Math 151A – Applied Numerical Methods I

University of California, Los Angeles

Duc Vu

Fall 2021

This is math 151A – Applied Numerical Methods taught by Professor Jiang. We meet weekly on MWF from 1:00 pm to 1:50 pm for lecture. The recommended textbook for the class is *Numerical Analysis* 10^{th} by *Burden*, *Faires* and *Burden*. Other course notes can be found at my blog site. Please let me know through my email if you spot any typos in the note.

Contents

List of Theorems

List of Definitions

§1 Lec 1: Sep 24, 2021

§1.1 Calculus Review

- Intermediate Value Theorem (IVT): For continuous function $C([a,b])$, let $f \in C([a,b])$. Let $k \in \mathbb{R}$ s.t. k is strictly between $f(a)$ and $f(b)$. Then, $\exists$ some $c \in (a,b)$ s.t. $f(c) = k$.

  Question 1.1. Why is IVT useful?

  It guarantees the existence of solution to some nonlinear equations.

  Example 1.1

  Let $f(x) = 4x^2 - e^x$. IVT tells us $\exists x^*$ s.t. $f(x^*) = 0$.

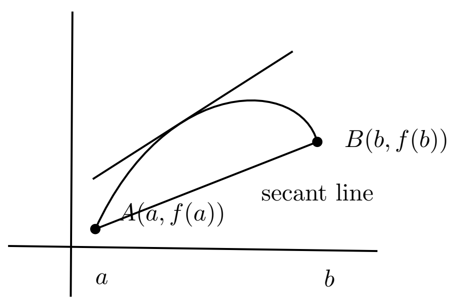
  $$f(0) = 0 - e^0 = -1 < 0$$
  $$f(1) = 4 - e > 0$$

  With $k = 0$, by IVT, $\exists c \in (0,1)$ s.t. $f(c) = 0$.

- Mean Value Theorem (MVT): If $f \in C([a,b])$ and f is differentiable in (a,b), then $\exists c \in (a,b)$ s.t.

  $$f'(c) = \frac{f(b) - f(a)}{b - a}$$

  in which $f'(c)$ is essentially the slope of the tangent line at $(c, f(c))$.



- Taylor's Theorem: Apply for a differentiable function, $f \in C^m([a,b])$ – f is m times continuously differentiable.

  Theorem 1.2 (Taylor)

  Let $f \in C^n([a,b])$. Let $x_0 \in [a,b]$. Assume $f^{(n+1)}$ exists on $[a,b]$. Then $\forall x \in [a,b]$, $\exists \xi(x) \in \mathbb{R}$ s.t. $x_0 < \xi < x$ or $x < \xi < x_0$. Then, we can express f as

  $$f(x) = P_n(x) + R_n(x)$$

  where

  $$P_n(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\frac{(x - x_0)^2}{2!} + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

  and

  $$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

Example 1.3

$f(x) = \cos(x)$, $x_0 = 0$

$$\begin{aligned} f(x) = \cos(x) &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(\xi(x))}{3!}x^3 \\ &= 1 + 0 - \frac{1}{2}x^2 + \frac{1}{6}x^3 \sin(\xi(x)) \end{aligned}$$

Note: Saying $f \in C^1$ is different from saying $f'(x)$ exists.

Example 1.4

Consider

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Have

$$\begin{aligned} f'(0) &= \lim_{h\to 0} \frac{f(0+h) - f(0)}{h} \\ &= \lim_{h\to 0} h \sin\left(\frac{1}{h}\right) \\ &= 0 \end{aligned}$$

But $f'(x)$ is not continuous. Specifically,

$$f'(x) = \begin{cases} 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Take sequence $\frac{1}{2k\pi}$, $f' \to -1$ and $\frac{1}{(2k+1)\pi}$, $f' \to 1$. Thus, the function is not continuous as it converges to two different values.

§2 Lec 2: Sep 27, 2021

§2.1 Errors and Convergence Rate

Fact 2.1.

1. Computers have finite memory
2. Only a subset of rational numbers $\mathbb{Q}$ can be exactly represented/stored.
3. Working with floating numbers instead of reals produces round-off error

Definition 2.1 (Error) — Let $p \in \mathbb{R}$, $\tilde{p}$ approximate to p. We define absolute error as

$$e_{\text{abs}} := |p - \tilde{p}|$$

We define relative error as

$$e_{\text{rel}} := \left|\frac{p - \tilde{p}}{p}\right|$$

Example 2.2

- $p = 1$, $\tilde{p} = 0.9$. In this case,

$$\begin{aligned} e_{\text{abs}} &= 0.1 \\ e_{\text{rel}} &= 0.1 \end{aligned}$$

- $p = 1000$, $\tilde{p} = 900$

$$\begin{aligned} e_{\text{abs}} &= 100 \\ e_{\text{rel}} &= 0.1 \end{aligned}$$

Finite Digit Arithmetic

Example 2.3

- π is rounded/chopped by computers
- $x = \frac{5}{7} = 0.\overline{714285}$, $y = \frac{1}{3} = 0.\overline{3}$
  Let $fl(x)$ is the floating point approx. to x. For example, we assume 5 digit rounding.

$$fl(x) = 0.71428, \qquad fl(y) = 0.33333$$

  Say if we want to add $x + y$ on computer

$$fl\left(fl(x) + fl(y)\right) = fl\left(1.04761\right) = 1.0476$$

Example 2.4

$f(x) = x^3 - 6.1x^2 + 3.2x + 1.5$ where $x = 4.71$. The exact value of $f(x)$ at $x = 4.71$ is -14.263899. Let's assume 3 digit rounding

$$\begin{aligned} fl(x^2) &= fl(4.71 \cdot 4.71) = fl(22.1841) = 22.2 \\ fl(x^3) &= 105 \\ fl(3.2x) &= 15.1 \\ fl\left(f(4.71)\right) &= -13.4 \end{aligned}$$

The relative error here is approximately 6% which is huge. Our example has 7 floating point operations (FLOPs). In order to reduce the floating point error, we want to nest the function

$$f(x) = ((x - 6.1)x + 3.2)x + 1.5 \quad - 5 \text{ FLOPs}$$

So $fl\left(f(4.71)\right) = -14.3$ and the $e_{\text{rel}} = 0.25\%$.

Remark 2.5. Every operation introduces error.

Order of convergence for Sequences:

Definition 2.6 (Order of Convergence) — For a convergent sequence $(p_n) = (p_1, p_2, p_3, \ldots)$. Let $p_n \to p$ as $n \to \infty$. Assume $p_n \neq p$. Then, if $\exists \lambda,\ \alpha$ with $0 < \lambda < \infty$ and $\alpha > 0$ s.t.

$$\lim_{n \to \infty} \frac{|p_{n+1} - p|}{|p_n - p|^\alpha} = \lambda$$

Then we say p_n converges to p with order α.

Example 2.7

$p_1 = 1, p_2 = \frac{1}{5}, \ldots, p_n = \frac{1}{5}p_{n-1}$ or $p_n = \frac{1}{5^{n-1}}$ where $p_n \to 0$ as $n \to \infty$.

$$\frac{|p_{n+1} - 0|}{|p_n - 0|^1} = \frac{\left(\frac{1}{5}\right)^n}{\left(\frac{1}{5}\right)^{n-1}} = \frac{1}{5}$$

So p_n converges with $\alpha = 1$

Problem 2.1. Test with $\alpha = 2$.

Definition 2.8 (Big O Notation) — We have $a(t) = \mathcal{O}\left(b(t)\right)$ where a is on the order of b $\iff$

$$\exists C > 0 \quad \ni \quad |a(t)| \leq Cb(t) \quad \text{for } t \to 0 \text{ or } t \to \infty$$

In practice, the definition is equivalent to

$$\lim_{t \to 0} \frac{|a(t)|}{b(t)} \text{ is bounded by a positive number}$$

§3 Lec 3: Sep 29, 2021

§3.1 Lec 2 (Cont'd)

Example 3.1

The Taylor's theorem for $\cos(h)$ about 0 is

$$\cos(h) = 1 - \frac{1}{2}h^2 + \frac{1}{24}h^4 \cos(\xi(h)) \text{ with some } 0 < \xi(h) < h$$

Denote $f(h) = \cos(h) + \frac{1}{2}h^2 - 1 = \frac{1}{24}h^4 \cos(\xi(h))$

$$\lim_{h \to 0} \frac{|f(h)|}{h^4} = \lim_{h \to 0} \frac{1}{24} |\cos(\xi(h))| \leq \frac{1}{24}$$

Thus y definition of big $\mathcal{O}$ notation,

$$f(h) = \mathcal{O}(h^4)$$

§3.2 Root Finding with Bisection

The goal is to find a root, or a zero, of a function f, i.e., find p s.t. $f(p) = 0$. First, let's assume

1. $f \in C([a,b])$
2. $f(a)f(b) < 0$

Then, $\exists p$ s.t. $f(p) = 0$ (by IVT).

Example 3.2

Consider:

$$f(x) = \sqrt{x} - \cos x, \qquad [a,b] = [0,1]$$

Then,

$$f(0) = -1, \qquad f(1) = 1 - \cos 1 > 0$$

Therefore, by IVT, $\exists p \in (0,1)$ s.t. $\sqrt{p} - \cos p = 0$.

Bisection Method (B.M): is an algorithm to approximate TBA

Algorithm 1: Bisection method (given $f(x) \in C([a,b])$, with $f(a)f(b) < 0$)

1. Set $a_1 = a$, $b_1 = b$
2. Set $p_1 = \frac{a_1+b_1}{2}$
3. *if* $f(p_1) == 0$ *then* we are done!
4. *else if* $f(p_1)$ has same sign as $f(a_1)$ *then* $p \in (p_1, b_1)$
   - Set $a_2 = b_1$, $b_2 = p_1$
5. *else if* $f(p_1)$ has same sign as $f(b_1)$ *then* $p \in (a_1, p_1)$
   - Set $a_2 = a_1$, $b_2 = p_1$
6. *end*

7. Set $p_2 = \frac{a_2+b_2}{2}$

8. Reset the entire if/else process.

> **Remark 3.3.** B.M. is similar to binary search in computer algorithms. If there exists multiple roots, e.g., $\{p, q, r\} \in [a, b]$, then the B.M. is guaranteed to find exactly one root, not all of them (but no guarantee exists for which one the method will find).

Stopping Criteria: We need a sequence $(p_1, p_2, \dots)$ and need specified tolerance ε. Choices for when to stop an algorithm:

- $|p_n - p_{n-1}| < \varepsilon$ – absolute difference between successive elements of the sequence
- $\frac{|p_n - p_{n-1}|}{|p_n|} < \varepsilon$ (assume $p_n \neq 0$) – relative difference
- $|f(p_n)| < \varepsilon$ – sometimes called a residual (how close are we to the answer).

§4 Lec 4: Oct 1, 2021

§4.1 Bisection Method (Cont'd)

Remark 4.1. B.M. is a global method, $f \in C([a, b])$ as long as the assumptions are satisfied, $f(a)f(b) < 0$, the B.M. will converge. In particular, it will converge to some point p s.t. $f(p) = 0$. Here "global" means the algorithm doesn't need a good initial guess p_0 unlike some "local" methods that we will cover later.

Example 4.2

The B.M. won't work for functions like $f(x) = x^2$ even though it has a root at $p = 0$ because we can't find any a, b that satisfies $f(a)f(b) < 0$.

Theorem 4.3 (Convergence Order of B.M.)

The sequence provided by B.M. satisfies

$$|p_n - p| \leq \frac{b-a}{2^n}$$

which approaches to 0 as $n \to \infty$.

This further tells us that the error bound of B.M. converges linearly. Recall from previous lectures that linear convergence for a convergent sequence (p_n) means that

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|^1} = \lambda \quad \text{for some finite positive } \lambda$$

and

$$p_n = \frac{b-a}{2^n}, \quad p = 0$$

We can easily show that $\lambda = \frac{1}{2}$.

Remark 4.4. The B.M. converges slowly compared to other methods. We will soon see that Newton's method has quadratic order of convergence.

§4.2 Fixed Points

Definition 4.5 (Fixed Point) — Let function g be $g : [a, b] \to \mathbb{R}$ and $p \in [a, b]$ s.t. $g(p) = p$. Then p is a fixed point of g.

Theorem 4.6

Let p be a fixed point of g, then p is also a root of $G(x) := g(x) - x$.

Proof. Obvious by definition. □

Given a root-finding problem $f(p) = 0$, we can define functions g with a fixed point at p in a number of ways. For example, as $g(x) = x - f(x)$ or as $g(x) = x + 3f(x)$.
A fixed point for g just corresponds to the intersection between $y = g(x)$ and $y = x$.
Fixed Point Iteration (F.P.I): the F.P.I method is quite simple. For $g \in C([a, b])$ and $p_0 \in [a, b]$ we set $p_{n+1} = g(p_n)$. We also need $g(x) \in [a, b]$, otherwise at some point of the algorithm we won't be able to proceed to evaluate g. Also note that the initial guess p_0 is arbitrary.

$$p_1 = g(p_0), \quad p_2 = g(p_1), \ldots, p_{n+1} = g(p_n)$$

We use the same stopping criteria as in B.M.

Example 4.7 (F.P.I Failure Case)
To solve $x^2 - 7 = 0$, it is equivalent to $x = \frac{7}{x}$. We want to use F.P.I to find $p = \sqrt{7} \approx 2.64575\ldots$, so we can set

$$g_1(x) = \frac{7}{x}$$

then the goal is to find p s.t. $p = g_1(p)$. Another option is to use

$$g_2(x) := \frac{x + \frac{7}{x}}{2} = x$$

Let $p_0 = 3$ we can show that

- $g_1(x)$: $p_0 = 3$, $p_1 = \frac{7}{3}$, $p_2 = 3, \ldots$, oscillates between 2 numbers
- $g_2(x)$: $p_0 = 3$, $p_1 = 2.666\ldots$, $p_2 = 2.645833\ldots, \ldots$

Example 4.8
$x^3 + 4x^2 - 10 = 0$ has a unique root in $[1, 2]$, i.e. $p = 1.365230013$.

a) $x = g_1(x) = x - x^3 - 4x^2 + 10$ – does not converge

b) $x = g_2(x) = \left(\frac{10}{x} - 4x\right)^{\frac{1}{2}}$ – does not converge

c) $x = g_3(x) = \frac{1}{2}(10 - x^3)^{\frac{1}{2}}$ – converge

d) $x = g_4(x) = \left(\frac{10}{x+4}\right)^{\frac{1}{2}}$ – converge

e) $x = g_5(x) = x - \frac{x^3+4x^2-10}{3x^2+8x}$ – converge

We can see that the choice of $g(x)$ is critical to determine whether the algorithm converges. Before delving into that problem, let's first establish a theorem about the existence of a fixed point.

Theorem 4.9 (Existence of a Fixed Point)
Let $g \in \mathbb{C}([a, b])$ with $a \leq g(x) \leq b$. Then, $\forall x \in [a, b]$, $\exists$ at least one fixed point p s.t. $g(p) = p$.