The Method of Bisection

Jonathan Lin

December 2, 2019

Abstract

I summarize a method called the "Bisection Argument", which can be used to prove several important results in real analysis.

1 The Nested Interval Theorem

All these results rely on the following theorem:

Theorem 1. *Suppose that $I = [a, b]$ is an interval. If we have a series of intervals I_n such that $I_0 = [a, b]$ and $I_n = [a_n, b_n]$ where we have the chain*

$$I_0 \supset I_1 \supset I_2 \supset \cdots .$$

Suppose also that

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

Then there is a unique point c contained in the I_n.

If we let go of the requirement that $\lim_{n\to\infty} a_n - b_n = 0$, then the point c still exists, but it's not necessarily unique (in fact, there will be uncountably many c.

Proof. The sequence $\{a_n\}$ is a monotonically increasing sequence which is bounded above by any b_n. Similarly, the sequence $\{b_n\}$ is a monotonically decreasing sequence which is bounded below by any a_n. It follows both limits exist, and we have

$$\lim_{n\to\infty} a_n = \sup(A)$$

and

$$\lim_{n\to\infty} b_n = \inf(B).$$

Since both limits are equal (by hypothesis) we have that their limit is the unique point contained in all the I_n (follows from consequences of the least upper bound). □

2 Bisection of an Interval

Here is the approach: Suppose we want to establish the existence of a point in a closed interval $[a, b]$ with property P. A priori we know this proposition to be true. We would like to “seek out” the point in question. What we might do is divide our interval in half. Then we might be able to deduce that in all cases the property P holds for either $\left[a, \frac{a+b}{2}\right]$ or $\left[\frac{a+b}{2}, b\right]$. If we iteratively do this, we end up with intervals which are contained in each other with length $2^{-n}(b-a)$. After this we will be able to apply the nested interval theorem to this sequence of intervals. We summarize these above remarks by defining a **bisection** below.

Definition 1. Let $I = [a, b]$ and let P be some property which holds for I, and also has the property that it holds for $\left[a, \frac{a+b}{2}\right]$ or $\left[\frac{a+b}{2}, b\right]$. We define $I_0 = 0$ and define a sequence I_n inductively as follows: given $I_n = [a_n, b_n]$ for which the property P holds, then P holds for either $\left[a, \frac{a_n+b_n}{2}\right]$ or $\left[\frac{a_n+b_n}{2}, b\right]$ in which case we define I_{n+1} to be the interval for which the property P holds. We call this sequence a **bisection** of I.

Note that any bisection of I has the following properties: We have the chain

$$I = I_0 \supset I_1 \supset \cdots \supset I_n \supset \cdots$$

and if $I_n = [a_n, b_n]$ we have that $b_n - a_n = (b-a)2^{-n}$, so we have that

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

It follows that the nested interval theorem can be applied to the bisection of I to get a point c contained in all the I_n.

3 The Intermediate Value Theorem

Theorem 2. *Suppose f is continuous on an interval $[a, b]$. Also, suppose that $f(a) \leq 0 \leq f(b)$. It follows that there exists a point c in $[a, b]$ such that $f(c) = 0$.*

Proof. Let P be the property on an interval $[a, b]$ such that f is continuous on that interval and $f(a) \leq 0 \leq f(b)$. The point $(a+b)/2$ is either non-negative or non-positive. This means that P holds on one of the half-intervals of $[a, b]$. Inductively, this holds as well. Hence we can construct a bisection $\{I_n\}$ where P holds all the I_n. By the nested interval theorem there is a unique point c contained in the I_n.

We claim that $f(c) = 0$. We can see this as follows: Since c is contained in the I_n we have that $f(a_n) \leq 0$ for all a_n and $f(b_n) \geq 0$ for all b_n. Since f is continuous, we have that $f(a_n) \to f(c)$ and $f(b_n) \to f(c)$. Combining these facts, we have that $f(c) \leq 0$ and $f(c) \geq 0$, implying that $f(c) = 0$. $\square$

4 The Extreme Value Theorem

Theorem 3. *If f is continuous on $[a, b]$ then f is bounded above on $[a, b]$. Moreover, f attains its supremum, so that f has a maximum value.*

Proof. Suppose that f is unbounded on $[a, b]$. Then f is unbounded on at least one of its half-intervals. We define the bisection $\{I_n\}$ with the property that I_n is unbounded. By the nexted interval theorem, we have that

$$\bigcap_{i=1}^{\infty} I_n = \{c\}$$

for $c \in [a, b]$. But since f is continuous at c, there exists a $\delta > 0$ such that $|x - c| < \delta \implies |f(x) - f(c)| < 1$. So f is bounded on $(c - \delta, c + \delta)$. HOwever, there exists $N \in \mathbb{N}$ such that $I_N \subset (c - \delta, c + \delta)$, which is a contradiction because I_N is an unbounded set.

Clearly then we have that $S = \{f(x) \mid x \in [a, b]\}$ is bounded and non-empty. Hence $\alpha = \sup(S)$ exists. Clearly either one of the half-intervals of $[a, b]$ must have supremum α as well. We define the bisection $\{I_n\}$ with the property that $sup(f(I_n)) = \alpha$. By the nested interval theorem, we have that there is a unique common point ℓ in the I_n. We claim that $f(\ell) = \alpha$. This is because if $f(\ell) \neq \alpha$, then if $\varepsilon = \frac{1}{2}(\alpha - f(\ell))$ there would exist a $\delta > 0$ such that $|x - \ell| < \delta \implies |f(x) - f(\ell)| < \varepsilon$. However, there is an interval I_n inside $(x - \delta, x + \delta)$, which is a contradiction since that would mean there is a value less than α which is an upper bound for I_n. □

5 An application to Uncountability

We will show that the interval $[0, 1]$ is uncountable. This argument is identical to the Cantor diagonalization argument, but the argument is made using the nested interval theorem and repeated bisections. First, note that our bisection will be defined as a trisection of the first and last third intervals. This is so we can use the nested interval theorem and there are no intersections.

Theorem 4. *The interval* $[0, 1]$ *is uncountable. That is, there is no bijection from* $[0, 1]$ *to the natural numbers* $\mathbb{N}$.

Proof. Suppose that $[0, 1]$ is countable. That is, there was some list $\{x_n\}$ of all natural numbers under some ordering. Define the bisection I_n to be the lower third of I_n if x_n is not in that interval and the upper third of I_n otherwise.

By the nested interval theorem, the I_n contain a unique point x. However, x is contained in the I_n, but none of the x_i are contained in all of the I_n. In particular, $x_j \notin I_j$ for all $j \in \mathbb{N}$. This is a contradiction, so it follows that $[0, 1]$ is uncountable. □

6 The Rationals are Dense

Here we show the rational numbers are dense. Actually we show a stronger statement that the dyadic rational numbers of the form $k/2^n$ are dense in $\mathbb{R}$.

How do we show this? We want to show that given any basic open neighborhood $(x - \varepsilon, x + \varepsilon)$ there exists some dyadic rational in this point. Assume x is not itself a dyadic rational, else x is a desired rational number of that form. Otherwise, there exists an integer

k such that $k < x < k+1$. Let $\{I_n\}_{n=0}^{\infty}$ be the bisection of $[k, k+1]$ where the I_n contain x. Clearly such I_n are unambiguous because x is not itself a dyadic rational and the endpoints of I_n are. By the nested interval the intersection of the I_n is a singular point. This point is x by construction. If we let $I_n = [a_n, b_n]$, we have two sequences of **dyadic** rational numbers $\{a_n\}$ and $\{b_n\}$ which converge to x. It follows by definition of convergence of sequences that $(x - \varepsilon, x + \varepsilon)$ contains some a_j so are are done.