# MA 8019: Numerical Analysis I
# Mathematical Preliminaries

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

First version: May 2, 2018 Last updated: September 25, 2024

## A quick review of Calculus

- **$\varepsilon$-$\delta$ definition of limit:** *Let $\emptyset \neq A \subseteq \mathbb{R}$, $c$ be an accumulation point of $A$, and $f : A \to \mathbb{R}$ be a real-valued function. Then*

$$\lim_{x \to c} f(x) = L \iff \forall\, \varepsilon > 0\ \exists\, \delta > 0 \text{ such that if } x \in A \text{ and } 0 < |x - c| < \delta \text{ then } |f(x) - L| < \varepsilon.$$

  **Exercise:** Use $\varepsilon$-$\delta$ argument to show that $\lim_{x \to 3} 2x = 6$.

- Not all functions have limits everywhere.

  **Example:** $\lim_{x \to 0} \frac{|x|}{x}$ does not exist.

  $\because \lim_{x \to 0^+} \frac{|x|}{x} = 1 \neq -1 = \lim_{x \to 0^-} \frac{|x|}{x}$

## Intermediate-Value Theorem for continuous functions

- **Definition (continuity):** Let $f : A \to \mathbb{R}$ and $c \in A$.
  $f(x)$ *is said to be continuous at* $x = c \iff \lim_{x \to c} f(x) = f(c)$.

- **Examples:**

  (1) $f(x) = 2x$ is continuous at $x = 3$.

  (2) $f(x) = \dfrac{|x|}{x}$ is not continuous at $x = 0$.
  (no matter how it is defined at 0)

- **Intermediate-Value Theorem:** *If $f$ is a continuous function on $[a, b]$ and $K$ is any number between $f(a)$ and $f(b)$ (i.e., $f(a) < K < f(b)$ or $f(b) < K < f(a)$), then $\exists\, c \in (a, b)$ such that $f(c) = K$.*

- **Bolzano's Theorem:** *If $f$ is a continuous function on $[a, b]$ and $f(a)f(b) < 0$, then $\exists\, c \in (a, b)$ such that $f(c) = 0$.*

## Derivative

- **Definition:** *Let $f : A \to \mathbb{R}$ and $c \in A$. The derivative of $f$ at $c$ is defined by*

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c},$$

*if the limit exists. If $f'(c)$ exists then $f$ is said to be differentiable at $c$.*

- **Alternative definition:**

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

- **Theorem:** *If $f$ is differentiable at $c$, then $f$ must be continuous at $c$.*

  But the converse is not true! For example, $f(x) = |x|$ at $x = 0$.

# Pseudocode

A pseudocode to compute $f'(x)$ at $x = 0.5$ with $f(x) = \sin(x)$:

```
program numerical differentiation
    integer parameter n ← 10
    integer i
    real error, h, x, y
    x ← 0.5
    h ← 1
    for i = 1 to n do
        h ← 0.25h
        y ← (sin(x + h) − sin(x))/h
        error ← |cos(x) − y|
        output i, h, y, error
    end for
end program numerical differentiation
```

## Some notations

- $C(\mathbb{R})$ or $C^0(\mathbb{R})$: the set of all functions that are continuous on the real line $\mathbb{R}$.
- $C^1(\mathbb{R})$: the set of all functions for which $f'$ is continuous on the real line $\mathbb{R}$.
- $C^n(\mathbb{R})$: the set of all functions for which $f^{(n)}$ is continuous on the real line $\mathbb{R}$.
- $C^\infty(\mathbb{R}) \subset \cdots \subset C^n(\mathbb{R}) \subset C^1(\mathbb{R}) \subset C^0(\mathbb{R})$.

  **Example:** $f(x) = e^x \in C^\infty(\mathbb{R})$.
- $C^n([a,b])$: the set of all functions for which $f^{(n)}$ is continuous on the interval $[a,b]$.

# Taylor's Theorem with Lagrange remainder

*If $f \in C^n[a,b]$ and $f^{(n+1)}$ exists on $(a,b)$, then for any points $c$ and $x$ in $[a,b]$ we have*

$$f(x) = P_n(x) + E_n(x),$$

*where the n-th Taylor polynomial $P_n(x)$ is given by*

$$P_n(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(c)(x-c)^k$$

*and the remainder (error) term $E_n(x)$ is given by*

$$E_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x-c)^{n+1}$$

*for some point $\xi$ between $c$ and $x$ (either $c < \xi < x$ or $x < \xi < c$).*

## Some remarks

- The Taylor series of $f$ at $c$ is $\sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(c)(x-c)^k$.

  ($c = 0$, also called the Maclaurin series)
- If $E_n(x) \to 0$ as $n \to \infty$, then $P_n(x) \to f(x)$ as $n \to \infty$.

  i.e., $f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(c)(x-c)^k$.
- The special case $n = 0$ of Taylor's Theorem is the

  **Mean-Value Theorem:** *If $f \in C[a,b]$ and $f'$ exists on $(a,b)$, then for $x, c \in [a,b]$, $f(x) = f(c) + f'(\xi)(x-c)$ for some $\xi$ between $x$ and $c$.*
- A special case of the Mean-Value Theorem is **Rolle's Theorem:**

  *If $f$ is continuous on $[a,b]$, $f'$ exists on $(a,b)$, and $f(a) = f(b)$, then $\exists\, \xi \in (a,b)$ such that $f'(\xi) = 0$.*

## Example

Find the Taylor polynomial and the remainder term of $f(x) = \sin(x)$ at $c = 0$ and for which interval we get an error less than $3 \times 10^{-4}$ using 2 terms in the Taylor polynomial.

**Solution:**

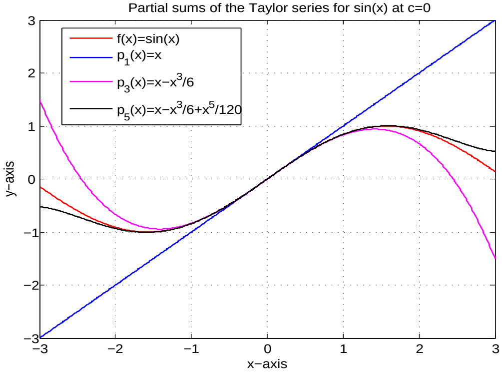
$$\text{Taylor polynomial} = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1},$$

$$\text{Remainder term} = \frac{(-1)^{n+1} \cos\xi}{(2n+3)!} x^{2n+3}.$$

$$n = 1: \quad |\text{Remainder term}\,| \leq \frac{|x|^{2n+3}}{(2n+3)!} = \frac{|x|^5}{5!} < 3 \times 10^{-4}.$$

$$\Longrightarrow |x - 0| < \left(360 \times 10^{-4}\right)^{1/5} \approx 0.514.$$
$$\Longrightarrow -0.514 < x < 0.514.$$

# Partial sums of the Taylor series for $f(x) = \sin(x)$ at $c = 0$




**Note:** *A Taylor series converges rapidly near the point of expansion and slowly (or not at all) at more remote points.*

## Taylor's Theorem with integral remainder

*If $f \in C^{n+1}[a,b]$ then for any points $c$ and $x$ in $[a,b]$ we have*

$$f(x) = P_n(x) + E_n(x),$$

*where the n-th Taylor polynomial $P_n(x)$ is given by*

$$P_n(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(c)(x-c)^k$$

*and the remainder term $E_n(x)$ is given by*

$$E_n(x) = \frac{1}{n!} \int_c^x f^{(n+1)}(t)(x-t)^n dt.$$

## Alternative form of Taylor's Theorem with L. remainder

*If $f \in C^n[a,b]$ and $f^{(n+1)}$ exists on $(a,b)$, then for any points $x$ and $x+h$ in $[a,b]$ we have*

$$f(x+h) = P_n(x) + E_n(h),$$

*where the n-th Taylor polynomial $P_n(x)$ is given by*

$$P_n(x) = \sum_{k=0}^{n} \frac{h^k}{k!} f^{(k)}(x)$$

*and the remainder term $E_n(h)$ is given by*

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

*for some point $\xi$ between $x$ and $x+h$.*

## Taylor's Theorem in two variables

*If* $f \in C^{n+1}([a,b] \times [c,d])$*, then for any points* $(x,y)$, $(x+h,y+k) \in [a,b] \times [c,d]$ *we have*

$$f(x+h,y+k) = \sum_{i=0}^{n} \frac{1}{i!}(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y})^i f(x,y) + E_n(h,k),$$

*where*

$$E_n(h,k) = \frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n+1} f(x+\theta h, y+\theta k)$$

*for some* $0 < \theta < 1$.

**Exercise:** What are the first few terms in the Taylor formula for $f(x,y) = \cos(xy)$?

For example, Taylor's formula with $n = 1$ is

$$\cos\big((x+h)(y+k)\big) = cos(xy) - hy\sin(xy) - kx\sin(xy) + E_1(h,k).$$

How about $n = 2$?

## Convergent sequences

- In numerical calculations, it often happens that a sequence of approximate answers is produced and hopefully converges to the desired solution.
- **Definition:** Let $\{x_n\}$ be a real sequence.

$$\lim_{n\to\infty} x_n = L \iff \forall\, \varepsilon > 0\ \exists\, n_0 \in \mathbb{N}\ \textit{s.t. if } n > n_0 \textit{ then } |x_n - L| < \varepsilon.$$

- **Example:** $\displaystyle\lim_{n\to\infty} \frac{n+1}{n} = 1.$

## Almost linear convergence

- For example, the sequence $x_n = \left(\frac{1+\frac{1}{2n}}{1-\frac{1}{2n}}\right)^n = \left(1 + \frac{2}{2n-1}\right)^n$ converges to the irrational number $e \approx 2.71828183$, $\lim_{n\to\infty} x_n = e$, also the famous sequence $y_n = (1 + \frac{1}{n})^n$ converges to $e$.

| $n$ | $x_n \downarrow$ | $y_n \uparrow$ |
|---|---|---|
| 1 | 3.00000000 | 2.00000000 |
| 10 | 2.72055141 | 2.59374246 |
| 30 | 2.71853357 | 2.67431878 |
| 50 | 2.71837244 | 2.69158803 |
| 100 | 2.71830448 | 2.70481383 |
| 1000 | 2.71828205 | 2.71692393 |

- $\{x_n\}$ converges faster than $\{y_n\}$, but both very slow.
- The ratio $\left|\frac{x_{n+1}-e}{x_n-e}\right| \to 1$ as $n \to \infty$ and similarly for $\{y_n\}$. This property is worse than linear convergence, we say *"almost linear convergence."*

## Superlinear convergence

- An example of a sequence that converges to $\sqrt{2}$ is

$$x_{n+1} = x_n - (x_n^2 - 2)\left(\frac{x_n - x_{n-1}}{x_n^2 - x_{n-1}^2}\right).$$

- Selecting two initial values, we have

$$\begin{aligned} &x_1 = 2.0, \quad x_2 = 1.5, \quad x_3 = 1.428571, \\ &x_4 = 1.414634, \quad x_5 = 1.414216, \quad x_6 = 1.414214, \cdots \end{aligned}$$

  The convergence to $\sqrt{2} \approx 1.41421356237310$ is quite rapid.

- Using double-precision computations, we find numerical evidence that

$$\frac{|x_{n+1} - \sqrt{2}|}{|x_n - \sqrt{2}|^{1.62}} \leq 0.77.$$

  We say *"superlinear convergence."*

## Rapid convergent sequences

- **Example:**

$$\begin{cases} x_1 = 2, & \\ x_{n+1} = \dfrac{1}{2}x_n + \dfrac{1}{x_n} & (n \geq 1). \end{cases}$$

  Few elements of this sequence: $x_1 = 2.000000$, $x_2 = 1.500000$, $x_3 = 1.416667$, $x_4 = 1.414216$.

  In fact, we can show that $\lim_{n\to\infty} x_n = \sqrt{2} \quad (\approx 1.41421356237310)$.

  **Hint:** First, show that $\{x_n\}$ is decreasing and bounded below. Then $\lim_{n\to\infty} x_n$ exists, say $x \cdots\cdots$.

- We find that $\dfrac{|x_{n+1} - \sqrt{2}|}{|x_n - \sqrt{2}|^2} \leq 0.36$. We say that this sequence converges quadratically *(quadratic convergence)*.

## Rate (order) of convergence

Let $\{x_n\}$ be a sequence of real numbers converges to $x^* \in \mathbb{R}$. We say the rate of convergence is

- at least linear: if $\exists\ 0 < C < 1$, $\exists\ n_0 \in \mathbb{N}$ such that

$$|x_{n+1} - x^*| \le C|x_n - x^*| \qquad \forall\ n \ge n_0.$$

- at least superlinear: if $\exists\ \{\varepsilon_n\}$ with $\varepsilon_n \to 0$ and $\exists\ n_0 \in \mathbb{N}$ s.t.

$$|x_{n+1} - x^*| \le \varepsilon_n|x_n - x^*| \qquad \forall\ n \ge n_0.$$

- at least quadratic: if $\exists\ C > 0$, $\exists\ n_0 \in \mathbb{N}$ such that

$$|x_{n+1} - x^*| \le C|x_n - x^*|^2 \qquad \forall\ n \ge n_0.$$

- of order $\alpha > 1$: if $\exists\ C > 0$, $\exists\ n_0 \in \mathbb{N}$ such that

$$|x_{n+1} - x^*| \le C|x_n - x^*|^\alpha \qquad \forall\ n \ge n_0.$$

## Big $O$ and little $o$ notation

- $x_n = O(\alpha_n)$ for two sequences $\{x_n\}$ and $\{\alpha_n\}$ if $\exists\, C > 0$ and $\exists\, n_0 \in \mathbb{N}$ s.t. $|x_n| \le C|\alpha_n|, \forall\, n \ge n_0$.

  **Example:** $\dfrac{n+1}{n^2} = O(\dfrac{1}{n})$.

- $x_n = o(\alpha_n)$ for two sequences $\{x_n\}$ and $\{\alpha_n\}$ if $\lim\limits_{n\to\infty} \dfrac{x_n}{\alpha_n} = 0$.

  (To avoid dividing by zero, sometimes modified as follows: if $\exists\, \{\varepsilon_n\}, \varepsilon_n \ge 0, \varepsilon_n \to 0$ and $\exists\, n_0 \in \mathbb{N}$ s.t. $|x_n| \le \varepsilon_n|\alpha_n|, \forall\, n \ge n_0$).

  **Example:** $e^{-n} = o(\dfrac{1}{n^2})$.

- These two notations give a coarse method of comparing two sequences. They are often used when both sequences converge to 0. *If $x_n \to 0, \alpha_n \to 0$, and $x_n = O(\alpha_n)$, then $x_n$ converges to 0 at least rapidly as $\alpha_n$. If $x_n = o(\alpha_n)$, then $x_n$ converges to 0 more rapidly than $\alpha_n$ does.*

## Big $O$ and little $o$ notation for functions

- $f(x) = O(g(x))\ (x \to \infty)$ for functions $f$ and $g$ if $\exists\, C > 0$ and $r > 0$ s.t. $|f(x)| \le C|g(x)|, \forall\, x \ge r$.

  **Example:** $\sqrt{x^2+1} = O(x)\ (x \to \infty)$.

  $\because\ \sqrt{x^2+1} \le 2x$ when $x \ge 1$.

- $f(x) = O(g(x))\ (x \to x^*)$ for functions $f$ and $g$ if $\exists\, C > 0$ and a neighborhood of $x^*$ s.t. $|f(x)| \le C|g(x)|, \forall\, x$ in the neighborhood.

- $f(x) = o(g(x))\ (x \to \infty)$ for functions $f$ and $g$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.

- $f(x) = o(g(x))\ (x \to x^*)$ for functions $f$ and $g$ if $\lim_{x\to x^*} \frac{f(x)}{g(x)} = 0$.

## Order of accuracy (order of convergence)

Let $u(x) = \sin(x)$ and $\bar{x} = 1$. Then $u'(1) = \cos(1) = 0.5403023\cdots$

$$D_+u(\bar{x}) := (u(\bar{x}+h) - u(\bar{x}))/h = u'(\bar{x}) + \frac{1}{2}hu''(\bar{x}) + \frac{1}{6}h^2u'''(\bar{x}) + O(h^3).$$

Then $D_+u(\bar{x}) \approx u'(\bar{x})$ as $h \to 0^+$.

**Table 1.1.** *Errors in various finite difference approximations to $u'(\bar{x})$.*

| $h$ | $D_+u(\bar{x})$ | $D_-u(\bar{x})$ | $D_0u(\bar{x})$ | $D_3u(\bar{x})$ |
|---|---|---|---|---|
| 1.0e−01 | −4.2939e−02 | 4.1138e−02 | −9.0005e−04 | 6.8207e−05 |
| 5.0e−02 | −2.1257e−02 | 2.0807e−02 | −2.2510e−04 | 8.6491e−06 |
| 1.0e−02 | −4.2163e−03 | 4.1983e−03 | −9.0050e−06 | 6.9941e−08 |
| 5.0e−03 | −2.1059e−03 | 2.1014e−03 | −2.2513e−06 | 8.7540e−09 |
| 1.0e−03 | −4.2083e−04 | 4.2065e−04 | −9.0050e−08 | 6.9979e−11 |

From the data in the above table, we have

$$D_+u(\bar{x}) - u'(\bar{x}) \approx -0.42h.$$

*(why and how? see page 23)*

## Log-log plot

If the error $E(h)$ behaves like $E(h) \approx Ch^p$, then

$$\log |E(h)| \approx \log |C| + p \log h.$$

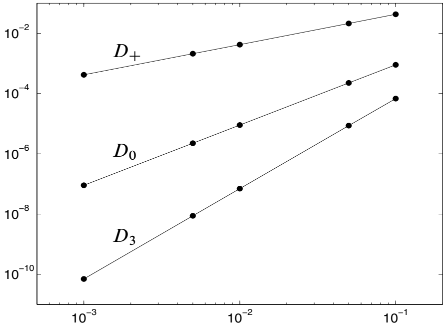



**Figure 1.2.** *The errors in* $Du(\bar{x})$ *from Table* 1.1 *plotted against* $h$ *on a log-log scale.*

## How to estimate the order of accuracy?

Assume a method is $p$-th order accurate, i.e., $E(h) \approx Ch^p$ for sufficiently small $h$. Then for $0 < h_2 < h_1$ small, we expect $E(h_1) \approx Ch_1^p$ and $E(h_2) \approx Ch_2^p$.

$$|E(h_1)| \approx |C|h_1^p, \quad |E(h_2)| \approx |C|h_2^p \Longrightarrow \frac{|E(h_1)|}{|E(h_2)|} \approx \frac{|C|h_1^p}{|C|h_2^p} = \left(\frac{h_1}{h_2}\right)^p$$

$$\Longrightarrow \log\left(\frac{|E(h_1)|}{|E(h_2)|}\right) \approx p \log\left(\frac{h_1}{h_2}\right) \Longrightarrow \boxed{p \approx \log\left(\frac{|E(h_1)|}{|E(h_2)|}\right) / \log\left(\frac{h_1}{h_2}\right)}$$

For example, for $D_+u(\bar{x})$, we have

```
log(4.2939e-02/2.1257e-02)/log(1.0e-01/5.0e-02) = 1.0144
log(2.1257e-02/4.2163e-03)/log(5.0e-02/1.0e-02) = 1.0052
log(4.2163e-03/2.1059e-03)/log(1.0e-02/5.0e-03) = 1.0015
log(2.1059e-03/4.2083e-04)/log(5.0e-03/1.0e-03) = 1.0005
```

p $\approx (1.0144 + 1.0052 + 1.0015 + 1.0005)/4 = 1.0054$, first order

*If the exact solution is not available, we can use an approximate solution with a very small h instead of the exact solution to estimate the order of accuracy.*