MA 109 Week 2

Mayukh Mukherjee

Department of Mathematics

December 06, 2021

Limits from the left and right

If $f : (a, b) \to \mathbb{R}$ is a function and $c \in (a, b)$, then it is possible to approach c from either the left or the right on the real line.

We can define the limit of the function $f(x)$ as x approaches c from the left (if it exists) as a number l^- such that for all $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - l^-| < \epsilon$ whenever $|x - c| < \delta$ and $x \in (a, c)$.

Our notation for this is $\lim_{x \to c-} f(x) = l^-$, and it is also called the left hand (side) limit.

Exercise 1: Write down a definition for the limit of a function from the right. We usually denote the right hand (side) limit by $lim_{x \to c+} f(x)$. Show, using the definitions, that $\lim_{x \to c} f(x)$ exists if and only if the left hand and right hand limits both exist and are equal.

We can also think of the left hand limit as follows. We restrict our attention to the interval (a, c), that is we think of f as a function only on this interval. Call this restricted function f_a. Then, another way of defining the left hand limit is

$$\lim_{x \to c-} f(x) = \lim_{x \to c} f_a(x).$$

It should be easy to see that it is the same as the definition before One can make a similar definition for the right hand limit.

The notions of left and right hand limits are useful because sometimes a function is defined in different ways to the left and right of a particular point. For instance, $|x|$ has different definitions to the left and right of 0.

Calculating limits explicitly

As with sequences, using the rules for limits of functions together with the Sandwich theorem allows one to treat the limits of a large number of expressions once one knows a few basic ones:

(i) $\lim_{x\to 0} x^{\alpha} = 0$ if $\alpha > 0$, (ii) $\lim_{x\to\infty} x^{\alpha} = 0$ if $\alpha < 0$,
(iii) $\lim_{x\to 0} \sin x = 0$, (iv) $\lim_{x\to 0} \sin x/x = 1$
(v) $\lim_{x\to 0}(e^x - 1)/x = 1$, (vi) $\lim_{x\to 0} \ln(1+x)/x = 1$
We have not concentrated on trying to find limits of complicated expressions of functions using clever algebraic manipulations or other techniques. However, I can't resist mentioning the following problem.
Exercise 2: Find

$$\lim_{x\to 0} \frac{\sin(\tan x) - \tan(\sin x)}{\arcsin(\arctan x) - \arctan(\arcsin x)}.$$

I will give the solution next time, together with the history of the problem (if I mention the history right away you will be able to get the solution by googling!), but feel free to use any method you like.

Continuity - the definition

Definition

If $f : [a, b] \to \mathbb{R}$ is a function and $c \in [a, b]$, then f is said to be continuous at the point c if and only if

$$\lim_{x \to c} f(x) = f(c).$$

Thus, if c is one of the end points we require only the left or right hand limit to exist.

A function f on (a, b) (resp. $[a, b]$) is said to be continuous if and only if it is continuous at every point c in (a, b) (resp. $[a, b]$).

If f is not continuous at a point c we say that it is discontinuous at c, or that c is a point of discontinuity for f.

Intuitively, continuous functions are functions whose graphs can be drawn on a sheet of paper without lifting the pencil of the sheet of paper. That is, there should be no "jumps" in the graph of the function.

Continuity of familiar functions: polynomials

What are the functions we really know or understand? What does "knowing" or understanding a function $f(x)$ even mean? Presumably, if we understand a function f, we should be able to calculate the value of the function $f(x)$ at any given point x. But if you think about it, for what functions $f(x)$ can you really do this?

One class of functions is the polynomial functions. More generally we can understand rational functions, that is functions of the form $R(x) = P(x)/Q(x)$ where $P(x)$ and $Q(x)$ are polynomials, since we can certainly compute the values of $R(x)$ by plugging in the value of x. How do we show that polynomials or rational functions are continuous (on $\mathbb{R}$)?

It is trivial to show from the definition that the constant functions and the function $f(x) = x$ are continuous. Because of the rules for limits of functions, the sum, difference, product and quotient (with non-zero denominator) of continuous functions are continuous. Applying this fact we see easily that $R(x)$ is continuous whenever the denominator is non-zero.

Continuity of other familiar functions

What are the other (continuous) functions we know? How about the trigonometric functions? Well, here it is less clear how to proceed. After all we can only calculate $\sin x$ for a few special values of x ($x = 0, \pi/6, \pi/4, \ldots$ etc.). How can we show continuity when we don't even know how to compute the function?

Of course, if we define $\sin x$ as the y-coordinate of a point on the unit circle it seems intuitively clear that the y-coordinate varies continuously as the point varies on the unit circle, but knowing the precise definition of continuity this argument should not satisfy you.

We will indicate a bit later how to prove the continuity of $\sin x$. For now, let us assume that $\sin x$ is continuous. Using that, how can we show that $\cos x$ is continuous?

The composition of continuous functions

Theorem

Let $f : (a, b) \to (c, d)$ and $g : (c, d) \to (e, f)$ be functions such that f is continuous at x_0 in (a, b) and g is continuous at $f(x_0) = y_0$ in (c, d). Then the function $g(f(x))$ (also written as $g \circ f(x)$ sometimes) is continuous at x_0. So the composition of continuous functions is continuous.

Exercise 3: Prove the theorem above starting from the definition of continuity.

Using the theorem above we can show that $\cos x$ is continuous if we show that $\sqrt{x}$ is continuous, since $\cos x = \sqrt{1 - \sin^2 x}$ and we know that $1 - \sin^2 x$ is continuous since it is the product of the sums of two continuous functions ($(1 + \sin x)$ and $(1 - \sin x)$!).

Once we have the continuity of $\cos x$ we get the continuity of all the rational trignometric functions, that is functions of the form $P(x)/Q(x)$, where P and Q are polynomials in $\sin x$ and $\cos x$, provided $Q(x)$ is not zero.

The continuity of the square root function

Thus in order to prove the continuity of $\cos x$ (assuming the continuity of $\sin x$) we need only prove the continuity of the square root function.

The main observation is that continuity is a local property, that is, only the behaviour of the function near the point being investigated is important.

Let $x_0 \in [0, \infty)$. To show that the square root function is continuous at x_0 we need to show that $\lim_{y \to x_0} \sqrt{y} = \sqrt{x_0}$, that is we need to show that $|\sqrt{y} - \sqrt{x_0}| < \epsilon$ whenever $0 < |y - x_0| < \delta$ for some δ. First assume that $x_0 \neq 0$. Then

$$|\sqrt{y} - \sqrt{x_0}| = \left| \frac{y - x_0}{\sqrt{y} + \sqrt{x_0}} \right| < \frac{|y - x_0|}{\sqrt{x_0}}.$$

If we choose $\delta = \epsilon \sqrt{x_0}$, we see that

$$|\sqrt{y} - \sqrt{x_0}| < \epsilon,$$

which is what we needed to prove. When $x_0 = 0$, I leave the proof as an exercise. $\square$

The intermediate value theorem

One of the most important properties of continuous functions is the Intermediate Value Property (IVP). We will use this property repeatedly to prove other results.

Theorem

Suppose $f : [a, b] \to \mathbb{R}$ is a continuous function. For every u between $f(a)$ and $f(b)$ there exists $c \in [a, b]$ there such that $f(c) = u$.

Functions which have this property are said to have the Intermediate Value Property. Theorem 9 can thus be restated as saying that continuous functions have the IVP.

We will not be proving this property - it is a consequence of the completeness of the real numbers. Intuitively, this is clear. Since one can draw the graph of the function without lifting one's pencil off the sheet of paper, the pencil must cut every line $y = e$ with e between $f(a)$ and $f(b)$.

The IVT in a picture

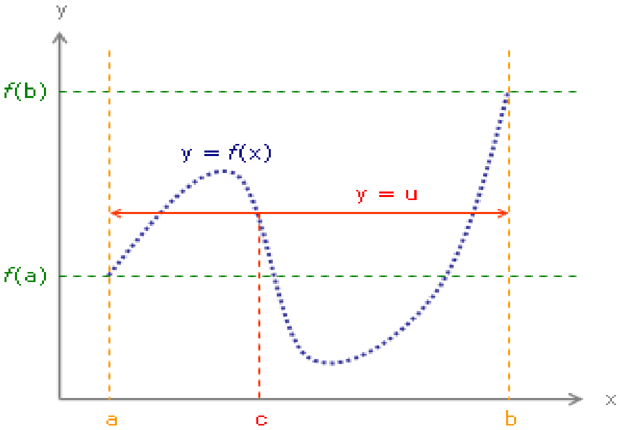

Image created by Enoch Lau see http://en.wikipedia.org/wiki/File:Intermediatevaluetheorem.png (Creative Commons Attribution-Share Alike 3.0 Unported license).

Zeros of functions

One of the most useful applications of the intermediate value property is to find roots of polynomials, or, more generally, to find zeros of continuous functions, that is to find points $x \in \mathbb{R}$ such that $f(x) = 0$.

Theorem

Every polynomial of odd degree has at least one real root.

Proof: Let $P(x) = a_n x^n + \ldots + a_0$ be a polynomial of odd degree. We can assume without loss of generality that $a_n > 0$. It is easy to see that if we take $x = b > 0$ large enough, $P(b)$ will be positive. On the other hand, by taking $x = a < 0$ small enough, we can ensure that $P(a) < 0$. Since $P(x)$ is continuous, it has the IVP, so there must be a point $x_0 \in (a, b)$ such that $f(x_0) = 0$. $\square$

The IVP can often be used to get more specific information. For instance, it is not hard to see that the polynomial $x^4 - 2x^3 + x^2 + x - 3$ has a root that lies between 1 and 2.

Continuous functions on closed, bounded intervals

The other major result on continuous functions that we need is the following. A closed bounded interval is one of the form $[a, b]$, where $-\infty < a$ and $b < \infty$.

Theorem

A continuous function on a closed bounded interval $[a, b]$ is bounded and attains its infimum and supremum, that is, there are points x_1 and x_2 in $[a, b]$ such that $f(x_1) = m$ and $f(x_2) = M$, where m and M denote the infimum and supremum respectively.

Again, we will not prove this, but will use it quite often. Note the contrast with open intervals. The function $1/x$ on $(0, 1)$ does not attain a maximum - in fact it is unbounded. Similarly the function $1/x$ on $(1, \infty)$ does not attain its minimum, although, it is bounded below.

The function $\sin \frac{1}{x}$

Let us look at Exercise 2.3 part (i).
Consider the function defined as $f(x) = \sin \frac{1}{x}$ when $x \neq 0$, and $f(0) = 0$.
The question asks if this function is continuous at $x = 0$.
How about $x \neq 0$? Why is $f(x)$ continuous? Because it is a compostion of the sin function and a rational function $1/x$. Since both of these are continuous when $x \neq 0$, so is $f(x)$.

Let us look at the sequence of points $x_n = 2/(2n+1)\pi$. Clearly $x_n \to 0$ as $n \to \infty$.
For these points $f(x_n) = \pm 1$. This means that no matter how small I take my δ, there will be a point $x_n \in (0, \delta)$, such that $|f(x_n)| = 1$. But this means that $|f(x) - f(0)| = |f(x)|$ cannot be made smaller than 1 no matter how small δ may be. Hence, f is not continuous at 0. The same kind of argument will show that there is no value that we can assign $f(0)$ to make the function $f(x)$ continuous at 0.

You can easily check that $f(x)$ has the IVP. However, we have proved that it is not continuous. So IVP $\not\Rightarrow$ continuity.

Sequential continuity

The preceeding example showed that in order to demonstrate that a function, say $f(x)$, is not continuous at a point x_0 it is enough to find a sequence x_n tending to x_0 such that the value of the function $|f(x_n) - f(x_0)|$ remains large. Suppose it is not possible to find such a sequence. Does that mean the function is continuous at x_0? The following theorem answers the question affirmatively.

Theorem

A function $f(x)$ is continuous at a point a if and only if for every sequence $x_n \to a$, $\lim_{x_n \to a} f(x_n) = f(a)$.

A function that satisfies the property that for every sequence $x_n \to a$, $\lim_{x_n \to a} f(x_n) = f(a)$ is said to be sequentially continuous. The theorem says that sequential continuity and continuity are the same thing. Indeed, it is clear that a continuous function is necessarily sequentially continuous. It is the reverse that is slightly harder to prove.

Limits of functions of several variables

Just like we did for sequences in $\mathbb{R}^2$ and $\mathbb{R}^3$, we can define the notion of the limit of a function for functions from $\mathbb{R}^2$ to $\mathbb{R}$.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function. The function $f(x_1, x_2)$ is said to tend to a limit l as $(x_1, x_2) \to (a_1, a_2)$ if for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f(x_1, x_2) - l| < \epsilon$$

whenever $0 < \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2} < \delta$. Notice, that one can now approach the point (a_1, a_2) from any direction in the plane. Our definition requires that the limits from the different directions all exist and be equal. This is quite a powerful condition.

If we have functions from $\mathbb{R}^2 \to \mathbb{R}^2$ we can make exactly the same definition. But this time $l = (l_1, l_2)$ will be in $\mathbb{R}^2$ and so will $f(x_1, x_2) = (f_1(x_1, x_2), f_2(x_1, x_2))$, so we will have to replace the modulus function by the distance between these two quantities:

$$\sqrt{[f_1(x_1, x_2) - l_1]^2 + [f_2(x_1, x_2) - l_2]^2}.$$

Limits of functions of several variables

The definitions we have made go through for functions from $\mathbb{R}^m$ to $\mathbb{R}^n$, where m and n may be different. For instance, we have considered the case when $m = 2$ and $n = 1$ and also the case $m = 2$ and $n = 2$ above. But we could allow m and n to take any of the values 1, 2 or 3 (in fact, we can allow values greater than 3 as well!).

Exercise 1: Show that

$$\lim_{y \to x} f(y) = l$$

for $f = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2 \iff$

$$\lim_{y \to x} f_1(y) = l_1 \quad \text{and} \quad \lim_{y \to x} f_2(y) = l_2,$$

where $l = (l_1, l_2)$ In other words, when dealing with limits of functions which are vector-valued, it is enough to study the limits of the coordinate functions.

Continuous functions of several variables

Once the definition of the limit is clear it makes sense to talk of continuity as well. All the definitions remain the same, only the definition of the distance function changes depending on the domain and the range.

For instance, provided we know what "closed and bounded sets" are in $\mathbb{R}^2$ or $\mathbb{R}^3$, Theorem 11 goes through for continuous functions $f : \mathbb{R}^m \to \mathbb{R}$. $(m = 2, 3)$. For functions with more than one variable in the range the first part of Theorem 11 still works, but for the second part things are more compacted (again there is no "ordering" in $\mathbb{R}^2$ or $\mathbb{R}^3$).

While it is easy to see what a bounded set in $\mathbb{R}^m$ should be, closed is a little more complicated and we will not give the definition here. However, a rectangle of the form $[a, b] \times [c, d]$ in $\mathbb{R}^2$ is an example of a closed and bounded set (also called "compact sets" of this form).
Theorem 12 goes through without any problems even when the range is in $\mathbb{R}^2$ or $\mathbb{R}^3$.

Derivatives: the definition

For now, if you did not understand the rigorous definition of the limit, forget about it. You will be able to understand what follows as long as you remember your 11th standard treatment of limits.

Recall that $f : (a, b) \to \mathbb{R}$ is said to be differentiable at a point $c \in (a, b)$ if

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h}$$

exists. In this case the value of the limit is denoted $f'(c)$ and is called the derivative of f at c. The derivative may also be denoted by $\frac{df}{dx}(c)$ or by $\frac{dy}{dx}|_c$, where $y = f(x)$.

In general, the derivative measures the rate of change of a function at a given point. Thus, if the function we are studying is the position of a particle on the x-coordinate, then $x'(t)$ is the velocity of the particle. If the function we are studying is the velocity $v(t)$ of the particle, then the derivative $v'(t)$ is the acceleration of the particle. If the function we are studying is the population of India, then the derivative measures the rate of change of the population.

The slope of the tangent

From the point of view of geometry, the derivative $f'(c)$ gives us the slope of the curve, that is, the slope of the tangent to the curve $y = f(x)$ at $(c, f(c))$. This becomes particularly clear if we rewrite the derivative as the following limit:

$$\lim_{y \to c} \frac{f(y) - f(c)}{y - c}.$$

The expression inside the limit obviously represents the slope of a line passing through $(c, f(c))$ and $(y, f(y))$, and as y approaches c this line obviously becomes tangent to $y = f(x)$ at the point $(c, f(c))$.

Another way of thinking of the derivative

Another way of thinking of the derivative of the function f at the point x_0 is as follows. If f is differentiable at x_0 we know that

$$\frac{f(x_0+h)-f(x_0)}{h} - f'(x_0) \to 0$$

as $h \to 0$. Since we are keeping x_0 fixed, we can treat the above quantity as a function of h. Thus we can write

$$\frac{f(x_0+h)-f(x)}{h} - f'(x_0) = o(h)$$

for some function $o(h)$ with the property that $o(h) \to 0$ as $h \to 0$. Taking a common denominator,

$$\frac{f(x_0+h)-f(x_0)-f'(x_0)h}{h} = o(h) \tag{1}$$

We can use the above equality to give an equivalent definition for the derivative. A function f is said to be differentiable at the point x_0 if there exists a real number (denoted $f'(x_0)$) such that (1) holds for some function $o(h)$ such that $o(h) \to 0$ as $h \to 0$.

Exercise 2.6: Let $f : (a, b) \to \mathbb{R}$ be a function such that

$$|f(x + h) - f(x)| \leq C|h|^{\alpha}$$

for all $x, x + h \in (a, b)$, where C is a constant and $\alpha > 1$. Show that f is differentiable on (a, b) and compute $f^{'}(x)$ for $x \in (a, b)$.

Solution:

$$\left|\lim_{h\to 0} \frac{f(x + h) - f(x)}{h}\right| \leq C \lim_{h\to 0} |h|^{\alpha-1} = 0.$$

Note: Functions that satisfy the property above for some α (not necessarily greater than 1) are said to be Lipschitz continuous with exponent α.

Equivalent condition for differentiability

Let $f : (a, b) \to \mathbb{R}$, and let $c \in (a, b)$.

Recall that f is said to be differentiable at c if the limit

$$f'(c) := \lim_{h \to 0} \frac{f(c+h) - f(c)}{h} \quad \text{exists.}$$

Theorem (Carathéodory lemma: C-lemma)

Let $f : (a, b) \to \mathbb{R}$, and $c \in (a, b)$. Then f is differentiable at $c \iff$ there is a function $f_1 : (a, b) \to \mathbb{R}$ which is continuous at c and satisfies

$$f(x) - f(c) = (x - c) f_1(x) \quad \text{for all } x \in (a, b).$$

In this case, the function f_1 is unique and $f'(c) = f_1(c)$.

The function $f_1 : (a, b) \to \mathbb{R}$ is called the **increment function** associated with f and c.

Proof: Suppose f is differentiable at c. Define

$$f_1(x) := \begin{cases} \frac{f(x)-f(c)}{x-c} & \text{if } x \in (a,b) \setminus \{c\}, \\ f'(c) & \text{if } x = c. \end{cases}$$

Then f_1 is continuous at c since $\lim_{x\to c} f_1(x) = f'(c) = f_1(c)$.

Conversely, suppose there is $f_1 : (a,b) \to \mathbb{R}$ as stated. Let $h := x - c$ for $x \in (a,b)$, and so $c + h = x$. Then

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = \lim_{x\to c} \frac{f(x) - f(c)}{x - c} = \lim_{h\to 0} f_1(x) = f_1(c),$$

since f_1 is continuous at c. Hence f is differentiable at c.
The uniqueness of the increment function f_1 is obvious (left as an exercise). $\square$

The C-lemma relates the concepts of differentiability and continuity. It enables us to prove the following results neatly.

Differentiability $\implies$ Continuity

Theorem

Suppose $f : [a, b] \to \mathbb{R}$ is differentiable at $c \in (a, b)$. Then f is continuous at c.

Proof: Let f_1 be the increment function associated with f and c. Then $f(x) = f(c) + (x - c)f_1(x)$ for $x \in I$. Since $f_1 : I \to \mathbb{R}$ is continuous at c, so is f.

Remarks:
1. If a function f is not continuous at $c \in (a, b)$, then it cannot be differentiable at c.
Example: Let $f(x) := [x]$ for $x \in \mathbb{R}$, and $c := 1$.

2. The converse of the above theorem is false.
Example: $f(x) = |x|$ for $x \in \mathbb{R}$, and $c := 0$.

Differentiability: Algebraic rules

Let f and g be differentiable at c. Then $f \pm g$ and $f \cdot g$ are differentiable at c, and

$$(f \pm g)'(c) = f'(c) \pm g'(c) \text{ resp.,}$$

$$(f \cdot g)'(c) = f'(c)g(c) + f(c)g'(c).$$

Further, if $g(c) \neq 0$, then f/g is differentiable at c, and

$$\left(\frac{f}{g}\right)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{g(c)^2}.$$

Note: If f is differentiable at c, then f^n is differentiable at c for each $n \in \mathbb{N}$, and $(f^n)'(c) = n\,f(c)^{n-1}f'(c)$.
This also holds for a negative integer n if $f(c) \neq 0$.

Examples: Every polynomial function p is differentiable on $\mathbb{R}$.
A rational function p/q is differentiable at c if $q(c) \neq 0$.

Differentiability: Chain Rule

Theorem

Let $f : (a_1, b_1) \to \mathbb{R}$. Also, let $g : (a_2, b_2) \to \mathbb{R}$ be such that $f((a_1, b_1)) \subset (a_2, b_2)$. Suppose c is an interior point of (a_1, b_1) and $f(c)$ is an interior point of (a_2, b_2). If f is differentiable at c, and if g is differentiable at $f(c)$, then $g \circ f : (a_1, b_1) \to \mathbb{R}$ is differentiable at c, and $(g \circ f)'(c) = g'(f(c))f'(c)$.

Proof. Let f_1 be the increment functions associated with f and c, and let g_1 be the increment functions associated with g and $d := f(c)$. Then $f(x) - f(c) = (x - c)f_1(x)$ for all $x \in (a_1, b_1)$, where f_1 is continuous at c, and $g(y) - g(d) = (y - d)g_1(y)$ for all $y \in (a_2, b_2)$, where g_1 is continuous at d.

Define $h := g \circ f : (a_1, b_1) \to \mathbb{R}$. Then for $x \in (a_1, b_1)$,

$$\begin{aligned} h(x) - h(c) &= g(f(x)) - g(f(c)) \\ &= (f(x) - f(c))g_1(f(x)) = (x - c)f_1(x)(g_1 \circ f)(x). \end{aligned}$$

Let $h_1 := f_1 \cdot (g_1 \circ f)$. Then h_1 is continuous at c since f and f_1 are continuous at c and g_1 is continuous at $d = f(c)$. It follows that h_1 is the increment function associated with $g \circ f$ and c, and

$$(g \circ f)'(c) = h_1(c) = g'(f(c))f'(c). \qquad \square$$

<u>Conclusion of the Chain Rule in the Leibnitz notation:</u>

If $y := f(x)$ and $z := g(y)$, then $z = (g \circ f)(x)$ and

$$\left.\frac{dz}{dx}\right|_{x=c} = \left.\frac{dz}{dy}\right|_{y=f(c)} \cdot \left.\frac{dy}{dx}\right|_{x=c}$$

Warning:

Cancelling out dy in the RHS above to obtain the LHS does not yield a proof of the Chain Rule!

Example: Consider

$$f(x) := \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if} \quad x \in \mathbb{R} \setminus \{0\}, \\ 0 & \text{if} \quad x := 0. \end{cases}$$

Let

$$f_1(x) := \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{if} \quad x \in \mathbb{R} \setminus \{0\}, \\ 0 & \text{if} \quad x := 0. \end{cases}$$

Then f_1 is continuous at 0, and $f(x) - f(0) = (x - 0)f_1(x)$ for all $x \in \mathbb{R}$. By the C-lemma, f is differentiable at 0, and $f'(0) = f_1(0) = 0$. Also, if $c \neq 0$, then f is differentiable at c, and $f'(c) = 2c \sin \frac{1}{c} - \cos \frac{1}{c}$. (Multiplication rule and chain rule.)

Note: The derivative f' of f is not continuous at 0. Why? (We shall discuss this point later again!)

Maxima and minima

Let $X \subset \mathbb{R}$ and let $f : X \to \mathbb{R}$ be a function (you can think of X as an open, closed or half-open interval, for instance).

Definition: The function f is said to attain a maximum (resp. minimum) at a point $x_0 \in X$ if $f(x) \leq f(x_0)$ (resp. $f(x) \geq f(x_0)$) for all $x \in X$.

Once again, I remind you that, in general, f may not attain a maximum or minimum at all on the set X. The standard example being $X = (0, 1)$ and $f(x) = 1/x$ (can you find an example on the closed interval $[0, 1]$?). However, if X is a closed bounded interval and f is a continuous function, we have already seen the theorem that tells us that the maximum and minimum are actually attained.

Maxima and minima and the derivative

If f has a maximum at the point x_0 and if it also differentiable at x_0, we can reason as follows. We know that $f(x_0 + h) - f(x_0) \leq 0$ for every $h > 0$ such that $x + h \in X$. Hence, we see that (one half of the Sandwich Theorem!)

$$\lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h} \leq 0.$$

On the other hand, when $h < 0$, we get

$$\lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0.$$

Because f is assumed to be differentiable at x_0 we know that left and right hand limits must be equal. It follows that we must have $f'(x_0) = 0$. A similar argument shows that $f'(x_0) = 0$ if f has a minimum at the point x_0.

Local maxima and minima

The preceeding argument is purely local. Before explaining what this means, we give the following defintion.

Definition

Let $f : X \to \mathbb{R}$ be a function and x_0 be in X. Suppose there is an sub-interval $x_0 \in (c, d) \subset X$ such that $f(x_0) \geq f(x)$ (resp. $f(x_0) \leq f(x)$) for all $x \in (c, d)$, then f is said to have a local maximum (resp. local minimum) at x_0.

Sometimes we use the terms global maximum or global minimum instead of just maximum or minimum in order to emphasize the points are not just local maxima or minima. The argument of the previous slide actually proves the following

Theorem

If $f : X \to \mathbb{R}$ is differentiable and has a local minimum or maximum at a point $x_0 \in X$, $f'(x_0) = 0$.

Rolle's Theorem

The last theorem is known as Fermat's theorem. It can be used to prove Rolle's Theorem.

Theorem

Suppose $f : [a, b] \to \mathbb{R}$ *is continuous, differentiable in* (a, b) *and* $f(a) = f(b)$. *Then there is a point* x_0 *in* (a, b) *such that* $f'(x_0) = 0$.

Proof: Since f is a continuous function on a closed bounded interval, f must attain its minimum and maximum somewhere in $[a, b]$. If both the minimum and maximum are attained at the end points, f must be the constant function, in which case we know that $f'(x) = 0$ for all $x \in (a, b)$. Hence, we can assume that at least one of the minimum or maximum is attained at an interior point x_0 and Theorem 13 shows that $f'(x_0) = 0$ in this case. $\square$

One easy consequence: If $P(x)$ is a polynomial of degree n with n real roots, then all the roots of $P'(x)$ are also real. (How do we know that polynomials are differentiable?)

Problems centered around Rolle's Theorem

Exercise 3.3: Let $f : [a, b] \to \mathbb{R}$ be continuous and suppose f is differentiable on (a, b). If $f(a)$ and $f(b)$ are of opposite signs and $f'(x) \neq 0$ for all $x \in (a, b)$, then there is a unique point x_0 in (a, b) such that $f(x_0) = 0$.

Solution: Since the Intermediate Value Theorem guarantees the existence of a point x_0 such that $f(x_0) = 0$, the real point of this exercise is the uniqueness.

Suppose there were two points $x_1, x_2 \in (a, b)$ such that $f(x_1) = f(x_2) = 0$. Applying Rolle's Theorem, we see that there would exist $c \in (x_1, x_2)$ such that $f'(c) = 0$ contradicting our hypothesis, This proves the exercise.

Let us look at Exercise 2.8(i): Find a function f which satisfies all the given conditions, or else show that no such function exists: $f''(x) > 0$ for all $x \in \mathbb{R}$ and $f'(0) = 1, f'(1) = 1$.

Solution: Apply Rolle's Theorem to $f'(x)$ to conclude that such a function cannot exist.

The Mean Value Theorem

Rolle's theorem is a special case of the Mean Value Theorem (MVT).

Theorem

Suppose that $f : [a, b] \to \mathbb{R}$ is a continuous function and that f is differentiable in (a, b). Then there is a point x_0 in (a, b) such that

$$\frac{f(b) - f(a)}{b - a} = f'(x_0).$$

Proof: Apply Rolle's Theorem to the function

$$g(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a} \cdot (x - a).$$

$\square$

(Why does one think of the function $g(x)$?)

Applications of the MVT

Here is an application of the MVT which you have probably always taken for granted:

Theorem

If f satisfies the hypotheses of the MVT, and further $f'(x) = 0$ for every $x \in (a, b)$, f is a constant function.

Indeed, if $f(c) \neq f(d)$ for some two points $c < d$ in $[a, b]$,

$$0 \neq \frac{f(d) - f(c)}{d - c} = f'(x_0),$$

for some $x_0 \in (c, d)$, by the MVT. This contradicts the hypothesis. $\square$

Applications of the MVT continued

Consider Exercise 2.6.:

Let f be continuous on $[a, b]$ and differentiable on (a,b). If $f(a) = a$ and $f(b) = b$, show that there exist distinct $c_1, c_2 \in (a, b)$ such that $f'(c_1) + f'(c_2) = 2$.

Solution: The idea is that the function clearly has an average rate of growth equal to 1 on the interval $[a, b]$. If the derivative at some point is less than 1, there must be another point where it is greater than 1 so that the sum adds up to 2. How to use this idea?

Split the interval into two pieces - $[a, \frac{a+b}{2}]$ and $[\frac{a+b}{2}, b]$ - and apply the MVT to each interval.

Darboux's Theorem

Another interesting property of differentiable functions is that their derivatives have the IVP (intermediate value property). This fact is sometimes called Darboux's Theorem.

Theorem

Let $f : (a, b) \to \mathbb{R}$ be a differentiable function. If c, d, $c < d$ are points in (a, b), then for every u between $f'(c)$ and $f'(d)$, there exists an x in $[c, d]$ such that $f'(x) = u$.

Proof: We can assume, without loss of generality, that $f'(c) < u < f'(d)$, otherwise we can take $x = c$ or $x = d$. Define $g(t) = ut - f(t)$. This is a continuous function on $[c, d]$, and hence must attain its extreme values. The maximum value cannot occur at c or d since $g'(c) = u - f'(c) > 0$ and $g'(d) = u - f'(d) < 0$.

Suppose g takes a maximum at c. Since $g'(c) > 0$, for $h > 0$ small enough, we must have $g(c + h) - g(c) > 0$, contradiction.

Suppose g takes a maximum at d. Since $g'(d) < 0$, for $h < 0$ small enough, we must have $g(d + h) - g(d) > 0$, contradiction.

It follows that there exists $x \in (c, d)$ where g takes a maximum. By Fermat's Theorem $g'(x) = 0$ which yields $f'(x) = u$. $\square$

Continuity of the first derivative

We have just seen that the derivative satisfies the IVP. Can we find a function which is differentiable but for which the derivative is not continuous?

Here is the standard example:

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & \text{if } x \neq 0 \\ = 0 & \text{if } x = 0. \end{cases}$$

This function will be differentiable at 0 but its derivative will not be continuous at that point. In order to see this you will need to study the function in Exercise 1.13(ii). This will show that $f'(0) = 0$. On the other hand, if we use the product rule when $x \neq 0$ we get

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x},$$

which does not go to 0 as $x \to 0$.

Back to maxima and minima

We will assume that $f : [a, b] \to \mathbb{R}$ is a continuous function and that f is differentiable on (a, b). A point x_0 in (a, b) such that $f'(x_0) = 0$ often called a stationary point. We will assume further that $f'(x)$ is differentiable at x_0, that is, that the second derivative $f''(x_0)$ exists. We formulate the Second Derivative Test below.

Theorem

With the assumptions above:

1. *If $f''(x_0) > 0$, the function has a local minimum at x_0.*
2. *If $f''(x_0) < 0$, the function has a local maximum at x_0.*
3. *If $f''(x_0) = 0$, no conclusion can be drawn.*

The proof of the Second Derivative Test

Proof: The proofs are straightforward. For instance, to prove the first part we observe that

$$0 < f''(x_0) = \lim_{h \to 0} \frac{f'(x_0 + h)}{h}.$$

It follows that for $|h|$ small enough, $f'(x_0 + h) < 0$, if $h < 0$ and $f'(x_0 + h) > 0$ if $h > 0$. It follows that $f(x_0)$ is decreasing to the left of x_0 and increasing to the right of x_0. Hence, x_0 must be a local minimum. A similar argument yields the second case. □

If the third case of the theorem above occurs, the function may be changing from concave to convex. In this case x_0 is called a point of inflection. An example of this phenomenon is given by $f(x) = x^3$ at $x = 0$.