§2.9 Intermediate Value Theorem (IVT)

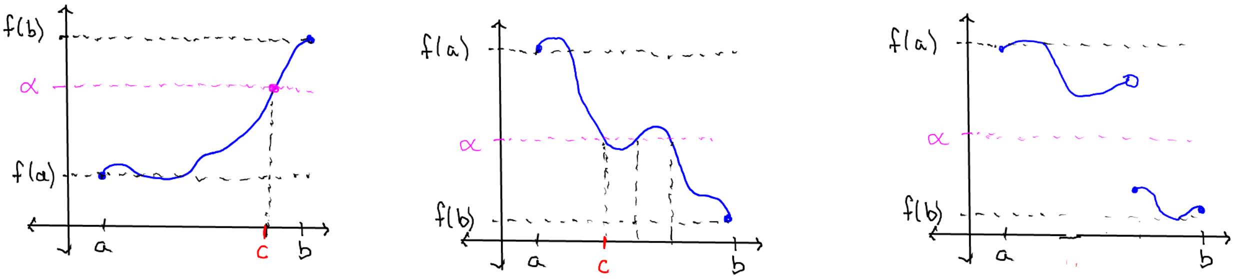
Theorem 2.16 (IVT)

Given a function f that is continuous on an interval $[a,b]$ and given a number $\alpha \in \mathbb{R}$, if either

$$f(a) < \alpha < f(b) \quad \text{or} \quad f(a) > \alpha > f(b)$$

then there exists a point $c \in (a,b)$ such that $f(c) = \alpha$.

The intuition behind this theorem is clear. If one is to draw the graph of the function, the curve must at some point cross the horizontal line at $y = \alpha$. (See picture below)



$f(a) < \alpha < f(b)$ $\qquad$ $f(a) > \alpha > f(b)$ $\qquad$ IVT can fail if f is not continuous

---

The value α is the "intermediate value" since it lies in between the starting and ending values ($f(a)$ and $f(b)$).

Note: The theorem guarantees the existence of at least one point $c \in (a,b)$ for which $f(c) = \alpha$. There could be more than one. Also, theorem does not tell us anything about where c is, only that it exists.

The proof of the IVT is rather difficult, so we'll worry about that later.

Examples

- Consider the function defined by $f(x) = x^5 - 2x^3 - 2$.

  Note that $f(0) = -2$ and $f(2) = 2^5 - 2\cdot 2^3 - 2$
  $= 32 - 16 - 2 = 14$.

  Since f is continuous everywhere and

  $$f(0) = -2 < 0 < 14 = f(2),$$

the IVT tells us that there is at least one <u>root</u> of f in the interval $[0,2]$.

> Recall: A <u>root</u> of a function f is a point $x \in \mathbb{R}$ such that $f(x) = 0$.

- We can show that there exists a point $x \in (0,1)$ such that

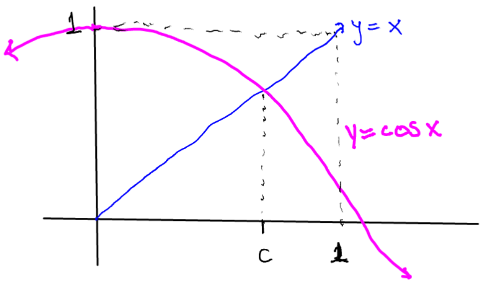
$$\cos x = x.$$

Define the function f as $f(x) = \cos(x) - x$, which is clearly continuous everywhere. Note that

$$f(0) = \cos 0 - 0 = 1$$

and $f(1) = \cos(1) - 1 < 0$ since $\cos(1) < 1$.

Hence $f(0) > 0 > f(1)$. By the IVT, there is a point $x \in (0,1)$ such that $f(x) = 0$. That is, $\cos x - x = 0$ and thus $\cos x = x$.



<u>Bisection Method for root finding</u>

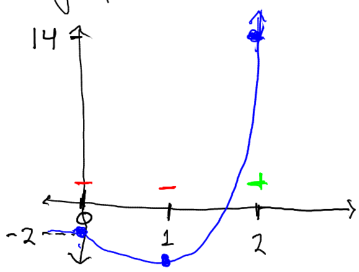
Let's return to the example from above, where we considered the function defined by $f(x) = x^5 - 2x^3 - 2$. We determined that there was at least one root of f in the interval $[0,2]$. There is no simple formula for finding the roots of a degree 5 polynomial. In fact, there is no "nice" expression for the exact position of the root in this interval. Using the IVT, however, we can get an <u>approximation</u> of the location of this root.

Since we know there is a root somewhere between 0 and 2, let's try some other points in this range to see if we can get a better idea of where the root is.

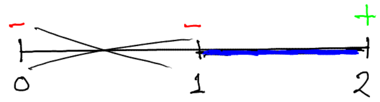
A good spot to check is $x = 1$, since 1 is exactly the midpoint of 0 and 2. Plugging this in, we see $f(1) = 1 - 2 - 2 = -3$, which is negative. Now we know that f is negative at 1 and positive at 2, so the IVT tells us there is a root between 1 and 2.

This cuts our interval where we need to look down by half!

The graph now must look something like: (not to scale)



So the root lies in $[1,2]$.



Checking the midpoint allows us to remove half of the original interval from the search.

To further narrow down where exactly the root lies, let's check the midpoint of our new interval:

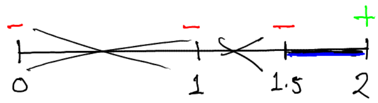
The midpoint of $[1,2]$ is $\frac{1+2}{2} = \frac{3}{2} \approx 1.5$.

Now $f(\frac{3}{2}) = \left(\frac{3}{2}\right)^5 - 2\left(\frac{3}{2}\right)^3 - 2$

$$= \frac{243}{32} - \frac{27}{4} - 2 = -\frac{37}{32} < 0,$$

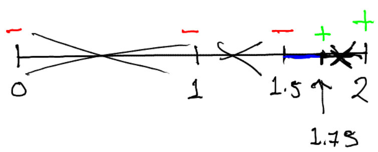
and thus $f(\frac{3}{2})$ is negative.

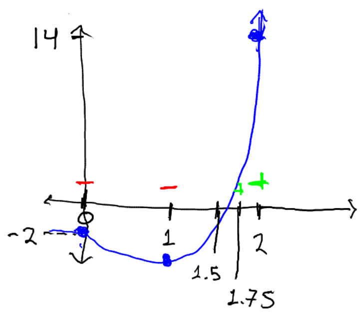
As we know $f(\frac{3}{2}) < 0 < f(2)$, we now know there is a root in the smaller interval $[1.5, 2]$.



Performing this procedure again, but on the interval $[1.5, 2]$, we check the midpoint $\frac{\frac{3}{2}+2}{2} = \frac{7}{4} = 1.75$

to find that $f(1.75) \approx 3.69$, which is positive. We now know the root lies between 1.5 and 1.75.





We can repeat this process indefinitely, at each step updating the lower and upper bounds of our estimate for the root location and checking the midpoint.

Let's summarize the results in the following table and repeat a few more times...

n step	a_n lower bound	b_n upper bound	c_n midpoint	$f(c_n)$ f(midpoint)	Sign	$\lvert b_n - a_n \rvert$
0	0	2	$\frac{0+2}{2} = 1$	-3	$-$	2
1	1	2	$\frac{1+2}{2} = \frac{3}{2} = 1.5$	≈ -1.16	$-$	1
2	$\frac{3}{2} = 1.5$	2	$\frac{7}{4} = 1.75$	≈ 3.69	$+$	$\frac{1}{2}$
3	1.5	$\frac{7}{4} = 1.75$	$\frac{13}{8} = 1.625$	≈ 0.75	$+$	$\frac{1}{4}$
4	1.5	$\frac{13}{8} = 1.625$	$\frac{25}{16} = 1.5625$	≈ -0.32	$-$	$\frac{1}{8}$
5	$\frac{25}{16} = 1.5625$	1.625	$\frac{51}{32} = 1.59375$	≈ 0.19	$+$	$\frac{1}{16}$
6	1.5625	$\frac{51}{32} = 1.59375$				$\frac{1}{32}$

Thus, after the 5th step we know that there is a root between

1.5625 and 1.59375.

We know the true location of the root up to a precision of

$$\frac{1}{32} = \left|\frac{51}{32} - \frac{50}{32}\right| = |1.59375 - 1.5625| = 0.03125.$$

Not bad!

Recap and algorithmic description of bisection method

At each step of the above procedure, we decrease the width of the interval of the known location of the root to $\frac{b-a}{2^n}$, where $a=0$ and $b=2$, where n is the number of the step. Since $\lim_{n\to\infty} \frac{b-a}{2^n} = 0$, we may continue this process to whatever degree of precision we wish!

In performing this procedure, we essentially defined three sequences,

$$a_1, a_2, a_3, \ldots, \qquad b_1, b_2, b_3, \ldots, \qquad c_1, c_2, c_3, \ldots$$

which are the lower bound, upper bound, and midpoint of the interval known to contain the root at step n. As n increases, each of these sequences seems to converge on the precise location of the true value of the root! If we only care about determining the root up to some given level of precision $\varepsilon > 0$, we can terminate the process once we get to the point where

$$|b_n - a_n| < \varepsilon.$$

On the other hand, if we ever get to a point where $f(c_n) = 0$, then we can also stop since we have found the true location of the root!

The Bisection Method is outlined explicitly in the following algorithm.

Algorithm (Bisection Method for root finding).

Let f be a continuous function on an interval $[a, b]$ and let $\varepsilon > 0$ be a given tolerance level for the precision.

- Define $a_0 = a$, $b_0 = b$, $c_0 = \frac{a_0 + b_0}{2}$.
- For each $n \in \mathbb{Z}$ with $n \geq 0$:
  - Define $c_n = \frac{a_n + b_n}{2}$.
  - If $|b_n - a_n| < \varepsilon$, stop! output c_n
  - If $f(c_n) = 0$, stop! Output c_n
  - If $f(c_n) < 0$, (increase lower bound)

    define $a_{n+1} = c_n$ and $b_{n+1} = b_n$.
  - If $f(c_n) > 0$, (decrease upper bound)

    define $a_{n+1} = a_n$ and $b_{n+1} = c_n$.
  - Continue to next n.

Note that the length of the search interval at step n is equal to $b_n - a_n = \frac{b-a}{2^n}$. Since we must continue the algorithm until $\frac{b-a}{2^n} < \varepsilon$, the maximum number of steps that needs to be performed is $\log_2\left(\frac{b-a}{\varepsilon}\right)$ where $\log_2$ is the logarithm base 2.

Proof of IVT

We are now ready to prove the IVT. Our technique will use the ideas outlined in the algorithm for the bisection method, as well as the Sequential Characterization of Continuity which we will state here. (See Theorem 2.11 in textbook)

Theorem (Sequential Characterization of Continuity).

Given a function f and a point $a \in \mathbb{R}$, we have the following equivalence:

$$f \text{ is continuous at } a \iff \text{For every sequence } x_1, x_2, x_3, \dots \text{ with } \lim_{n\to\infty} x_n = a, \text{ it holds that } \lim_{n\to\infty} f(x_n) = f(a).$$

This follows trivially from the definition of continuity and the sequential characterization of limits (Theorem 2.1).

We will prove a slightly simplified version of the IVT, which we state here:

Simplified IVT: Let f be a continuous function on the interval $[a,b]$ and suppose that $f(a) < 0 < f(b)$. Then there is a point $c \in (a,b)$ such that $f(c) = 0$.

This is just the IVT with $\alpha = 0$ and where $f(a) < f(b)$. The proof when $\alpha \neq 0$ and $f(a) > f(b)$ is similar.

Proof (of simplified IVT):

Following the procedure of the Bisection method, we define sequences recursively as follows. Define

$$a_0 = a \quad \text{and} \quad b_0 = b.$$

For each integer $n \geq 0$,

- Define $c_n = \dfrac{a_n + b_n}{2}$ (the midpoint of a_n and b_n).
- Consider the value of $f(c_n)$:
  - If $f(c_n) = 0$, then we are done, so suppose $f(c_n) \neq 0$.
  - If $f(c_n) < 0$, define $a_{n+1} = c_n$ and $b_{n+1} = b_n$.
  - If $f(c_n) > 0$, define $a_{n+1} = a_n$ and $b_{n+1} = c_n$.
- By construction, these sequences satisfy

- $f(a_n) < 0 < f(b_n)$,
- $a \le a_n < c_n < b_n \le b$,
- $a_n \le a_{n+1}$ and $b_{n+1} \le b_n$, and
- $|b_n - a_n| = \frac{b-a}{2^n}$

for every $n \geq 0$.

- In particular, the sequences $a_1, a_2, a_3, \ldots$ and $b_1, b_2, b_3, \ldots$ are both monotonic (since $\{a_n\}$ is non-decreasing and $\{b_n\}$ is non-increasing) and bounded (since $a_n, b_n \in [a,b]$ for each n). By the Monotone Convergence Theorem (MCT), both sequences converge so the limits

$$\lim_{n\to\infty} a_n \quad \text{and} \quad \lim_{n\to\infty} b_n$$

exist.

- Since $b_n - a_n = \frac{b-a}{2^n}$ for each $n \in \mathbb{N}$, we have that

$$\lim_{n\to\infty}(b_n - a_n) = \lim_{n\to\infty} \frac{b-a}{2^n} = 0$$

and thus $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

- Define $c = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.
- Since $a_n \in [a,b]$ and $b_n \in [a,b]$ for each $n \in \mathbb{N}$, it follows that $c \in [a,b]$.
- Because f is continuous on $[a,b]$, f is continuous at c.
- By the Sequential Characterization of Continuity, we have

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

- However, since $f(a_n) < 0 < f(b_n)$ holds for every $n \in \mathbb{N}$, we know that $\lim_{n\to\infty} f(a_n) \le 0 \le \lim_{n\to\infty} f(b_n)$.
- Now $f(c) = \lim_{n\to\infty} f(a_n) \le 0 \le \lim_{n\to\infty} f(b_n) = f(c)$, and thus $f(c) = 0$.

$\square$

§2.10 Extreme Value Theorem (EVT)

Here we present an important result that demonstrates why continuity on a closed interval is different from continuity on an open interval.

First we need a few definitions

Definition. Suppose f is a continuous function on some interval I, and let $c \in I$. We say that c is:

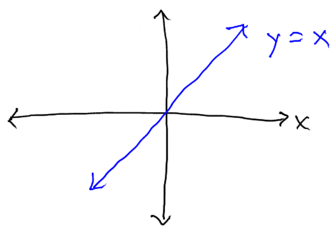
- a global maximum for f on I if $f(x) \leq f(c)$ holds for every $x \in I$,
- a global minimum for f on I if $f(x) \geq f(c)$ holds for every $x \in I$,
- a global extremum for f on I if it is either a global maximum or minimum.

Note: global minima/maxima are also called absolute extrema.

Question: Does every function f on an interval I always have a global maximum/global minimum?

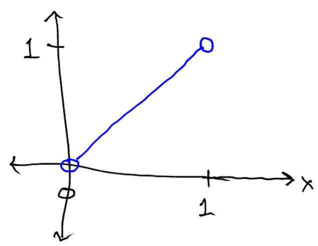
Ans: No! Consider the following examples.

- Consider $f(x) = x$ on $\mathbb{R}$. The function f is unbounded on this interval, and does not have a global min or max.



- Even if the interval is bounded, this function might not have a min/max if the interval is open.

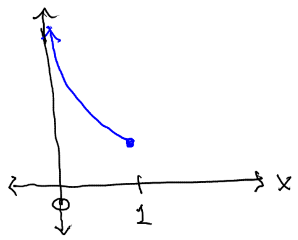
  The function $f(x) = x$ on $(0, 1)$ also does not have a global max/min.



  It seems like the min/max should be at $x = 0$ / $x = 1$, but these aren't in the interval $(0, 1)$.

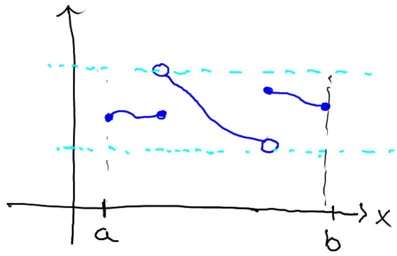
- The function could also be unbounded at the end points. E.g.

$$f(x) = \frac{1}{x} \quad \text{on } (0, 1]$$



has no global max, since it is unbounded above, but does have a global min at $x=1$.

- If the function is not continuous on the interval, then global max and min may not exist even if the interval is closed.



The function $f:[a,b]\to\mathbb{R}$ whose graph is depicted here does not achieve either it's max or min

However, if the interval I is closed and f is continuous on I, then it turns out that the global extrema are always achieved!

Theorem (EVT).

Suppose f is continuous on $[a,b]$ with $a<b$. There exist two numbers c_1 and c_2 in $[a,b]$ such that

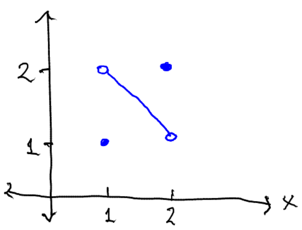
$$f(c_1)\le f(x)\le f(c_2)$$

for every $x\in[a,b]$.

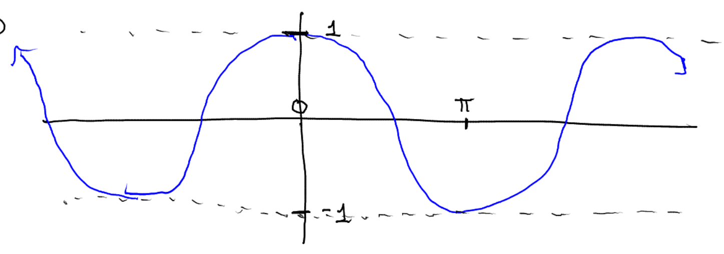
Note: There may be more than one global max/min. Also, the EVT does not say anything about how to find the extrema.

Note: Also note that the extrema can be achieved even if the conditions of the theorem are not met. For example, this discontinuous function $f:[1,2]\to\mathbb{R}$ achieves its global minimum at $x=1$ and global maximum at $x=2$:

$$f(x)=\begin{cases}1 & x=1\\ 3-x & 1<x<2\\ 2 & x=2\end{cases}$$



Also, the function defined by $g(x)=\cos x$ on $(-10,+10)$ achieves its global max at $x=0$ and global min at $x=\pi$.



While the Extreme Value Theorem is "intuitive", actually proving this result is rather difficult and beyond the scope of this course.