Introduction to Computational Mathematics (AMATH 242/CS 371)

University of Waterloo
Winter 2019

Root Finding

The growth of a population can often be modeled over short periods of time by assuming that the population grows continuously with time at a rate proportional to the number present at that time. Suppose that $N(t)$ denotes the number in the population at time t and λ denotes the constant birth rate of the population. Then the population satisfies the differential equation $\dfrac{dN(t)}{dt} = \lambda N(t)$. Solution of this differential equation is $N(t) = N_0 e^{\lambda t}$. If immigration is permitted at a constant rate ν, then the differential equation becomes $\dfrac{dN(t)}{dt} = \lambda N(t) + \nu$ with $N(t) = N_0 e^{\lambda t} + \dfrac{\nu}{\lambda}(e^{\lambda t} - 1)$ as its solution.

Root Finding

If N_0, is given and we also have ν and $N(1)$ then finding the birth rate of this population is not easy. We should solve $N(1) = N_0 e^{\lambda} + \frac{\nu}{\lambda}(e^{\lambda} - 1)$. It is not possible to solve explicitly for λ in this equation and we should use numerical methods.

Now, lets define a problem in general:

Given any function f(x), find x^* such that $f(x^*) = 0$. The value x^* is called a root of the equation f(x)=0.

Even if the existence of root of a function is guaranteed, there is no guaranteed method for finding x^*. In fact there are several techniques that can work for many cases.

• Since x^* may not be defined in a floating point number system, we will not be able to find x^* exactly. Therefore, we consider a computational version of the same problem:

Given any $f(x)$ and some error tolerance $\epsilon > 0$, find x^* such that $|f(x^*)| < \epsilon$.

Example:

• Find roots of the following functions:

$$f(x) = x^2 - x - 12$$
$$f(x) = x^3 - x - 1$$
$$f(x) = cos(e^x - 2)$$

• In finding roots of a function numerically, a sequence is generated and and we require that iterates eventually converge to the actual root.
• Important question that should be answered is that "how do we know where a root of a function may approximately be located?"

Intermediate value theorem

If $f(x)$ is continuous on a closed interval $[a, b]$ and $c \in [f(a), f(b)]$, then $\exists x^* \in [a, b]$ such that $f(x^*) = c$.

Four Algorithms for root finding: Bisection Method

- A simple method to find roots
- Convergence is guaranteed
- f(x) should be continuous on [a,b] and $f(a).f(b) < 0$
- This method works by bisecting the interval and recursively using the Intermediate Value Theorem

$$a_k = \begin{cases} a_{k-1} & \text{if} \quad f(a_{k-1}).f((a_{k-1}+b_{k-1})/2) \leq 0 \\ (a_{k-1}+b_{k-1})/2 & \text{otherwise} \end{cases}$$

$$b_k = \begin{cases} b_{k-1} & \text{if} \quad f(a_{k-1}).f((a_{k-1}+b_{k-1})/2) > 0 \\ (a_{k-1}+b_{k-1})/2 & \text{otherwise} \end{cases}$$

Therefore, convergence to a solution is guaranteed.
Question: How many steps does bisection method take to reach the desired tolerance?
$[a, b]$ is the initial interval of the solution. In each iteration the interval containing x^* is halved. Therefore, if it takes n steps to fulfill $|b-a| \leq t$ we have that:

$$2^{-n}|b-a| \leq t \Longrightarrow n \geq \frac{1}{\log 2} \log(\frac{|b-a|}{t})$$

This is the number of steps that bisection method should take to converge.
Examples:
• Show that $f(x) = x^3 + 4x^2 - 10 = 0$ has a root in [1,2]. Use the bisection method to determine an approximation to the root that is accurate of at least within 10^{-4}.
• Determine the number of iterations necessary to solve $f(x) = x^3 + 4x^2 - 10 = 0$ with accuracy 10^{-3} using $a = 1$ and $b = 2$.

Four Algorithms for root finding: Fixed Point Iteration

• Let $f(x) = (x+2)(x+1)x(x-1)^3(x-2)$. To which zero of f does the Bisection method converge when applied on $[-3, 2.5]$.
A *fixed point* for a function is a number at which the value of the function does not change when the function is applied.

Definition

x^* is a fixed point of g(x) if $g(x^*) = x^*$ i.e. if x^* is mapped to itself under g.

We can define root-finding problem in another way. Consider the real-valued function g, defined by g(x)=x-f(x) so f(x)=x-g(x). We observe that fixed point of g(x) is a root of f(x).

- Fixed point Iteration method is convergent if g(x) satisfies certain properties.
- g(x)=x-f(x) is not the only way to define g(x). In general we can write g(x)= x-H(f(x)) as long as we choose H such that H(0)=0.

Fixed Point Iteration

- It is not always straightforward to find the fixed point of a function explicitly. So, we start by an initial approximation x_0 and generate the sequence $\{x_n\}_{n=0}^{\infty}$ by letting $x_n = g(x_{n-1})$. If this sequence is convergent to x^*, then $x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} g(x_{n-1}) = g(x^*)$.

Example:

• The equation $x^3 + 4x^2 - 10 = 0$ has a unique root in [1,2]. There are many ways to change the equation to the fixed-point form $x = g(x)$:

$$(a)\ g_1(x) = x - x^3 - 4x^2 + 10 \qquad (b)\ g_2(x) = \left(\frac{10}{x} - 4x\right)^{\frac{1}{2}}$$

$$(c)\ g_3(x) = \frac{1}{2}\left(10 - x^3\right)^{\frac{1}{2}} \qquad (d)\ g_4(x) = \left(\frac{10}{x+4}\right)^{\frac{1}{2}}$$

How do these functions affect the convergence of fixed-point iteration if $x_0 = 1.5$?

Four Algorithms for root finding: Newton's Method

Newton's (or the Newton-Raphson) method is one of the most powerful and well-known numerical methods for solving a root-finding problem.
If f(x) is a continuous and differentiable function on $[a, b]$, then we can use Taylor series expansion around x_0 to derive Newton's method:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi(x))}{2!}(x - x_0)^2 \Longrightarrow$$
$$f(x^*) = 0 = f(x_0) + f'(x_0)(x^* - x_0) + \frac{f''(\xi(x^*))}{2!}(x^* - x_0)^2$$

Assuming that $|x^* - x_0|$ is small, the last term is negligible. So,

$$0 \approx f(x_0) + f'(x_0)(x^* - x_0) \Longrightarrow$$
$$x^* = x_0 - \frac{f(x_0)}{f'(x_0)} \equiv x_1$$

This sets the stage for Newton's method, which starts with an initial approximation x_0 and generates the sequence $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$

Four Algorithms for root finding: Newton's Method

The Taylor series derivation of Newton's method points out the importance of an accurate initial approximation. The crucial assumption is that the term involving $(x^* - x_0)^2$ is so small that it can be deleted. This will clearly be false unless x_0 is a good approximation to x^*. If x_0 is not sufficiently close to the actual root, there is little reason to suspect that Newton's method will converge to the root. However, in some instances, even poor initial approximations will produce convergence.

Four Algorithms for root finding: Secant Method

Newton's method is an extremely powerful technique, but it has a major weakness: the need to know the value of the derivative of f at each approximation. However, we can approximate the derivative using a numerical scheme that requires only evaluations of the function f(x).

$$f'(x_n) = \lim_{\eta \to x_n} \frac{f(x_n) - f(\eta)}{x_n - \eta} \approx \frac{f(x_n) - f(\eta)}{x_n - \eta}$$

using $\eta = x_{n-1}$ results in

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

Using this approximation in Newton's method, defines equation for the secant method:

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

Stopping Criteria for Iterative Functions

In all aforementioned root-finding methods, a stop criterion should be defined to terminate iteration:

- **Maximum number of steps.** We impose some maximum number of steps in the iteration n_{max} and stop when $n = n_{max}$.
- **Tolerance on the size of the correction.** Under this criterion, we are given some tolerance t and stop when $|x_{n+1} - x_n| \leq t$. This criterion does not guarantee that $|x_{n+1} - x^*| \leq t$, in general.
- **Tolerance on the size of the function value**. Under this criterion, we are given some tolerance t and stop when $|f(x_n)| \leq t$.

Error Analysis for Iterative Methods

For sequence $\{x_n\}_{n=0}^{\infty}$ and point x^*, the error at iteration n is defined as $e_n = x_n - x^*$. We will define the rate of convergence by how quickly the error converges to zero i.e. how quickly $\{x_n\}_{n=0}^{\infty}$ converges to x^*.

• The sequence $\{x_n\}_{n=0}^{\infty}$ converges to x^* with order q if and only if $\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^q} = N$ where $N \in (0, \infty)$. q is the order of convergence and N is asymptotic error constant.

Two cases of order are given special attention:

1. If q$=$ 1 (and $0 < N < 1$), the sequence is linearly convergent.
2. If q$=$ 2, the sequence is quadratically convergent.

Error Analysis for Iterative Methods

Why do we try to find methods that produce higher-order convergent sequences?

Illustration: Suppose that $\{x_n\}_{n=0}^{\infty}$ is convergent to 0 linearly and $\{\tilde{x}_n\}_{n=0}^{\infty}$ is convergent to 0 quadratically:

$$\lim_{n\to\infty} \frac{|x_{n+1}|}{|x_n|} = 0.5 \qquad \text{and} \lim_{n\to\infty} \frac{|\tilde{x}_{n+1}|}{|\tilde{x}_n|^2} = 0.5$$

For simplicity we assume that $N = 0.5$ for both sequences and for each n. Therefore

$$\frac{|x_{n+1}|}{|x_n|} \approx 0.5 \qquad \text{and} \frac{|\tilde{x}_{n+1}|}{|\tilde{x}_n|^2} \approx 0.5$$

For the linearly convergent scheme, this means that

$$|x_n - 0| = |x_n| \approx 0.5|x_{n-1}| \approx (0.5)^2|x_{n-2}| \approx \cdots \approx (0.5)^n|x_0|$$

whereas the quadratically convergent procedure has

$$|\tilde{x}_n - 0| = |\tilde{x}_n| \approx 0.5|\tilde{x}_{n-1}|^2 \approx (0.5)[0.5|x_{n-2}|^2]^2 = (0.5)^3|x_{n-2}|^4$$
$$\approx \cdots \approx (0.5)^{2^n-1}|\tilde{x}_0|^{2^n}$$

If $|x_0| = |\tilde{x_0}|$, $\{\tilde{x}_n\}_{n=0}^{\infty}$ converges to 0 quickly.

Bisection Method: Convergence of bisection method is guaranteed. Although the interval that contains the solution is halved at each iteration and consequently error becomes smaller, but error may increase for certain iterations. To find convergence rate of Bisection method, we consider sequence $\{L_n\}_{n=1}^{\infty}$ with $L_n = |b_n - a_n|$ the length of the interval at step n. We know that $L_{n+1} = \frac{1}{2}L_n$ and so the sequence $\{L_n\}_{n=1}^{\infty}$ converges to 0 linearly. We also know that $|e_n| \leq L_n$, and so we say that $\{e_n\}$ converges to 0 at least linearly.

Error Analysis for Iterative Methods

• The upper bound of error in bisection method is $|x_n - x^*| \leq \dfrac{b-a}{2^n}$.

Definition. Suppose that g is a real-valued function, defined and continuous on a bounded closed interval [a,b] of the real line. Then, g is said to be a **contraction** on [a,b] if there exists a constant $L \in (0, 1)$ such that $|g(x) - g(y)| \leq L|x - y|$ for $x, y \in [a, b]$.

Results:

- $|g(b) - g(a)|$ is smaller than $|b - a|$ i.e. the interval [a,b] has been contracted to a smaller interval [g(a),g(b)] or [g(b),g(a)].
- From the definition we have $\dfrac{|g(x) - g(y)|}{|x - y|} \leq L < 1$. Thus, slope of any secant line within the interval [a,b] cannot exceed L in absolute value.

Another way of defining contraction function is that if we have g(x) differentiable on [a,b] with $|g'(x)| < 1$ for all x in [a,b], then g(x) is a contraction on [a,b] with $L = \max_{x \in [a,b]} |g'(x)|$

Error Analysis for Iterative Methods

Question: How can we find a fixed-point problem that produces a sequence that reliably and rapidly converges to a solution to a given root-finding problem?

Theorem: Contraction Mapping Theorem

Let g(x) be a real-valued function, defined and continuous on a bounded closed interval [a,b] of the real line, and assume that $g(x) \in [a, b]$ for all $x \in [a, b]$. Suppose, further, that g(x) is a contraction on [a,b]. Then,

- g(x) has a unique fixed point * in the interval [a,b].
- The sequence $\{x_k\}$ by $x_{k+1} = g(x_k)$ converges to x^* as $k \to \infty$ for any starting value x_0 in [a,b].

Error Analysis for Iterative Methods

Note that conditions for the existence of a unique fixed point ($g(x) \in [a, b]$ for all $x \in [a, b]$ and $|g'(x)| \leq 1$ are sufficient to guarantee a unique fixed point but are not necessary.
For example function $g(x) = 3^{-x}$ on [0,1] satisfies the condition $g(x) \in [a, b]$ but $|g'(x)| > 1$ for x in [0,1]. However, it has a unique fixed-point in this interval.

Error Analysis for Iterative Methods

Fixed Point Iteration: The rate of convergence for this method is highly variable and depends greatly on the actual problem being solved. We can show that under special circumstance, fixed point iteration converges linearly

Corollary. Let g be a real-valued function, defined and continuous on a bounded closed interval [a,b] of the real line, and assume that $g(x) \in [a, b]$ for all $x \in [a, b]$. Let $x^* = g(x^*)$ be a fixed point of g(x) with $x^* \in [a, b]$. Assume there exists δ such that $g'(x)$ is continuous in $I_\delta = [x^* - \delta, x^* + \delta]$. Define the sequence $\{x_n\}_{n=0}^{\infty}$ by $x_{n+1} = g(x_n)$. Then:

- If $|g(x^*)| < 1$ then there exists ϵ such that $\{x_n\}$ converges to x^* for $|x_0 - x^*| < \epsilon$. Further, convergence is linear with $N = g'(x^*)$.
- If $|g'(x^*)| > 1$ then $\{x_n\}$ diverges for any starting value x_0.

Using this corollary, we can come up with a method of choosing our form for g(x) in terms of f(x) depending on the derivative at the point x^*.

This theorem implies that higher-order convergence for fixed-point methods of the form $g(x) = x$ can occur only when $g'(x^*) = 0$.

Theorem

Let x^* be a solution of the equation $x = g(x)$. Suppose that $g'(x^*) = 0$ and g" is continuous with $|g"(x)| < M$ on an open interval I containing x^*. Then there exists a $\delta > 0$ such that, for $x_0 \in [x^* - \delta, x^* + \delta]$, the sequence defined by $x_n = g(x_{n-1})$, converges at least quadratically to x^*. Moreover, for sufficiently large values of n, $|x_{n+1} - x^*| < \dfrac{M}{2}|x_n - x^*|^2$.

Error Analysis for Iterative Methods

Convergence theorem for Newton's method

If $f(x^*) = 0$, $f'(x^*) \neq 0$ and f, f' and f" are all continuous in $I_\delta = [x^* - \delta, x^* + \delta]$ with x_0 sufficiently close to x^* then the sequence $\{x_n\}_{n=0}^{\infty}$ converges quadratically to x^* with $N = \dfrac{f''(x^*)}{2f'(x^*)}$.

If $f'(x^*) = 0$, then the rate of convergence degrades to linear convergence. One method of handling the problem of multiple roots of a function f(x) is to define

$$\mu(x) = \frac{f(x)}{f'(x)}$$

in which $f(x) = (x - x^*)^m q(x)$. We can show that x^* is a simple zero of $\mu(x)$. Newton's method can then be applied to $\mu(x)$ to give $g(x) = x - \dfrac{\mu(x)}{\mu'(x)}$. If g(x) is continuous, functional iteration applied to g(x) will be quadratically convergent regardless of the multiplicity of the zero of f(x).

Theoretically, the only drawback to this method is the additional calculation of f''(x). In practice, however, multiple roots can cause serious round-off problems because the denominator of the modified Newton's method consists of the difference of two numbers that are both close to 0. **Secant Method.** The secant method has a very similar convergence theorem to that of Newton's method. In fact, all that changes is the order of convergence of the method. Similar to Newton's method, it can be shown that if $f'(x^*) = 0$ then the rate of convergence degrades to linear.

Convergence theorem for secant method

If $f(x^*) = 0$, $f'(x^*) \neq 0$ and f, f' and f'' are all continuous in $I_\delta = [x^* - \delta, x^* + \delta]$ with x_0 sufficiently close to x^* then the sequence $\{x_i\}_{n=0}^{\infty}$ converges to x^* with order $q = \frac{1}{2}(1 + \sqrt{5})$.