Nonlinear Equations: Bisection Method

Martin Licht

UC San Diego

Winter Quarter 2021

Bisection Method
Motivation

In this lecture, we discuss the algorithmic solution of the nonlinear equation

$$f(x) = 0$$

where f is a continuous function.

This means, we want to find a root of that function.

More generally, solving the system

$$g(x) = y$$

where g is a continuous function, can be written as finding a root of

$$f(x) = 0$$

where $f(x) = g(x) - y$.

Rule of thumb: solving any system of equations can be written as finding a root of a function.

That's why root finding algorithms receive so much attention in computational mathematics.

Analytical Background

Bisection Method

Analytical Background

One fundamental property of the real numbers is the ***principle of nested intervals***.

Theorem (Principle of Nested Intervals)

Given a sequence of intervals $[a_n, b_n]$ *that are nested,*

$$[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$$

and whose length goes to zero,

$$\lim_{n \to \infty} b_n - a_n = 0,$$

there exists a unique real number c contained within all the intervals.

We call c the ***limit*** of the nested intervals.

Remark

The principle of nested intervals is a fundamental property of the real numbers.

It can be used to axiomatically construct the real numbers from the rational numbers. There are different principles that could be used equivalently in the following proofs, such as the least upper bound axiom.

Analytical Background

Lemma

Given a sequence of nested intervals $[a_n, b_n]$ *and whose lengths go to zero, with limit point c, we have*

$$\lim_{n\to\infty} a_n = c, \qquad \lim_{n\to\infty} b_n = c.$$

Proof.

The lengths $L_n = b_n - a_n$ of the intervals go to zero, that is,

$$\lim_{n\to\infty} L_n = 0.$$

The number c is contained within all the intervals, so also have $|c - a_n| \leq L_n$ and $|c - b_n| \leq L_n$.

That just means, by definition, that

$$\lim_{n\to\infty} a_n = c, \quad \lim_{n\to\infty} b_n = c.$$

This had to be shown. □

Analytical Background

The principle of nested intervals allows us to prove a few intuitive facts about continuous functions.

We start with a discussion of boundedness properties of continuous functions and their maxima/minima.

Lemma (Continuous Functions over Closed Intervals have upper bounds)

Let $f : [a, b] \to \mathbb{R}$ be continuous. Then there exists $M \in \mathbb{R}$ such that $f(x) \leq M$ for all $x \in [a, b]$.

Proof.

We prove the statement by contradiction: suppose that no such M exists. That means that f has no upper bound over $[a, b]$, that is, for all $M \in \mathbb{R}$ there exists $z \in [a, b]$ such that $f(z) > M$.

We set $a_0 = a$ and $b_0 = b$. So f has no upper bound over $[a_0, b_0]$. Define the midpoint $c_0 = \frac{1}{2}a_0 + \frac{1}{2}b_0$.

Then f has no upper bound over $[a_0, c_0]$ or $[c_0, b_0]$. If f has no upper bound over $[a_0, c_0]$, then we pick $a_1 = a_0$ and $b_1 = c_0$, and otherwise we pick $a_1 = c_0$ and $b_1 = b_0$.

In either case, note that

$$[a_1, b_1] \subseteq [a_0, b_0], \quad b_1 - a_1 = \frac{1}{2}(b_0 - a_0)$$

and that f has no upper bound over $[a_1, b_1]$.

Proof.

We repeat this construction to get a sequence of new intervals.

Suppose that a_n and b_n have been defined such that f has no upper bound over $[a_n, b_n]$. Define $c_n = \frac{1}{2}a_n + \frac{1}{2}b_n$. Then f has no upper bound over $[a_n, c_n]$ or $[c_n, b_n]$.

If f has no upper bound over $[a_n, c_n]$, then we pick $a_{n+1} = a_n$ and $b_{n+1} = c_n$. Otherwise, f has no upper bound over $[c_n, b_n]$, and we pick $a_{n+1} = c_n$ and $b_{n+1} = b_n$.

Note that

$$[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n], \quad b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n)$$

and f has no upper bound over $[a_{n+1}, b_{n+1}]$.

Proof.

We see that

$$(b_n - a_n) = \left(\frac{1}{2}\right)^n (b_0 - a_0)\,.$$

So we have succession over nested intervals $[a_n, b_n]$ such that

$$\lim_{n\to\infty} b_n - a_n = 0.$$

The principle of nested intervals shows that there exists $c \in [a, b]$ that is contained in all these intervals.

Proof.

Since f has no upper bound over all the intervals $[a_0, b_0], [a_1, b_1], \ldots$, we can pick a sequence $z_n \in [a_n, b_n]$ such that $f(z_n) > n$.

We observe that

$$\lim_{n \to \infty} z_n = c.$$

Since f is continuous, we must have

$$\lim_{n \to \infty} f(z_n) = f(c).$$

This contradicts the fact $f(z_n)$ increases to infinity.

We conclude that f must have an upper bound over the interval $[a, b]$, that is, M as in the statement of the lemma must exist. □

Not only do we have an upper bound, we even have a least upper bound.

Lemma

*Let $f : [a, b] \to \mathbb{R}$ be continuous. Then there exists a **least** upper bound $M \in \mathbb{R}$ such that $f(x) \leq M$ for all $x \in [a, b]$.*

A least upper bound is an upper bound such that all other upper bounds are larger.

Proof.

Let $M \in \mathbb{R}$ be an upper bound of f. This upper bound exists by the previous result. Pick any $c \in [a, b]$ be arbitrary and Define the set of values of f by

$$V := \{f(x) \mid x \in [a, b]\}.$$

Then the interval $[y_0, z_0]$ with $y_0 = f(c)$ and $z_0 = M$ has non-empty intersection with V, and its right endpoint is an upper bound for V.

Suppose we have an interval $[y_n, z_n]$ such that $[y_n, z_n] \cap V \neq \emptyset$ and z_n is an upper bound for the values in V. Define $w_n = \frac{1}{2}y_n + \frac{1}{2}z_n$.

If $[w_n, z_n] \cap V = \emptyset$, then we set $y_{n+1} = y_n$ and $z_{n+1} = w_n$. Otherwise, we set $y_{n+1} = w_n$ and $z_{n+1} = z_n$. In any case $[y_{n+1}, z_{n+1}]$ has non-empty intersection with V, and its right endpoint is an upper bound for V.

Thus we get a sequence of nested intervals $[y_n, z_n]$, each of which has non-empty intersection with V and whose right endpoints are an upper bound for V.

Proof.

By the principle of nested intervals, there exists a unique w contained in all the intervals.

The number w is an upper bound of V. To see this, recall that all z_n are upper bounds of V and that w is a limit of the z_n: hence for all $x \in [a, b]$, we have

$$0 \leq \lim_{n \to \infty} z_n - f(x) = w - f(x),$$

and thus $w \geq f(x)$. Furthermore, w is the least upper bound. If w' is another upper bound for V such that $w' < w$, then $w' < w \leq z_n$. Since $y_n \to w$, there exists a smallest $n \geq 1$ such that $w < y_n$. But then $[y_n, z_n] \cap V = \emptyset$, which contradicts our assumptions.

Hence all other upper bounds are larger than w. □

Not only do we have a least upper bound, that least upper is actually attained.

Lemma

Let $f : [a, b] \to \mathbb{R}$ be continuous. Then there exists $c \in [a, b]$ such that $f(x) \leq f(c)$ for all $x \in [a, b]$.

Proof.

Let $M \in \mathbb{R}$ be the least upper bound for the values of f over $[a, b]$. We write $a_0 = a$ and $b_0 = b$.

Suppose I have an interval $[a_n, b_n]$ such that M is the least upper bound of f over $[a_n, b_n]$. Then we get that for every $\epsilon > 0$ there exists $x \in [a_n, b_n]$ such that $f(x) > M - \epsilon$.

Set $c_n = \frac{1}{2}a_n + \frac{1}{2}b_n$. If M is the least upper bound of f over $[a_n, c_n]$, then we set $[a_{n+1}, b_{n+1}] = [a_n, c_n]$.

Otherwise, M must be the least upper bound of f over $[c_n, b_n]$, and then we set $[a_{n+1}, b_{n+1}] = [c_n, b_n]$.

By the principle of nested intervals, there exists a unique c contained all the intervals.

Proof.

By assumption, for every $n \geq 0$ there exists $x_n \in [a_n, b_n]$

$$M - 1/n \leq f(x_n) \leq M.$$

But $x_n \to c$, so it follows that

$$\lim_{n\to\infty}(M - 1/n) \leq \lim_{n\to\infty} f(x_n) \leq M,$$

that is

$$M = \lim_{n\to\infty} f(x_n) = f(c).$$

This had to be shown. □

Using the preceding results, we can prove the extreme value theorem for continuous functions over bounded closed intervals.

Lemma (Extreme Value Theorem)

Let $f : [a, b] \to \mathbb{R}$ be continuous. Then there exist $c, d \in [a, b]$ such that for all $x \in [a, b]$

$$f(c) \leq f(x) \leq f(d).$$

Analytical Background

If a function f over some interval into the real numbers is continuous, then all intermediate values between any two function values are attained by that function f somewhere. More formally, this is known as ***intermediate value theorem***.

Theorem (Intermediate Value Theorem)

Let $f : I \to \mathbb{R}$ be a continuous function over some interval $I \subseteq \mathbb{R}$. Let $a, b \in I$ with $a < b$ with $f(a) \leq f(b)$. Then for all $z \in [f(a), f(b)]$ there exists $c \in [a, b]$ such that $f(c) = z$.

Proof.

Let $a, b \in I$ with $a < b$ with $f(a) \leq f(b)$. Let $z \in [f(a), f(b)]$

We construct two sequences a_n and b_n such that the intervals $[a_n, b_n]$ get smaller and smaller while $z \in [f(a_n), f(b_n)]$ always holds.

We define $a_0 = a$ and $b_0 = b$ as starting values.

Now, assuming we have defined a_n and b_n for some n, let $c_n = \frac{1}{2}a_n + \frac{1}{2}b_n$ be the midpoint of the interval $[a_n, b_n]$. We make a case distinction depending on the value of $f(c_n)$:

- If $z < f(c_n)$, then we set $a_{n+1} = a_n$ and $b_{n+1} = c_n$.
- If $z \geq f(c_n)$, then we set $a_{n+1} = c_n$ and $b_{n+1} = b_n$.

In either of these two cases, we end up with $z \in [f(a_{n+1}), f(b_{n+1})]$.

Furthermore, the new interval $[a_{n+1}, b_{n+1}]$ is contained within the old interval $[a_n, b_n]$ and has half its length: $b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n)$.

Proof.

Repeating this construction, we get a sequence of intervals $[a_n, b_n]$ where each interval has half the length of its predecessor and always $z \in [f(a_n), f(b_n)]$.

By the principle of nested intervals, there exists a unique c contained all the intervals. We have

$$\lim_{n\to\infty} a_n = c, \quad \lim_{n\to\infty} b_n = c.$$

Since f is continuous, we get the following two limits:

$$\lim_{n\to\infty} f(a_n) = f(c) = \lim_{n\to\infty} f(b_n)$$

Since $f(a_n) \leq z$ is always true, we must have $f(c) \leq z$. And since $f(b_n) \geq z$ is always true, we must have $f(c) \geq z$. Thus $z = f(c)$ must be true. □

Analytical Background

Bisection Method

Bisection Method

We want to study the solution of general nonlinear equations

$$g(x) = y$$

where g is a continuous function.

Equivalently, we study how to find roots of functions. Namely, we define $f(x) = g(x) - y$, and want to solve

$$f(x) = 0.$$

Because of that argument, solving nonlinear equations is generally (re)formulated as solving $f(x) = 0$ for some x.

Bisection Method Algorithm

Suppose that we have interval $[a, b]$ and a continuous function $f : [a, b] \to \mathbb{R}$. Suppose that f has different signs on the endpoint of the interval,

$$\operatorname{sgn} f(a) \neq \operatorname{sgn} f(b).$$

If f is positive at one endpoint and negative at the other one, then the intermediate value theorem implies that there must exist $x^* \in [a, b]$ such that $f(x^*) = 0$.

We would like to have an algorithm that finds (approximately) such a root of f in the interval $[a, b]$.

One such method is the ***bisection method***.

Bisection Method Algorithm

Example (root of $f(x) = x^2 - 5$)

We review a classic method of computing square roots. Consider the function $f(x) = x^2 - 5$ and the interval $[1, 3]$. We have

$$f(1) = -4, \quad f(3) = 4.$$

So there is a root of f in $[1, 3]$. We see that $f(2) = -1$ at the midpoint of the interval. So there exists a root of f between $[2, 3]$, which is the “right half” of the original interval.

Bisection Method Algorithm

Example (root of $f(x) = x^2 - 5$)

Now consider the interval $[2, 3]$ with midpoint 2.5. We have

$$f(2) = -1, \quad f(2.5) = 1.25, \quad f(3) = 4.$$

This time, the endpoints of the left half interval have different signs. So we pick that as the interval to seek the root of f.

Bisection Method Algorithm

Example (root of $f(x) = x^2 - 5$)

Consider the interval $[2, 2.5]$ with midpoint 2.25. We have

$$f(2) = -1, \quad f(2.25) = 0.0625, \quad f(2.5) = 1.25.$$

Again, the endpoints of the left half interval have different signs. So we pick that as the interval to seek the root of f.

Bisection Method Algorithm

Example (root of $f(x) = x^2 - 5$)

Consider the interval $[2, 2.25]$ with midpoint 2.125. We have

$$f(2) = -1, \quad f(2.125) = -0.484375, \quad f(2.25) = 0.0625$$

Now, the endpoints of the right half interval have different signs. So we pick that as the interval to seek the root of f.

Bisection Method Algorithm

Example (root of $f(x) = x^2 - 5$)

Consider the interval $[2.125, 2.25]$ with midpoint 2.1875. We have

$$f(2.125) = -0.484375,$$
$$f(2.1875) = -0.21484375,$$
$$f(2.25) = 0.0625$$

Now, the endpoints of the right half interval have different signs. So we pick that as the interval to seek the root of f.

Bisection Method Algorithm

Example

We can proceed with this process over and over again. The center of the last interval $c = 2.1875$ is already somewhat of a guess for $\sqrt{5} = 2.236067...$, so clearly this method has some potential.

Bisection Method Algorithm

We formalize this method. We start with an initial interval $[a_0, b_0]$ with $a_0 < b_0$ and a continuous function $f : [a_0, b_0] \to \mathbb{R}$ such that

$$\operatorname{sgn} f(a_0) \neq \operatorname{sgn} f(b_0).$$

By the intermediate value theorem, f has a root in $[a_0, b_0]$.
We define the midpoint of the initial interval

$$c_0 := \frac{1}{2}(a_0 + b_0).$$

If $f(c_0) = 0$, then we are already done. Else, $f(c_0)$ is either positive or negative. Consequently we either have $\operatorname{sgn} f(a_0) \neq \operatorname{sgn} f(c_0)$ or $\operatorname{sgn} f(c_0) \neq \operatorname{sgn} f(b_0)$ but not both. In the first case, where $\operatorname{sgn} f(a_0) \neq \operatorname{sgn} f(c_0)$, we set

$$a_1 = a_0, \quad b_1 = c_0,$$

and in the second case, where $\operatorname{sgn} f(b_0) \neq \operatorname{sgn} f(c_0)$, we set

$$a_1 = c_0, \quad b_1 = b_0.$$

Bisection Method Algorithm

More generally, suppose that we have defined the k-th interval $[a_k, b_k]$ with $a_k < b_k$ such that

$$\operatorname{sgn} f(a_k) \neq \operatorname{sgn} f(b_k).$$

By the intermediate value theorem, f has a root in $[a_k, b_k]$. We define the midpoint of the k-th interval

$$c_k := \frac{1}{2}(a_k + b_k).$$

If $f(c_k) = 0$, then we are already done. Else, $f(c_k)$ is either positive or negative. Hence we either have $\operatorname{sgn} f(a_k) \neq \operatorname{sgn} f(c_k)$ or $\operatorname{sgn} f(c_k) \neq \operatorname{sgn} f(b_k)$ but not both. In the first case, where $\operatorname{sgn} f(a_k) \neq \operatorname{sgn} f(c_k)$, we set

$$a_{k+1} = a_k, \quad b_{k+1} = c_k,$$

and in the second case, where $\operatorname{sgn} f(b_k) \neq \operatorname{sgn} f(c_k)$, we set

$$a_{k+1} = c_k, \quad b_{k+1} = b_k.$$

Again, $a_{k+1} < b_{k+1}$, and by construction,

$$\operatorname{sgn} f(a_{k+1}) \neq \operatorname{sgn} f(b_{k+1}).$$

Bisection Method Algorithm

We can proceed in this manner iteratively to create a sequence of intervals

$$[a_0, b_0], \quad [a_1, b_1], \quad [a_2, b_2], \quad [a_3, b_3], \quad \ldots$$

with a sequence of midpoints

$$c_0, c_1, c_2, c_3, \ldots$$

In particular, these intervals are nested:

$$[a_{k+1}, b_{k+1}] \subset [a_k, b_k].$$

Furthermore, the length of the intervals goes to zero,

$$\lim_{k \to \infty} b_k - a_k = 0.$$

Indeed, by construction we have

$$b_{k+1} - a_{k+1} = \frac{1}{2} (b_k - a_k),$$

so the length of intervals halves at every iteration. By the principle of nested intervals, there exists a unique number x^* that is contained in all the intervals.

Bisection Method Algorithm

This number x^* is a good candidate for a root of the function. To rigorously verify that guess, we first discuss how this number can be approximated.

A good choice for approximations are sequences of numbers a_k, b_k, c_k.

For technical convenience, we introduce the notation

$$L_k = b_k - a_k$$

for the length of the k-th interval. Since the interval is halved at each step, we have

$$L_{k+1} = \frac{1}{2} L_k.$$

By recursion, we have the formula

$$L_k = 2^{-k} L_0.$$

Bisection Method Algorithm

Since x^* is in the interval $[a_k, b_k]$ and every point in that interval has distance at most L_k from the interval endpoints, for all k we have

$$|a_k - x^*| \leq 2^{-k} L_0, \quad |b_k - x^*| \leq 2^{-k} L_0, \quad |c_k - x^*| \leq 2^{-k} L_0.$$

We have

$$\lim_{k \to \infty} a_k = \lim_{k \to \infty} b_k = \lim_{k \to \infty} c_k = x^*.$$

The three sequences converge to the point x^*. Since f is continuous, we have

$$\lim_{k \to \infty} f(a_k) = \lim_{k \to \infty} f(b_k) = \lim_{k \to \infty} f(c_k) = f(x^*).$$

We see that $f(x^*)$ is the limit of a sequence of non-positive numbers and the limit of a sequence of non-negative numbers. Consequently, $f(x^*) = 0$.

Thus have shown that x^* is a root of f.

So the interval boundary points and midpoints produced by the bisection method converge to a root of the function f.

Bisection Method Algorithm

The above estimates formalize the intuition that the approximations will become better the more iterations we perform and the smaller the initial enclosing interval $[a_0, b_0]$.

We finish this discussion with a subtlety of the convergence estimates. We have

$$|a_k - x^*| \leq 2^{-k} L_0, \quad |b_k - x^*| \leq 2^{-k} L_0, \quad |c_k - x^*| \leq 2^{-k} L_0.$$

In fact, since every point in $[a_k, b_k]$ has distance at most $\frac{1}{2} L_k$ from the midpoint, we get slightly better estimate

$$|c_k - x^*| \leq \frac{1}{2} L_k = 2^{-k-1} L_0.$$

Thus, typically, c_k will be a better approximation than a_k or b_k.

Bisection Method Algorithm

Example

Let us consider another example:
$[a_0, b_0] = [0, 2]$ and $f(x) = x^2 - 2$. We can put the calculation results in the following table:

k	a_k	c_k	b_k	$f(a_k)$	$f(c_k)$	$f(b_k)$
0	0	1	2	-2	-1	2
1	1	1.5	2	-1	0.25	2
2	1	1.25	1.5	-1	-0.4375	0.25
3	1.25	1.345	1.5	-0.4375	-0.190975	0.25
4	1.345	1.4225	1.5	-0.190975	0.02350625	0.25

After four calculation steps, the midpoint begins resembling the exact solution $x^* = 1.414213562...$

Bisection Method Algorithm

A pseudocode for the bisection applied to some function f looks like

1: **BisectionMethod**
2: If $a > b$ then abort
3: $f_a := f(a), f_b := f(b)$
4: $k = 0$
5: **while** no termination criterion satisfied **do**
6: $\quad c = \frac{1}{2}a + \frac{1}{2}b$
7: $\quad f_c := f(c)$
8: $\quad$ If $\operatorname{sgn} f_a = \operatorname{sgn} f_b$ then abort
9: $\quad$ **if** $\operatorname{sgn} f_a = \operatorname{sgn} f_c$ **then**
10: $\quad\quad a = c$ and $f_a = f_c$
11: $\quad$ **else**
12: $\quad\quad b = c$ and $f_b = f_c$
13: $\quad$ **end if**
14: $\quad$ increment k
15: **end while**

Bisection Method Algorithm

In this context, the termination criterion can be a (combination of a) number of things:

- Maximal number of iterations: $k \leq k_{\max}$
- Maximal interval length: $b - a < \epsilon$
- Maximum value at midpoint $|f_c| < \epsilon$

The precise choice of termination criterion depends on practical considerations.