6 Continuity, Part II (On an Interval)

6.1 Intermediate Value Theorem

Theorem: if f is continuous on $[a, b]$ and $f(a) < 0 < f(b)$, then there exists $x \in [a, b]$ such that $f(x) = 0$.

Or intuitively, if $f(a)$ is below zero and $f(b)$ is above zero, f must cross the x-axis somewhere.

Proof: intuitively, we will locate the smallest number x on the x-axis where $f(x)$ first crosses from negative to positive, and show that $f(x)$ must be zero.

First, we define a set A that contains all inputs to f before f crosses from negative to positive for the first time:

$$A = \{x : a \leq x \leq b, \text{and } f \text{ is negative on the interval } [a, x]\}$$

We know $A \neq \emptyset$ since $a \in A$, and b is an upper bound of A. Thus A has a least upper bound α such that $a \leq \alpha \leq b$. By nonzero neighborhood lemma (see 4.1) we know there is some interval around a on which f is negative, and some interval around b on which f is positive. Thus we can further refine the bound on α to $a < \alpha < b$.

We now show $f(\alpha) = 0$ by eliminating the possibilities $f(\alpha) < 0$ and $f(\alpha) > 0$.

Case 1. Suppose for contradiction $f(\alpha) < 0$. By nonzero neighborhood lemma there exists $\delta > 0$ such $|x - \alpha| < \delta$ implies $f(x) < 0$ for all x. But that means numbers in $(\alpha - \delta, \alpha + \delta)$ are in A. E.g. $(\alpha + \delta/2) \in A$. Since $\alpha + \delta/2 > \alpha$, α is not an upper bound of A, and is thus not the least upper bound.

Case 2. Suppose for contradiction $f(\alpha) > 0$. By nonzero neighborhood lemma there exists $\delta > 0$ such $|x - \alpha| < \delta$ implies $f(x) > 0$ for all x. But that means numbers in $(\alpha - \delta, \alpha + \delta)$ are *not* in A, and there exist many upper bounds of A less than α. E.g. $\alpha - \delta/2$ is an upper bound of A, and since $\alpha - \delta/2 < \alpha$, α is not the *least* upper bound.

Both cases lead to contradiction, therefore $f(\alpha) = 0$. QED.

IVT generalization

The intermediate value theorem is usually presented in a more general way. If f is continuous on $[a, b]$ and $f(a) < c < f(b)$ or $f(a) > c > f(b)$ then there is some x in $[a, b]$ such that $f(x) = c$.

Intuitively, f takes on any value between $f(a)$ and $f(b)$ at some point in the interval $[a, b]$.

Proof. This trivially follows from the the theorem as initially stated. There are two cases:

Case 1: $f(a) < c < f(b)$. Let $g = f - c$. Then g is continuous and $g(a) < 0 < g(b)$. Thus there is some x in $[a, b]$ such that $g(x) = 0$. But that means $f(x) = c$.

Case 2: $f(a) > c > f(b)$. Observe that $-f$ is continuous on $[a, b]$ and $-f(a) < -c < -f(b)$. By case 1 there is some x in $[a, b]$ such that $-f(x) = -c$, which means $f(x) = c$.

QED.

6.2 Boundedness theorem

The boundedness theorem states that if f is continuous on $[a, b]$, then f is bounded above (i.e. f lies below some line). Before we prove this, we first prove a simple lemma.

Bounded neighborhood lemma: if f is continuous at a, then there is $\delta > 0$ such that f is bounded above on the interval $(a - \delta, a + \delta)$.

Intuitively, if f is continuous at a then there is some interval around a on which f is bounded above.

Proof: The proof is trivial. Inlining the definition of continuity, for any $\epsilon > 0$ there exists $\delta > 0$ such that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$ for all x. Thus $f(a) + \epsilon$ is the upper bound on f within $(a - \delta, a + \delta)$, as desired.

(Note that we can pick any ϵ to concretize the proof, for example $\epsilon = 1$.)

Boundedness theorem: if f is continuous on $[a, b]$, then f is bounded above on $[a, b]$. I.e. there is some numbers N such that $f(x) \leq N$ for all x in $[a, b]$.

Proof: intuitively, we will try to find the smallest number x on the x-axis where $f(x)$ becomes unbounded above, and discover that there is no such number in $[a, b]$.

First, we define a set A that contains all inputs to f before f stops being bounded above:

$$A = \{x : a \leq x \leq b, \text{ and } f \text{ is bounded above on } [a, x]\}$$

By bounded neighborhood lemma f is bounded above in the neighborhood of a[9]. Thus we know $A \neq \emptyset$ because $a \in A$. Further, b is an upper bound of A. Thus A has a least upper bound.

[9]We are being sloppy here as we actually need a left-sided and right-sided version of the bounded neighborhood lemma. I am papering over this for now, but will need to fix at some point by giving proper one sided proofs

Let $\alpha = \sup A$. To prove the boundedness theorem we must prove two claims:

1. $\alpha = b$, i.e. f does not ever stop being bounded above before b.
2. $(\alpha = b) \in A$, as $\sup A$ is not necessarily a member of A.

First, we prove $\alpha = b$. Suppose for contradiction $\alpha < b$. By bounded neighborhood lemma there is some $\delta > 0$ such that f is bounded above in $(\alpha - \delta, \alpha + \delta)$. But that means there are many upper bounds greater than α, for example $\alpha + \delta/2$. Thus α is not the *least* upper bound. We have a contradiction, and so $\alpha = b$.

Second, we prove $(\alpha = b) \in A$. By bounded neighborhood lemma there is some $\delta > 0$ such that f is bounded above in $(b - \delta, b]$. Pick any x_0 such that $b - \delta < x_0 < b$. Then:

- $x_0 < b = \alpha$. Since α is the least upper bound it follows $x_0 \in A$. Thus f is bounded above on $[a, x_0]$.
- f is bounded above on $[x_0, b]$.

Since f is bounded above on $[a, x_0]$ and on $[x_0, b]$, it follows f is bounded above on $[a, b]$ as desired. QED.

Boundedness theorem generalization

The *boundedness theorem* is usually presented slightly more generally: it proves f is bounded above *and* below. We already proved the former. Put more formally, the latter states:

If f is continuous on $[a, b]$, then f is bounded *below* on $[a, b]$. I.e. there is some numbers N such that $f(x) \geq N$ for all x in $[a, b]$.

Proof: observe that $-f$ is continuous on $[a, b]$. By claim 2 there exists a number M such that $-f(x) \leq M$ for all x in $[a, b]$. But that means $f(x) \geq -M$ for all x in $[a, b]$. QED.

6.3 Extreme Value Theorem

The *extreme value theorem* states that is f is continuous on $[a, b]$, then f attains its maximum on $[a, b]$. To see why we need the extreme value theorem, consider $f = \frac{1}{x}$. f is discontinuous at 0 and approaches infinity. Thus f does not attain a maximum value on the interval $[0, 1]$.

Extreme value theorem: If f is continuous on $[a, b]$, then there is some number y in $[a, b]$ such that $f(y) \geq f(x)$ for all x in $[a, b]$.

Proof: Let A be the set of f's outputs on $[a, b]$:

$$A = \{f(x) : x \text{ in } [a, b]\}$$

Since $[a, b]$ isn't empty, $A \neq \emptyset$. By boundedness theorem, f is bounded on $[a, b]$, and so A has an upper bound. Thus A has a least upper bound. Let $\alpha = \sup A$. By definition $\alpha \geq f(x)$ for x in $[a, b]$. Thus it suffices to show $\alpha \in A$ (i.e. $\alpha = f(y)$ for some y in $[a, b]$).

Let's consider a function g[10]:

$$g = \frac{1}{\alpha - f(x)}, \quad x \text{ in } \; [a, b]$$

Suppose for contradiction $\alpha \notin A$. Then the denominator is never zero and g is continuous. Therefore:

$$\begin{aligned}
&\frac{1}{\alpha - f(x)} < M && \text{by boundedness theorem} \\
& && \text{for some bound } M \\
&\implies \alpha - f(x) > \frac{1}{M} && \text{take reciprocal} \\
&\implies -f(x) > \frac{1}{M} - \alpha && \\
&\implies f(x) < \alpha - \frac{1}{M} && \text{times } -1
\end{aligned}$$

But this contradicts that α is the *least* upper bound. Thus $\alpha \in A$ as desired. QED.

EVT generalization

The extreme value theorem is usually presented slightly more generally: a continuous f attains both its maximum and its minimum. We already proved the former. Put more formally, the latter states:

If f is continuous on $[a, b]$, then there is some number y in $[a, b]$ such that $f(y) \leq f(x)$ for all x in $[a, b]$.

Proof: Observe that $-f$ is continous on $[a, b]$. By claim 3 there is some y in $[a, b]$ such that $-f(y) \geq -f(x)$ for all x in $[a, b]$. But that means that $f(y) \leq f(x)$ for all x in $[a, b]$. QED.

6.4 IVT and EVT consequences

Claim 1a: Every positive number has a square root. I.e. if $\alpha > 0$, then there is some number x such that $x^2 = \alpha$.

Proof: Consider the function $f(x) = x^2$. If f takes on the value of α as its output, then $x = \sqrt{\alpha}$ is the input (i.e. $x^2 = \alpha$). Thus all we must show is that f takes on the value of α.

[10] g is a bit of a rabbit pulled out of a magic hat, but to quote a great British statesman, them's the breaks

We can do it as follows. Show there exist a, b such that $f(a) < \alpha < f(b)$. Since f is continuous, by intermediate value theorem there exists x such that $f(x) = \alpha$. So, let's find a and b:

- First, find a such that $f(a) < \alpha$. Observe that $f(0) = 0 < \alpha$, thus fix $a = 0$.
- Second, find b such that $\alpha < f(b)$.
  - If $\alpha < 1$ then $f(1) = 1 > \alpha$. Thus fix $b = 1$.
  - If $\alpha > 1$ then $f(\alpha) = \alpha^2 > \alpha$. Thus fix $b = \alpha$.

By intermediate value theorem, there is some x in $[0, b]$ such that $f(x) = \alpha$. QED.

Claim 1b: Every positive number has an nth root. I.e. if $\alpha > 0$, then there is some number x such that $x^n = \alpha$.

Proof: We can use the exact same argument as 1a, just consider $f(x) = x^n$.

Claim 1c: Let n be odd. Then every number has an nth root. I.e. there is some number x such that $x^n = \alpha$ for all α.

Proof: This is also easy:

- *Case* $\alpha > 0$. By claim 2b, there is an x such that $x^n = \alpha$.
- *Case* $\alpha < 0$. By claim 2b, there is an x such that $x^n = -\alpha$. Then $(-x)^n = \alpha$.

QED.

Claim 2: If n is odd, then any equation of the form

$$x^n + a_{n-1}x^{n-1} + \ldots + a_0 = 0$$

has a root.

Proof: Let $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$. Here is an intuitive outline of the proof:

1. We will show that f must take on negative and positive values. Thus by the intermediate value theorem, there exists some x such that $f(x) = 0$.
2. To do that we will show that as $|x|$ gets large, x^n completely dominates other terms. (This is obvious if you consider Big-Oh of each term.)
3. Since n is odd, x^n takes on a negative value when x is negative, and a positive value when x is positive. And since x^n dominates other terms, when x is sufficiently large, f takes on both negative and positive values.

We must find a way to bound the magnitude of $a_{n-1}x^{n-1} + \ldots + a_0$ to show that for large enough x, it's smaller than the magnitude of x^n. This way we guarantee $f(x)$ has the same sign as x^n. This is trivial to do by adopting Big-Oh notation, but both math books I looked at do it the old-fashioned way, so we will too.

Let's start with some obvious transformations we can make:

$$\begin{aligned} |a_{n-1}x^{n-1} + \ldots + a_0| &\le |a_{n-1}x^{n-1}| + \ldots + |a_0| && \text{by triangle inequality} \\ &= |a_{n-1}||x^{n-1}| + \ldots + |a_0| && \text{in general } |ab| = |a||b| \end{aligned}$$

If we only consider behavior of f on large x (i.e. when $|x| > 1$), we can further bound the expression. Observe that when $|x| > 1$ then $x^{n-1} > x^{n-2} > \ldots > x > 1$. Therefore:

$$\begin{aligned} |a_{n-1}x^{n-1} + \ldots + a_0| &\le |a_{n-1}||x^{n-1}| + \ldots + |a_0| \\ &\le |a_{n-1}||x^{n-1}| + \ldots + |a_0||x^{n-1}| \\ &= x^{n-1}(|a_{n-1}| + \ldots + |a_0|) \end{aligned}$$

Let $M = |a_{n-1}| + \ldots + |a_0| + 1$, i.e. a bound on the sum of the coefficients, plus a little extra to ensure $M > 1$. Then

$$\begin{aligned} |a_{n-1}x^{n-1} + \ldots + a_0| &\le x^{n-1}(|a_{n-1}| + \ldots + |a_0|) \\ &< M|x^{n-1}| \end{aligned}$$

Given this bound it follows that for all $|x| > 1$:

$$x^n - M|x^{n-1}| < x^n + (a_{n-1}x^{n-1} + \ldots + a_0) < x^n + M|x^{n-1}|$$

or put differently:

$$x^n - M|x^{n-1}| < f(x) < x^n + M|x^{n-1}|$$

We will now find x_1 and x_2 such that $f(x_1) < 0$ and $f(x_2) > 0$. Let $x_1 = -2M$ (note that x_1 satisfies our condition $|x_1| > 1$ since $M > 1$). Then for all $x \le x_1$:

$$\begin{aligned} f(x) &< x^n + M|x^{n-1}| \\ &= x^n + Mx^{n-1} && n \text{ is odd, thus } n-1 \text{ is even, thus } x^{n-1} > 0 \\ &= x^{n-1}(x + M) && \text{factor out } x^{n-1} \\ &\le -2^{n-1}M^n && \text{substitute } -2M \text{ and simplify} \\ &< 0 \end{aligned}$$

Similarly, let $x_2 = 2M$. Then for all $x \geq x_2$:

$$\begin{aligned} f(x) &> x^n - M|x^{n-1}| \\ &= x^n - Mx^{n-1} \\ &= x^{n-1}(x - M) \\ &\geq 2^{n-1}M^n \\ &> 0 \end{aligned}$$

QED.

Claim 3: If n is even and $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$, then there is a number y such that $f(y) \leq f(x)$ for all x.

Intuitively, even degree polynomials achieve their minimum on $\mathcal{R}$ because when you zoom out enough they are U-shaped (consider the graph $f(x) = x^2$ as a simple example).

Proof: It's easy to intuitively see why the claim makes sense. x^n dominates the rest of the terms when x is very large. Since n is even, $x^n > 0$. Thus on very large $|x|$ the graph shoots up (i.e. it has a U shape).

Here is the outline for our proof:

1. Observe that $f(0) = a_0$.
2. We will prove f is U-shaped by proving there exist two points:
   - $x_0 < 0$ such that $f(x) > a_0$ on $(-\infty, x_0]$.
   - $x_1 > 0$ such that $f(x) > a_0$ on $[x_1, \infty]$.
3. By extreme value theorem f achieves a minimum m on $[x_0, x_1]$. Note $m \leq a_0$ (otherwise it wouldn't be a minimum).
4. Thus f achieves a minimum m on $\mathcal{R}$, as we've shown that outside $[x_0, x_1]$, $f(x) > a_0$ (and thus $f(x) > m$) for all x.

All we must do now to complete the proof is find $x_0 < 0 < x_1$. Let $M = |a_{n-1}| + \ldots + |a_0| + 1$, i.e. a bound on the sum of the coefficients, plus a little extra to ensure $M > 1$. In Claim 2 we discovered that for $|x| > 1$

$$x^n - M|x^{n-1}| < f(x) < x^n + M|x^{n-1}|$$

Let $x_1 = -2M$. Note that x_1 satisfies our condition $|x_1| > 1$ since $M > 1$. Then for all $x < x_1$:

$$\begin{aligned} f(x) &> x^n - M|x^{n-1}| \\ &= x^n + Mx^{n-1} && x \text{ is negative, and } n-1 \text{ is odd} \\ &= x^{n-1}(x + M) \\ &\geq 2^{n-1}M^n && \text{substitute } -2M \text{ and simplify} \end{aligned}$$

Similarly let $x_2 = 2M$. Then for all $x > x_1$:

$$
\begin{aligned}
f(x) &> x^n - M|x^{n-1}| \\
&= x^n + Mx^{n-1} && x \text{ is positive} \\
&= x^{n-1}(x + M) \\
&\geq 2^{n-1}M^n && \text{substitute } 2M \text{ and simplify}
\end{aligned}
$$

Since $M > 1$ we have

$$2^{n-1}M^n \geq M \geq |a_n + 1| \geq a_n + 1 > a_n$$

Therefore for all $x < x_1$ and $x > x_2$, $f(x) > a_n$ as desired.

Claim 4: Consider the equation

$$x^n + a_{n-1}x^{n-1} + \ldots + a_0 = c$$

and suppose n is even. Then there is a number m such that the equation has a solution for $c \geq m$ and has no solution for $c < m$.

Proof: In claim 3 we saw that even degree polynomials achieve a minimum. Let that be m. There are three cases:

- If $c < m$ there is no solution, as the polynomial doesn't take on values less than m.
- If $c = m$ there is a solution, as the polynomial obviously takes on the value m (by claim 3).
- Suppose $c > m$. Let $y, z \in \mathcal{R}$ such that $f(y) = m$ and $z > y, f(z) > c$.[11] Then $f(y) = m < c < f(z)$. By intermediate value theorem there is a number k in $[y, z]$ such that $f(k) = c$.

QED.

6.5 Uniform continuity

TODO (skipping until it comes up in Spivak or I hit it in Galvin's notes)

[11] Technically we have to prove such a z exists, but somehow Spivak rolls right past this.