, ϵ_s. If the absolute relative approximate error $|\epsilon_a|$ is less than or equal to the pre-specified tolerance ϵ_s, that is, $|\epsilon_a| \leq \epsilon_s$, then the acceptable error has been reached and no more iterations would be required.

Alternatively, one may pre-specify how many significant digits they would like to be correct in their answer. In that case, if one wants at least m significant digits to be correct in the answer, then you would need to have the absolute relative approximate error, $\boxed{|\epsilon_a| \leq 0.5 \times 10^{2-m}\%}$.

Significant Digits × $| \in_a |$

An example

If one chooses 6 terms of the Maclaurin series for e^x to calculate $e^{0.7}$, how many significant digits can you trust in the solution? Find your answer without knowing or using the exact answer.

Solution

The Maclaurin series for e^x is given by

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots\ldots\ldots\ldots\ldots\ldots \qquad (E1.1)$$

Using 5 terms in Equation (E1.1), we get the previous approximation as

$$e^{0.7} \cong 1 + 0.7 + \frac{0.7^2}{2!} + \frac{0.7^3}{3!} + \frac{0.7^4}{4!}$$
$$= 2.0122$$

Using 6 terms in Equation (E1.1), we get the current approximation as

$$e^{0.7} \cong 1 + 0.7 + \frac{0.7^2}{2!} + \frac{0.7^3}{3!} + \frac{0.7^4}{4!} + \frac{0.7^5}{5!}$$
$$= 2.0136$$

The percentage absolute relative approximate error is

$$|\epsilon_a| = \left| \frac{\text{Current Approximation - Previous Approximation}}{\text{Present Approximation}} \right| \times 100$$
$$= \left| \frac{2.0136 - 2.0122}{2.0136} \right| \times 100$$
$$= 0.069527\%$$

At least *m* significant digits are expected to be correct when absolute relative approximate error, $|\epsilon_a| \leq 0.5 \times 10^{2-m}\%$. Since $|\epsilon_a| \leq 0.5 \times 10^{2-2}\% = 0.5\%$, but $|\epsilon_a| \geq 0.5 \times 10^{2-3}\% = 0.05\%$, at least 2 significant digits are correct.

Significant Digits × | $\in_a$ |

An example

Alternatively, to find the least number of significant digits correct in the answer, $|\epsilon_a| \leq 0.5 \times 10^{2-2}\%$

$$0.069527 \leq 0.5 \times 10^{2-m}$$

$$\frac{0.069527}{0.5} \leq 10^{2-m}$$

$$0.1391 \leq 10^{2-m}$$

$$\log_{10}(0.1391) \leq \log_{10}(10^{2-m})$$

$$-0.8567 \leq 2 - m$$

$$m \leq 2.8567$$

Hence 2 significant digits are at least correct in the estimated value

$$e^{0.7} \approx 2.0136$$

Sources of Error

Types of errors in Numerical Analysis

Error in solving an engineering or science problem can arise due to several factors.

First, the error may be in the modeling technique. A mathematical model may be based on using assumptions that are not acceptable.

Second, errors may arise from mistakes in programs themselves or in the measurement of physical quantities.

In applications of numerical methods themselves, the two errors we focus on are inherent to the subject, namely

- **Round-off Error** — caused by approximating **number**
- **Truncation Error** — caused by approximating **mathematical processes** or **modeling technique**

Round-off Error

What is Round-off Error?

A computer can only represent a number approximately. For example, a number like $\frac{1}{3}$ may be represented as 0.333333 on a PC. Then the round-off error, in this case, is $\frac{1}{3} - 0.333333 = 0.0000003\bar{3}$. Then there are other numbers that cannot be represented exactly. For example, π and $\sqrt{2}$ are numbers that need to be approximated in computer calculations.

Computers would need an **infinite number of bits** to store the numbers in binary!

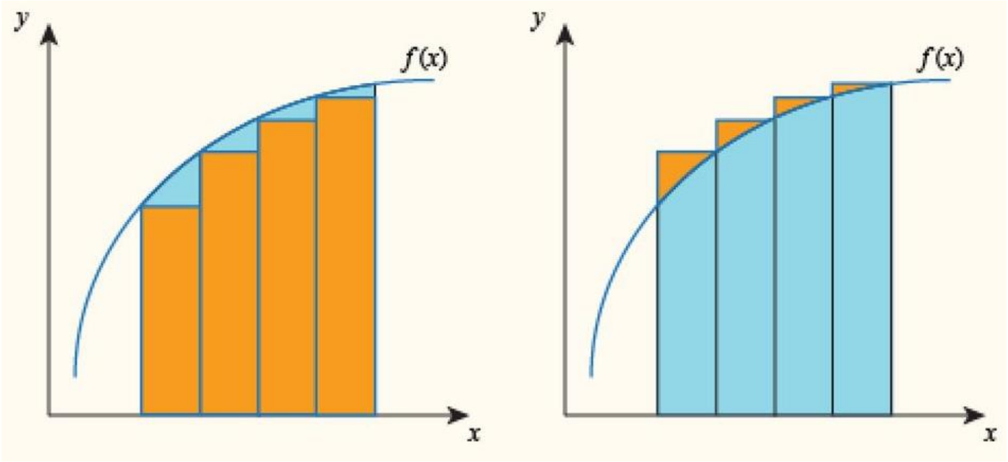
But, registers have a **finite number of bits**. So we need to ***chop*** or ***round***.

Truncation Error

What is Truncation Error?

Truncation error is defined as the error caused by approximating a mathematical procedure. Examples of mathematical procedure getting approximated include an infinite series where only the first few terms are used, integrals where the number of rectangles drawn to show the area under the curve is finite in number as opposed to infinite, and the derivative of a function is approximated by using the slope of a secant line as opposed to a tangent line.

$$e^x \approx 1 + x + \frac{x^2}{2!} \qquad f'(x) \approx \frac{f(x+h) - f(x)}{h}$$



Truncation Error

Examples (Chopping of terms in a series)

For example, the Maclaurin series for e^x is given as

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

This series has an infinite number of terms, but when using this series to calculate e^x, only a finite number of terms can be used. For example, if one uses three terms to calculate e^x, then

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

the truncation error for such an approximation for any value of x is

$$\begin{aligned} \text{Truncation error} &= e^x - \left(1 + x + \frac{x^2}{2!}\right) \\ &= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\right) - \left(1 + x + \frac{x^2}{2!}\right) \\ &= \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

Think: How to reduce this?

Truncation Error

Examples (Using a finite step size instead of an infinitesimal value)

To find the exact derivative of a function, we define

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

But since we cannot use $h \to 0$, in numerical methods, we use a finite value of h, to give

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

So the truncation error is caused by choosing a finite value of h as opposed to a $h \to 0$.

directly using the formula from your differential calculus class

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

Truncation Error

Examples (Using a finite step size instead of an infinitesimal value)

In finding $f'(3)$ for $f(x) = x^2$, we have the exact value of the derivative calculated as follows. Given

$$f(x) = x^2$$

from the definition of the derivative of a function,

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{(x+h)^2 - (x)^2}{\Delta x} \\ &= \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{\Delta x} \\ &= \lim_{h \to 0} (2x + h) \\ &= 2x \end{aligned}$$

The above derivative is the same expression you would have obtained by directly using the formula from your differential calculus class

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

By this formula for

$$f(x) = x^2$$

$$f'(x) = 2x$$

Given the step size of $h = 0.2$, we get

$$f'(3) = \frac{f(3+0.2) - f(3)}{0.2} = \frac{10.24 - 9}{0.2} = 6.2$$

$$\begin{aligned} \text{Truncation Error} &= \text{Exact Value} - \text{Approximate Value} \\ &= 6 - 6.2 \\ &= -0.2 \end{aligned}$$

Think: How to reduce this?

Mini Quiz

Theoretical vs Practical notion of infinitesimally small values

Is it a good idea to take a step size of $h = 10^{-20}$ to approximate the derivative $f'(3)$?

If so, why? If not, why not?

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \qquad f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

Truncation Error

Examples (Approximating Area under curve with Rectangles)

We have the exact value as

$$\int_3^9 x^2 dx = \left[\frac{x^3}{3}\right]_3^9 = \left[\frac{9^3 - 3^3}{3}\right] = 234$$

Using two rectangles of equal width to approximate the area (see Figure 1) under the curve, the approximate value of the integral

$$\int_3^9 x^2 dx \approx \left.(x^2)\right|_{x=3}(6-3) + \left.(x^2)\right|_{x=6}(9-6) = (3^2)3 + (6^2)3 = 27 + 108 = 135$$

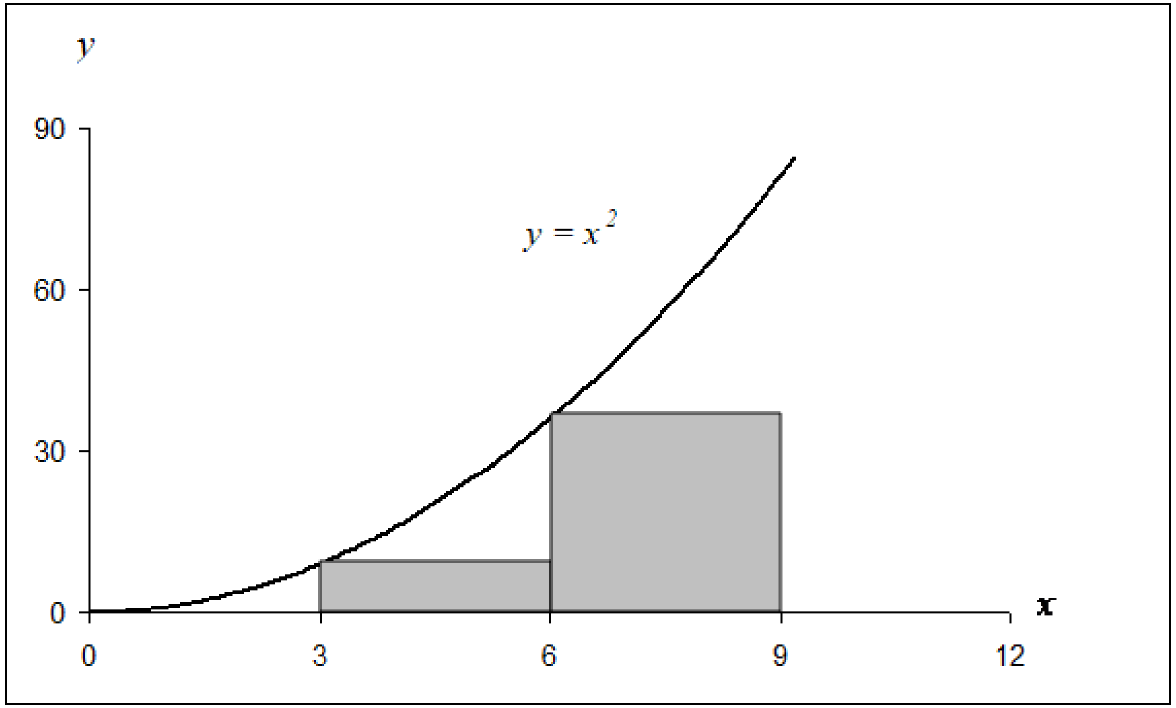


Figure 1 Plot of $y = x^2$ showing the approximate area under the curve from $x = 3$ to $x = 9$ using two rectangles.

$$\text{Truncation Error} = \text{Exact Value} - \text{Approximate Value} = 234 - 135 = 99.$$

Think: How to reduce this?