Numerical Analysis Notes - Week 1

Riya Mokashi

January 26, 2022

1 Fixed Point Iterations

1.1 The Babylonian Method

At the end of the last class we saw the Babylonian Algorithm, an iterative method for computing square roots by hand.

$$y_{k+1} = \frac{1}{2}\left(\frac{x}{y_k} + y_k\right) \tag{1}$$

We also learned that if $y_0 > 0$ (the initial iterate), then y_k converges to $\sqrt{x}$. With each additional iteration, the number of digits of accuracy also doubles and eventually plateaus. This ties into the intricacies of floating point arithmetic where double precision floats roughly have the capacity for 16 digits of agreement (capped at a fourth iteration of the Babylonian Algorithm)

The Babylonian Method is an example of a fixed-point iteration which is a term that describes methods used to compute fixed points of functions.

1.2 FPI - Definition

Fixed Point Iteration 1. Let $f : R \to R_1$ and let x_0 be fixed (somehow) then we'll call the sequence generated by $x_{k+1} = f(x_k)$ where $(k > 0)$ a simple iteration (aka a fixed point iteration).

1.3 Fixed Point Iteration Conditions

Before looking further there are a few things we need to think about.

1. First, is there anything special about fixed points?
2. Second, if we have an iteration like $x_{k+1} = f(x_k)$ should we automatically expect convergence to a fixed point?

As a side note, in class we learned that Numerical Analysis largely involves designing algorithms to solve continuous problems by finding approximate solutions. Often, if the solution for a problem exists at all and the solution is unique, we will be more likely to succeed in our endeavors. Alternatively, Computer

Science is more often concerned with finding the exact solution to a discrete problem.

In general Numerical Analysis is concerned with two main concepts:

1. Existence
2. Uniqueness

To apply these concepts to fixed point iterations we can ask - for a function f should a fixed point exist? If it does, is it unique?

To take things further, the Babylonian Algorithm shows us:

$$y_{k+1} = f(y_k)$$
$$f(y) = \frac{1}{2}\left(\frac{x}{y} + y\right)$$

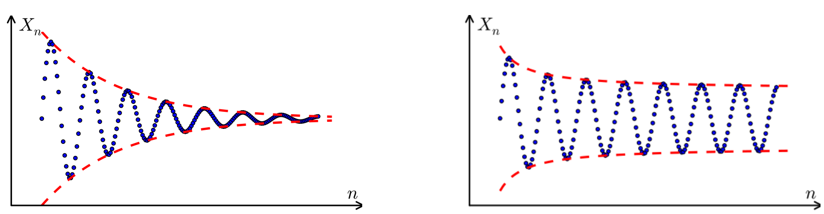
This begs the question, does $(y_k)_{k=0}^{\infty}$ converge?

Cauchy Sequences 1. $(y_k)_{k=0}^{\infty}$ is Cauchy (converges) if the following conditions are met:

$$\forall \epsilon > 0 \exists N(\epsilon) \geq 0 \tag{2}$$

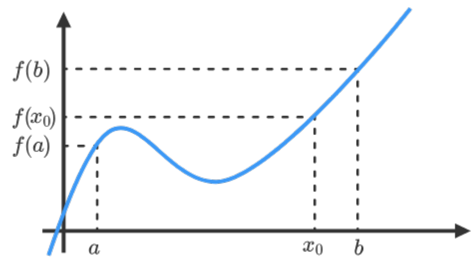
$$\forall m, n \geq N(\epsilon) : D(y_m, y_n) < \epsilon \tag{3}$$

- N = Some integer
- ϵ = Error tolerance
- D = Distance



(a) Cauchy sequence (b) Non-Cauchy sequence

Figure 1: As seen in the figures above, a Cauchy sequence converges to an "ultimate destination", or in other words, a limit clearly exists



(a) Intermediate Value Theorem

Figure 2: The Intermediate Value Theorem establishes the existence any given value between $f(a)$ and $f(b)$ at some point within $[a, b]$ given that the function in continuous throughout the interval.

1.4 Determining the Existence of Fixed Points

1.4.1 Intermediate Value Theorem

The first way to determine the existence of fixed points is to utilize the Intermediate Value theorem or IVT.

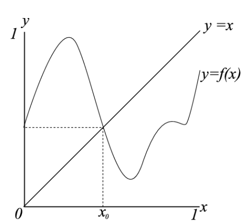
Intermediate Value Theorem 1. The theorem states that if f is real-valued: $f \in C^0([a, b])$ continuous on the interval a,b where a,b is closed and bounded. Then for each y such that

$$\min_{a \leq x \leq b} f(x) \leq y \leq \max_{a \leq x \leq b} f(x) \tag{4}$$

there exists $x \in [a, b]$ such that $y = f(x)$

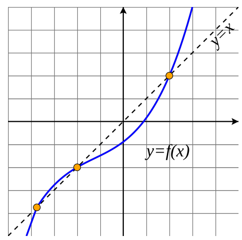
1.4.2 Brouwer's Fixed Point Theorem (in 1D)

Brouwer's Fixed Point Theorem 1. Assume that f is real-valued and assume that $f([a, b]) \leq [a, b]$. Then $\exists \xi$ such that $f(\xi) = \xi$



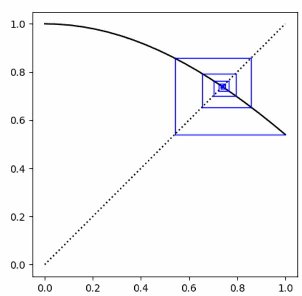
(a) Brouwer's Theorem

A fixed point for $f : R \to R$ is a point where $y = x$ intersects the graph of f. However, it is not always guaranteed that a fixed point will exist or that there will be only one fixed point.



(a) Multiple Fixed Points

As seen above, a function f can have multiple fixed points. If we have identified multiple points, how can we tell which fixed point the iteration $x_{k+1} = f(x_k)$. Food for thought: We can utilize a mapping technique called cobweb plotting to graphically solve for a fixed point on f. However, depending on which part of the function you start, the cobweb plot will reach a different fixed point given that there are multiple fixed points.



(a) Cobweb Plot

Figure 5: The cobweb plot reaches a fixed point by following a path back and forth across $f \subseteq y = x$.

Note: If f is continuous and $f([a, b]) \leq [a, b]$ and if x_k generated by $x_{k+1} = f(x_k)$ converges to ξ then:

$$\xi = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} f(x_k) = f(\lim_{k\to\infty} x_k) = f(\xi). \tag{5}$$

This tells us that if the preconditions for Brouwer's Theorem hold and the sequence converges, then it will converge to a fixed point. Given this, we are left with the questions - Does x_k converge and to which fixed point will it converge to?

1.5 Contractions and Contraction Mapping

To determine a unique fixed-point of convergence it is viable to use contractions or contraction mapping.

1.5.1 Contraction Definition and Lipschitz Constants

Contraction Definition 1. Let $f \leftarrow C^0[a,b]$ on a finite interval $[a,b]$. Then f is a contraction if $\exists L < 1$ such that $|f(x) - f(y) < L|x - y|$ for all x, y.

L is called a Lipschitz Constant for f over the interval $[a,b]$. To define the term further: If $L > \max_{a \leq x \leq b} |f'(x)|$ then L is a Lipschitz constant for f over $[a,b]$. Note: $L < 1$ is not required for a Lipschitz constant.

1.5.2 Contraction Mapping Theorem

Contraction Theorem 1. If $f \in C^0[a,b], f([a,b]) \subseteq [a,b]$, and f is a contraction on $[a,b]$. Then f has a unique fixed point ξ on $[a,b]$ and $x_{k+1} = f(x_k)$ converges to ξ. To define further - let $x_{k+1} = f(x_k)$ generate a sequence and assume it converges to ξ. Then the basin of attraction of ξ is all x_0 such that $x_k \to \xi$.

To apply the contraction mapping theorem to f over $[a,b]$:

1. Check if $f \leftarrow C^0[a,b]$ if $f([a,b]) \leq [a,b]$
2. Compute $L = max_{a \leq x \leq b}|f'(x)|$
3. Check if $L < 1$
4. If the above condition is met, there exists a unique fixed point for f, call it ξ in $[a,b]$.
5. Compute ξ: Pick any $x_0 \leftarrow [a,b]$ and run $x_{k+1} = f(x_k)$ until convergence.
   - At the fixed point $x_k = \xi$ and $x_{k+1} = f(x_k) = f(\xi) = \xi$
6. Compute $x_{k+1} - x_k$ to check accuracy

Note: If you are interested in further research: more general questions can be answered by discrete dynamical systems (Henon maps, chaotic iterated maps (2D), etc.).

2 Homework Notes

Sturm Chain Definition 1. Let p, where $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 ....$ Define $p_0 = p, p_1 = p'(x) = \frac{dp}{dx}$ For $n \geq 2$ define $q_n = [\frac{p_{n-2}}{p_{n-1}}]$ where q_n is the quotient result of polynomial long division (without the remainder). Then define, for the same n, $p_n = q_n(p_{n-1}) - p_{n-2}$ stopping when p_N is constant. The resulting sequence $p_0, ..., p_N$ is called a Sturm chain.

2.1 Sturm's Theorem

Sturm Chain Theorem 1. Let p be a polynomial. Let $a < b$ define a bounded interval $[a, b]$ and let $p_0, ..., p_N$ be the Sturm chain corresponding to p. Now consider the table below where Δ_a and Δ_b are the number of sign changes in each row.

$p_0 a$	$p_1 a$	...	$p_n a$	$\rightarrow$	Δ_a
$p_0 b$	$p_1 b$	...	$p_n b$	$\rightarrow$	Δ_b

The number of real roots in (a, b) is $|\Delta_b - \Delta_a|$.

2.1.1 Sturm's Theorem Example

For example, say you are given a polynomial $p_0 = x^5 - 3x - 1$. You can calculate the Sturm Chain below. Note that it took three iterations to reach a constant.

$$\begin{aligned} p_0 &= x^5 - 3x - 1 \\ p_1 &= 5x^4 - 3 \\ p_0 &= \frac{1}{5}(12x + 5) \\ p_3 &= \frac{59083}{20736} \end{aligned}$$

You can then use the Sturm chain to generate the following sign change table.

x	Sign p_0	Sign p_1	Sign p_2	Sign p_3	Δ
-2	$-$	$+$	$-$	$+$	3
0	$-$	$-$	$+$	$+$	1
2	$+$	$+$	$+$	$+$	0

This table can now be used to identify the number of roots in each interval. $\Delta_{x=0} - \Delta_{x=2} = 1$ therefore there is only one root in this interval. This is a good place to use a tool such as `brentq` to identify the single root. However, $\Delta_{x=-2} - \Delta_{x=0} = 2$ meaning there are multiple roots in this interval. To proceed, Sturm's theorem will need to be reapplied until you are able to separate the roots into separate intervals.

For the homework, a viable method is combining Sturm's theorem with binary search using bisection. Utilizing python generators could also be beneficial - link to the documentation can be found here.