Lecture 1: The IVT and the Itinerary Lemma

(Source materials: "Period three implies chaos," by Li and Yorke, and "From Intermediate Value Theorem To Chaos," by Huang.)

This lecture, roughly speaking, is about how the intermediate value theorem is a deeply strange and powerful piece of mathematics.

On its first glance, it looks pretty innocuous. Here's the theorem statement, as you've probably seen it in calculus:

Theorem. (Intermediate Value Theorem.) Suppose that f is a continuous function on some interval $[a, b]$, and L is a value between $f(a)$ and $f(b)$. Then there is some value $x \in [a, b]$ such that $f(x) = L$.

On its face, this looks pretty normal, and quite believable: if a continuous function starts at $f(a)$ and ends up at $f(b)$, then it must adopt every value between $f(a)$ and $f(b)$ along the way. Despite its simplicity, the intermediate value theorem has a lot of useful, obvious, and not-entirely-obvious applications:

1. Suppose that you are running a race and are in last place. If you finish in first place, then at some point in time you must have passed the other runners. To make this an intermediate value theorem problem: for each other runner i, let $f_i(t)$ denote the signed distance between you and that runner. At the point in time in which you are in last, the function $f_i(t)$ is negative; at the point in time when you finished the race, $f_i(t)$ is positive. Because f is continuous[1], then there must be some time t where $f_i(t) = 0$, at which point you pass that runner.

2. Suppose that $p(x)$ is a polynomial of odd degree: i.e. that there are coefficients $a_0, \ldots a_n$ such that $p(x) = a_0 + \ldots + a_n x^n$, with n odd and $a_n \neq 0$. Then $p(x)$ has a root: i.e. there is some value x_0 such that $p(x_0) = 0$. This is because for sufficiently large values of x, $p(x)$ will be dominated by its $a_n x^n$ term, and thus become whichever sign a_n is. Therefore, for sufficiently large values of x, $p(x)$ and $p(-x)$ are different signs! So we can apply the intermediate value theorem and choose L to be 0, which gives us that there is some value at which $f(x_0) = 0$.

3. Suppose that $f(x)$ is a continuous function on $[a, b]$, whose range contains the interval $[a, b]$. Then there is some point $x_0 \in [a, b]$ such that $f(x_0) = x_0$: i.e. there is a point in our interval that our function does not change.

   This is not hard to see. Because $[a, b]$ is within the range of $f(x)$, there are two values $c, d \in [0, 1]$ such that $f(c) = 0, f(d) = 1$. If $c = a$ or $d = b$, we've found our point! Otherwise: look at the function $g(x) = f(x) - x$. At c, we have $g(c) = a - c < 0$, because c is a number in $[a, b]$ not equal to a. At d, we have $g(d) = b - d > 0$, because

[1] Assuming that you're not cheating or (less likely) quantum-tunneling during said race.

d is a positive number not equal to b in $[a, b]$. Therefore, by the intermediate value theorem, there is some point x_0 between c and d such that $g(x_0) = 0$. But this means that $f(x_0) - x_0 = 0$; i.e. $f(x_0) = x_0$, and we have our result!

For an example, let's consider the function

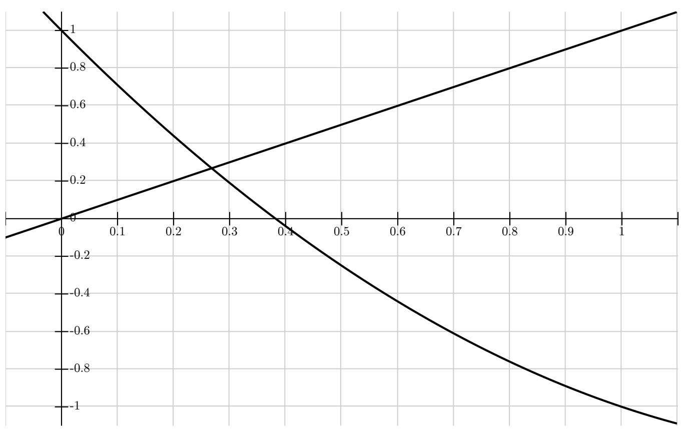
$$f(x) = x^2 - 3x + 1$$

on the interval $[0, 1]$. What is the range of this function? Well: its derivative $f'(x)$ is just

$$2x - 3,$$

which is negative on the entire interval $[0, 1]$. So our function is decreasing over all of $[0, 1]$: therefore its maximum occurs at the leftmost point of the interval $[0, 1]$, $x = 0$, where our function is 1, and its minimum occurs at the rightmost point of the interval $[0, 1]$, $x = 1$, where our function is $1 - 3 + 1 = -1$. Because our function is continuous, the intermediate value theorem tells us we adopt every value between this maximum and the minimum: i.e. that our function has range $[-1, 1]$. This contains $[0, 1]$, which means that 3 above applies, and there is some point x_0 such that $f(x_0) = x_0$!

This is visually obvious: if we graph $f(x) = y$ along with the relation $y = x$, we can see that there is clearly a point of intersection in the interval $[0, 1]$:



This intersection is precisely where we have $f(x) = x$! Algebraically, we can solve for this point:

$$f(x) = x \Leftrightarrow x^2 - 3x + 1 = x \Leftrightarrow x^2 - 4x + 1 = 0,$$

which holds for $x = \frac{4 \pm \sqrt{16-4}}{2} = 2 \pm \sqrt{3}$. $x = 2 - \sqrt{3}$ is in our interval, and therefore is the point we're looking for! I.e.

$$f(2 - \sqrt{3}) = (2 - \sqrt{3})^2 - 3(2 - \sqrt{3}) + 1 = 7 - 4\sqrt{3} - 6 + 3\sqrt{3} + 1 = 2 - \sqrt{3},$$

as claimed.

These points are particularly special objects in mathematics, and as such we should give them a name:

Definition. A point x_0 is called a **fixed point** of a function $f(x)$ if $f(x_0) = x_0$. (This name reflects the fact that this point is "fixed" under the mapping $f(x)$.)

We generalize these objects in the next section:

1 Periodic Points

Definition. Let $f(x)$ be some function. We say that a point x_0 in the domain of f is a **periodic point** with **period** n if the following two conditions hold:

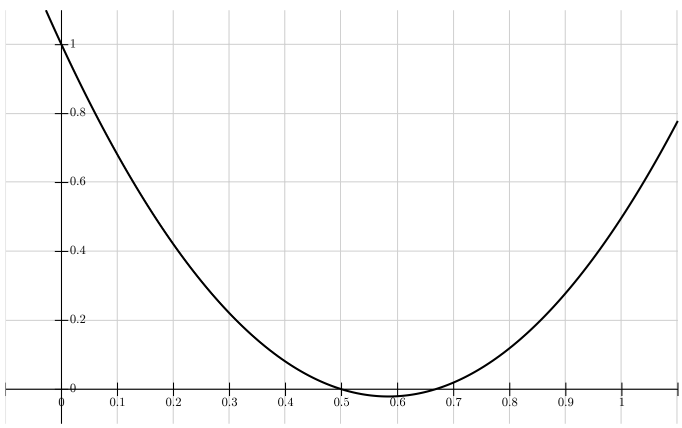
1. $f^n(x_0)$, the result of applying the function f n times in a row to x_0 (i.e. $\overbrace{f(f(\ldots f}^{\text{n times}}(x_0)\ldots))$, is equal to x_0.
2. For any k, $1 \leq k \leq n-1$, $f^k(x_0) \neq x_0$.

In other words, a point has period n if applying f to that point n times returns that point to itself, and n is the smallest value for which this point returns to itself.

In this sense, the **fixed points** we studied earlier were just points with period 1.

Finding examples of other such points is a bit trickier, but not too hard! Consider

$$p(x) = 3x^2 - \frac{7}{2}x + 1.$$



Notice that

- $p(0) = 1$,
- $p(1) = 1/2$, and
- $p(1/2) = 0$;

therefore, 0 is a point of period 3.

Determining whether this function has points with other periods, though: like points with period 5, or 7, or 6 ... seems hard. How can we do this? Well: the intermediate value theorem gave us a way to find fixed points. Perhaps we can build something out of the intermediate value theorem that can find periodic points!

As it turns out, we can do this via the following theorem:

Theorem. (The Itinerary Lemma.) Let $f(x)$ be a continuous function on the interval $[a, b]$, and $I_0, \ldots I_{n-1}$ denote a collection of closed intervals that are each contained within $[a, b]$. Assume that

1. $f(I_k) \supseteq I_{k+1}$, for every $k = 0 \dots n-2$, and
2. $f(I_{n-1}) \supseteq I_0$,

where by $f(I_k)$ we mean the set given by applying f to all of the points in the interval I_k. (In other words, applying f to any one interval I_k gives you a set that contains the next interval I_{k+1})

Then there is some point $x_0 \in I_0$ such that

1. $f^n(x_0) = x_0$, and
2. $f^k(x_0) \in I_k$, for every $k = 0, \dots n-1$.

Note that if we can make all of the I_k's for $k \geq 1$ not contain points in I_0, then any solution of the above is a point with period n, because each $f^k(x_0)$ will be contained in I_k, and therefore not a point in I_0 (and in particular not equal to x_0 itself!)

We prove this theorem here:

Proof. We start by observing the following useful fact:

Lemma 1. *If $f(I_k) \supseteq I_{k+1}$, then there is a subinterval of I_k such that $f(I_k) = I_{k+1}$.*

Proof. This is a consequence of the intermediate value theorem. Suppose that $I_{k+1} = [c, d]$, for some pair of endpoints c, d. Because $f(I_k) \supseteq [c, d]$, there are values that get mapped to c and d themselves. Pick x_1, x_2 such that $f(x_1) = c, f(x_2) = d$, and x_1, x_2 are the closest two such points with this property. (Question you should answer for yourself: why is this possible?)

Claim: this means that $f([x_1, x_2]) = [c, d]$. To see why, simply use the intermediate value theorem to see that $f([x_1, x_2])$ contains $[c, d]$. Moreover, if it contained a point $z \notin [c, d]$, then (if x_3 maps to z) the intermediate value theorem would tell us that we can find a point that maps to one of c, d in one of the intervals $[x_1, x_3], [x_3, x_2]$, in such a way that violates our "closest two points" property! So we've proven our lemma.

$\square$

Given this lemma, our proof is relatively simple. Repeatedly use the lemma above to construct intervals I_k^* as follows:

- First, find $I_{n-1}^* \subseteq I_{n-1}$ such that $f(I_{n-1}^*) = I_0$.
- Then, find $I_{n-2}^* \subseteq I_{n-2}$ such that $f(I_{n-2}^*) = I_{n-1}^*$.
- Then, find $I_{n-3}^* \subseteq I_{n-3}$ such that $f(I_{n-3}^*) = I_{n-2}^*$.
- ...
- Then, find $I_1^* \subseteq I_1$ such that $f(I_1^*) = I_2^*$.
- Finally, find $I_0^* \subseteq I_0$ such that $f(I_0^*) = I_1^*$.

Then, as a consequence, we must have that for any $k = 0, 1, \ldots n-2$,

$$f^k(I_0^*) = I_k^*.$$

In particular, this tells us that $f^n(I_0^*) = f^{n-1}(I_1^*) = \ldots = f(I_{n-1}^*) = I_0 \supseteq I_0^*$. In other words, f^n is a function whose range contains its domain! Therefore, we know from our result at the start of the lecture on fixed points that there is some $x_0 \in I_0^*$ such that $f^n(x_0) = x_0$. This proves our claim: we have shown that there is a point x_0 such that

1. $f^n(x_0) = x_0$, and
2. $f^k(x_0) \in I_k$, for every $k = 0, \ldots n-1$.

$\square$