Limit, Derivative, Extrema of a Function

Hayk Aprikyan, Hayk Tarkhanyan

Suppose you run a supermarket and your profits from the sales of apples vary like this:

$$f(x) = 20\sqrt{x} - 3x^3$$

where x is the price of apples in dollars.

Suppose you run a supermarket and your profits from the sales of apples vary like this:

$$f(x) = 20\sqrt{x} - 3x^3$$

where x is the price of apples in dollars.

As a decent businessperson, you want to make your profit as high as possible, so you are interested in the following question: How much should x be so the profit f(x) is maximized?

Suppose you run a supermarket and your profits from the sales of apples vary like this:

$$f(x) = 20\sqrt{x} - 3x^3$$

where x is the price of apples in dollars.

As a decent businessperson, you want to make your profit as high as possible, so you are interested in the following question: How much should x be so the profit f(x) is maximized?

In real life where you sell more goods rather than apples, the situation looks more complicated.

Suppose you run a supermarket and your profits from the sales of apples vary like this:

$$f(x) = 20\sqrt{x} - 3x^3$$

where x is the price of apples in dollars.

As a decent businessperson, you want to make your profit as high as possible, so you are interested in the following question: How much should x be so the profit f(x) is maximized?

In real life where you sell more goods rather than apples, the situation looks more complicated. For example, your profits look like

$$f(x, y, z, t) = 3xy^{2} - y \log t - (1 - y) \log(1 - t) + \frac{z^{3}}{t}$$

with real-time values x = 4, y = 0.4, z = 0.8, t = 55, and you should decide whether to increase or decrease each of x, y, z, t (and how much).

Aprikyan, Tarkhanyan Lecture 6 2/40

Suppose you run a supermarket and your profits from the sales of apples vary like this:

$$f(x) = 20\sqrt{x} - 3x^3$$

where x is the price of apples in dollars.

As a decent businessperson, you want to make your profit as high as possible, so you are interested in the following question: How much should x be so the profit f(x) is maximized?

In real life where you sell more goods rather than apples, the situation looks more complicated. For example, your profits look like

$$f(x, y, z, t) = 3xy^{2} - y \log t - (1 - y) \log(1 - t) + \frac{z^{3}}{t}$$

with real-time values x=4, y=0.4, z=0.8, t=55, and you should decide whether to increase or decrease each of x,y,z,t (and how much). In machine learning you often have 1.000.000+ such parameters.

To begin our journey, let's start with the definition of a sequence of numbers.

To begin our journey, let's start with the definition of a sequence of numbers. A *sequence of numbers* is an unending (infinite)... well, sequence of numbers.

Example

- 1, 2, 3, 4, 5, ...
- \bullet 1, -1, 1, -1, 1, ...
- 0, 0.2, 0.4, 0.6, 0.8, ...
- 6, 6, 6, 6, 6, ...

To begin our journey, let's start with the definition of a sequence of numbers. A *sequence of numbers* is an unending (infinite)... well, sequence of numbers.

Example

- 1, 2, 3, 4, 5, ...
- \bullet 1, -1, 1, -1, 1, ...
- 0, 0.2, 0.4, 0.6, 0.8, ...
- 6, 6, 6, 6, 6, ...

We usually fix a letter, say a, and denote the first term by a_1 , the second term by a_2 , and so on. In general, for the n^{th} term we write a_n , and to denote the whole sequence we use $\{a_n\}$ or $\{a_n\}_{n=1}^{\infty}$.

3 / 40

To begin our journey, let's start with the definition of a sequence of numbers. A *sequence of numbers* is an unending (infinite)... well, sequence of numbers.

Example

- 1, 2, 3, 4, 5, ...
- \bullet 1, -1, 1, -1, 1, ...
- 0, 0.2, 0.4, 0.6, 0.8, ...
- 6, 6, 6, 6, 6, ...

We usually fix a letter, say a, and denote the first term by a_1 , the second term by a_2 , and so on. In general, for the n^{th} term we write a_n , and to denote the whole sequence we use $\{a_n\}$ or $\{a_n\}_{n=1}^{\infty}$.

Sometimes it also comes in handy to give the formula of the general n^{th} term, e.g. $a_n = n^2$ or $\{a_n\} = \{n^2\}$, which means:

$$a_1 = 1,$$
 $a_2 = 4,$ $a_3 = 9,$...

There are many interesting examples of sequences. Take, for example, the sequence $a_n = \frac{1}{n}$:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

There are many interesting examples of sequences. Take, for example, the sequence $a_n = \frac{1}{n}$:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

Note that as n becomes larger, the sequence gets smaller and smaller.

Aprikyan, Tarkhanyan Lecture 6 4 / 40

There are many interesting examples of sequences. Take, for example, the sequence $a_n = \frac{1}{n}$:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

Note that as n becomes larger, the sequence gets smaller and smaller.

Question

Is there a number to which $\{a_n\}$ approaches (gets closer and closer to)?

There are many interesting examples of sequences. Take, for example, the sequence $a_n = \frac{1}{n}$:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

Note that as n becomes larger, the sequence gets smaller and smaller.

Question

Is there a number to which $\{a_n\}$ approaches (gets closer and closer to)?

Indeed, we can say that as $n \to \infty$, the sequence moves closer to 0.

There are many interesting examples of sequences. Take, for example, the sequence $a_n = \frac{1}{n}$:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

Note that as n becomes larger, the sequence gets smaller and smaller.

Question

Is there a number to which $\{a_n\}$ approaches (gets closer and closer to)?

Indeed, we can say that as $n \to \infty$, the sequence moves closer to 0. We say that 0 is *the limit* of the sequence.

There are many interesting examples of sequences. Take, for example, the sequence $a_n = \frac{1}{n}$:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

Note that as n becomes larger, the sequence gets smaller and smaller.

Question

Is there a number to which $\{a_n\}$ approaches (gets closer and closer to)?

Indeed, we can say that as $n \to \infty$, the sequence moves closer to 0. We say that 0 is *the limit* of the sequence.

Question

Does the sequence become equal to 0 at some point?

There are many interesting examples of sequences. Take, for example, the sequence $a_n = \frac{1}{n}$:

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$$

Note that as n becomes larger, the sequence gets smaller and smaller.

Question

Is there a number to which $\{a_n\}$ approaches (gets closer and closer to)?

Indeed, we can say that as $n \to \infty$, the sequence moves closer to 0. We say that 0 is *the limit* of the sequence.

Question

Does the sequence become equal to 0 at some point?

Interestingly, it does not: The numbers come arbitrarily close to 0 but they never actually become 0. This shows that the sequence may or may not eventually equal to its limit.

Definition

We say that $\{a_n\}$ converges to the number L (or that the number L is its *limit*), denoted as

$$\lim_{n\to\infty} a_n = L \qquad \text{(or } a_n \to L\text{)}$$

if as n becomes large enough, all terms a_n become arbitrarily close to L.

Definition

We say that $\{a_n\}$ converges to the number L (or that the number L is its *limit*), denoted as

$$\lim_{n\to\infty} a_n = L \qquad \text{(or } a_n \to L\text{)}$$

if as n becomes large enough, all terms a_n become arbitrarily close to L.

In other words, if you say:

"Will a_n be close enough to L after some point?"

we will say:

"Sure, how much close do you want it to be?"

Definition

We say that $\{a_n\}$ converges to the number L (or that the number L is its *limit*), denoted as

$$\lim_{n\to\infty} a_n = L \qquad \text{(or } a_n \to L\text{)}$$

if as n becomes large enough, all terms a_n become arbitrarily close to L.

In other words, if you say:

"Will a_n be close enough to L after some point?"

we will say:

"Sure, how much close do you want it to be?"

and then whatever number you say (e.g. "not further than 0.002"), we can point out some number N (say, N = 1000) such that after the $N^{\rm th}$ term, all others are close to L by 0.002, i.e.

$$|a_N - L| < 0.002$$
, $|a_{N+1} - L| < 0.002$, $|a_{N+2} - L| < 0.002$, ...

Aprikyan, Tarkhanyan Lecture 6 5 / 40

So more technically, $\lim_{n \to \infty} a_n = L$ means that

- for any positive number $\varepsilon > 0$
- there exists N large enough
- such that $|a_n L| < \varepsilon$ for all n = N, N+1, N+2, ...

So more technically, $\lim_{n \to \infty} a_n = L$ means that

- for any positive number $\varepsilon > 0$
- there exists N large enough
- such that $|a_n L| < \varepsilon$ for all $n = N, N + 1, N + 2, \dots$

Definition

Similarly, we say $\lim_{n\to\infty} a_n = +\infty$ if a_n becomes greater than any number after some number of steps.

So more technically, $\lim_{n \to \infty} a_n = L$ means that

- for any positive number $\varepsilon > 0$
- there exists N large enough
- such that $|a_n L| < \varepsilon$ for all n = N, N+1, N+2, ...

Definition

Similarly, we say $\lim_{n\to\infty} a_n = +\infty$ if a_n becomes greater than any number after some number of steps.

Definition

If $\{a_n\}$ has a **finite** limit, we say that it is *convergent*, otherwise it is *divergent*.

Aprikyan, Tarkhanyan Lecture 6 6 / 40

Example

In the previous examples,

• 1, 2, 3, 4, 5, ...

Example

In the previous examples,

• 1, 2, 3, 4, 5, ... $\to +\infty$

Example

- \bullet 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- \bullet 1, -1, 1, -1, 1, ...

Example

- 1, 2, 3, 4, 5, ... $\to +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit

Example

- 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, ...

Example

- 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, ... $\rightarrow +\infty$

Example

- 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, $\dots \to +\infty$
- 6, 6, 6, 6, 6, ...

Example

- 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, ... $\rightarrow +\infty$
- \bullet 6, 6, 6, 6, ... \rightarrow 6

Example

In the previous examples,

- 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, ... $\rightarrow +\infty$
- \bullet 6, 6, 6, 6, ... \rightarrow 6

In general,

• $n^k \to \infty$ for any k > 0

Example

In the previous examples,

- 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, ... $\rightarrow +\infty$
- \bullet 6, 6, 6, 6, ... \rightarrow 6

In general,

- $n^k \to \infty$ for any k > 0
- $\frac{1}{n^k} \to 0$ for any k > 0

Example

In the previous examples,

- 1, 2, 3, 4, 5, ... $\rightarrow +\infty$
- $1, -1, 1, -1, 1, \ldots$ does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, ... $\rightarrow +\infty$
- \bullet 6, 6, 6, 6, ... \rightarrow 6

In general,

- $n^k \to \infty$ for any k > 0
- $\frac{1}{n^k} \to 0$ for any k > 0
- $c^n \to +\infty$ if c > 1, but $c^n \to 0$ if |c| < 1

Example

In the previous examples,

- 1, 2, 3, 4, 5, ... $\to +\infty$
- \bullet 1, -1, 1, -1, 1, ... does not have a limit
- 0, 0.2, 0.4, 0.6, 0.8, ... $\rightarrow +\infty$
- \bullet 6, 6, 6, 6, ... \rightarrow 6

In general,

- $n^k \to \infty$ for any k > 0
- $\frac{1}{n^k} \to 0$ for any k > 0
- $c^n \to +\infty$ if c > 1, but $c^n \to 0$ if |c| < 1
- If a sequence consists of the same number (or if it becomes constant starting from some point), the limit is that number.

More examples (we will not go further into details):

Example

Consider the sequence $\{\frac{1}{n}\}$. We claim that $\lim_{n\to\infty}\frac{1}{n}=0$.

More examples (we will not go further into details):

Example

Consider the sequence $\{\frac{1}{n}\}$. We claim that $\lim_{n\to\infty}\frac{1}{n}=0$.

Proof: For any $\varepsilon > 0$, choose N such that $\frac{1}{N} < \varepsilon$. Then, for all $n \ge N$, we have

$$\left|\frac{1}{n}-0\right|=\frac{1}{n}<\frac{1}{N}<\varepsilon.$$

Therefore, the sequence converges to 0, as *n* approaches infinity.

More examples (we will not go further into details):

Example

Consider the sequence $\{\frac{1}{n}\}$. We claim that $\lim_{n\to\infty}\frac{1}{n}=0$.

Proof: For any $\varepsilon > 0$, choose N such that $\frac{1}{N} < \varepsilon$. Then, for all $n \ge N$, we have

$$\left|\frac{1}{n}-0\right|=\frac{1}{n}<\frac{1}{N}<\varepsilon.$$

Therefore, the sequence converges to 0, as n approaches infinity.

Example

Consider the sequence $\left\{ \left(\frac{2}{3}\right)^n \right\}$. $\lim_{n \to \infty} \left(\frac{2}{3}\right)^n = 0$.

More examples (we will not go further into details):

Example

Consider the sequence $\{\frac{1}{n}\}$. We claim that $\lim_{n\to\infty}\frac{1}{n}=0$.

Proof: For any $\varepsilon > 0$, choose N such that $\frac{1}{N} < \varepsilon$. Then, for all $n \ge N$, we have

$$\left|\frac{1}{n}-0\right|=\frac{1}{n}<\frac{1}{N}<\varepsilon.$$

Therefore, the sequence converges to 0, as n approaches infinity.

Example

Consider the sequence $\left\{ \left(\frac{2}{3}\right)^n \right\}$. $\lim_{n \to \infty} \left(\frac{2}{3}\right)^n = 0$.

Example

Consider the sequence $\{0.3n\}$. $\lim_{n\to\infty} 0.3n = \infty$ (it is divergent).

Properties

$$\lim_{n \to \infty} (a_n + b_n) = L + M$$

$$\lim_{n \to \infty} (a_n - b_n) = L - M$$

$$\lim_{n \to \infty} (a_n \cdot b_n) = L \cdot M$$

$$\lim_{n \to \infty} \frac{a_n}{b_n} = \frac{L}{M} \qquad (\text{if } M \neq 0)$$

Properties

① If $\lim_{n\to\infty} a_n = L$ and $\lim_{n\to\infty} b_n = M$, then

$$\lim_{n \to \infty} (a_n + b_n) = L + M$$

$$\lim_{n \to \infty} (a_n - b_n) = L - M$$

$$\lim_{n \to \infty} (a_n \cdot b_n) = L \cdot M$$

$$\lim_{n \to \infty} \frac{a_n}{b_n} = \frac{L}{M} \qquad (\text{if } M \neq 0)$$

② If $\lim_{n\to\infty} a_n = L$, then for any constant c,

$$\lim_{n\to\infty}(c\cdot a_n)=c\cdot L.$$

Properties

① If $\lim_{n\to\infty} a_n = L$ and $\lim_{n\to\infty} b_n = M$, then

$$\lim_{n \to \infty} (a_n + b_n) = L + M$$

$$\lim_{n \to \infty} (a_n - b_n) = L - M$$

$$\lim_{n \to \infty} (a_n \cdot b_n) = L \cdot M$$

$$\lim_{n \to \infty} \frac{a_n}{b_n} = \frac{L}{M} \qquad (\text{if } M \neq 0)$$

② If $\lim_{n\to\infty} a_n = L$, then for any constant c,

$$\lim_{n\to\infty}(c\cdot a_n)=c\cdot L.$$

Now that we have the notion of

$$\lim_{n\to\infty} \left(2+\frac{1}{n}\right)^3$$

what do you think the expression

$$\lim_{x\to 0} (2+x)^3$$

would mean?

Now that we have the notion of

$$\lim_{n\to\infty} \left(2+\frac{1}{n}\right)^3$$

what do you think the expression

$$\lim_{x\to 0} (2+x)^3$$

would mean?

Similarly to sequences, we can define the limit of the above expression, i.e. of the function

$$f(x)=(2+x)^3,$$

as x approaches 0.

How do we do that?

Now that we have the notion of

$$\lim_{n\to\infty} \left(2+\frac{1}{n}\right)^3$$

what do you think the expression

$$\lim_{x\to 0} (2+x)^3$$

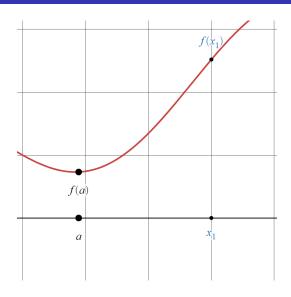
would mean?

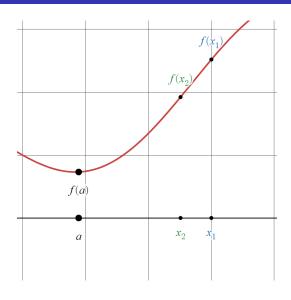
Similarly to sequences, we can define the limit of the above expression, i.e. of the function

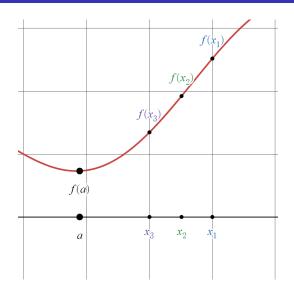
$$f(x)=(2+x)^3,$$

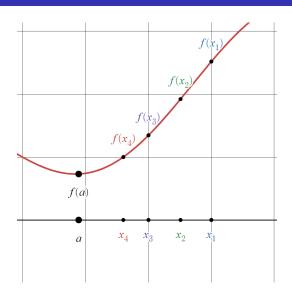
as x approaches 0.

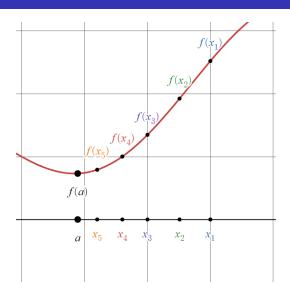
How do we do that? We can say: take any sequence x_n that converges to a, calculate the values of f(x) at x_1, x_2, \ldots , and see what happens.

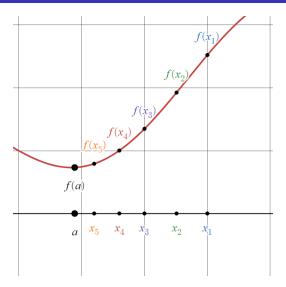






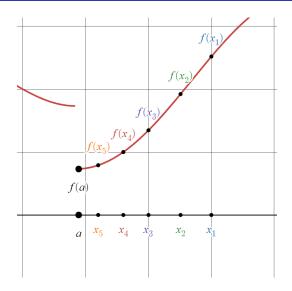






If numbers $f(x_1)$, $f(x_2)$, ... approach some limit L, then $\lim_{x\to a}f(x)=L$

Aprikyan, Tarkhanyan Lecture 6 15/40



It may happen that if $x_n \to a$ from the other side, we get another "limit".

In the second case, we say that the limit does not exist and the function is *discontinuous* at that point.

In the second case, we say that the limit does not exist and the function is *discontinuous* at that point.

Definition

If for **all** sequences $x_n \to a$ (no matter from left or right), the sequence

$$f(x_1), f(x_2), f(x_3), \ldots, f(x_n), \ldots$$

converges to a certain number L, then we say

$$\lim_{x\to a}f(x)=L$$

In the second case, we say that the limit does not exist and the function is *discontinuous* at that point.

Definition

If for **all** sequences $x_n \to a$ (no matter from left or right), the sequence

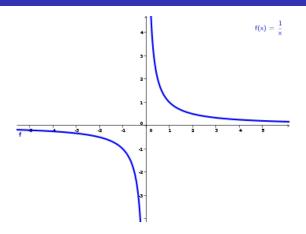
$$f(x_1), f(x_2), f(x_3), \ldots, f(x_n), \ldots$$

converges to a certain number L, then we say

$$\lim_{x\to a} f(x) = L$$

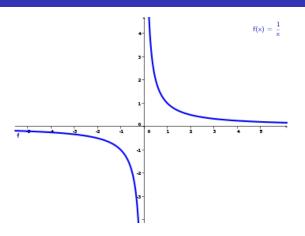
In other words, if the value of f(x) always approaches L, as its input approaches a.

Aprikyan, Tarkhanyan Lecture 6 17/40



Example

$$\lim_{x \to 3} \frac{1}{x} = \frac{1}{3}$$

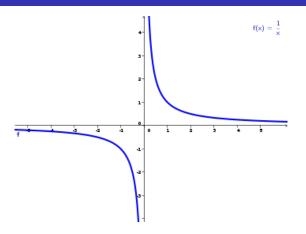


Example

$$\lim_{x \to 3} \frac{1}{x} = \frac{1}{3}$$

but

$$\lim_{x\to 0}\frac{1}{x}$$



Example

$$\lim_{x \to 3} \frac{1}{x} = \frac{1}{3}$$

but

$$\lim_{x \to 0} \frac{1}{x}$$
 does not exist.

Notice how the graph of $f(x) = \frac{1}{x}$ looks like it consists of two separate graphs put together at x = 0.

Notice how the graph of $f(x) = \frac{1}{x}$ looks like it consists of two separate graphs put together at x = 0.

You can draw its graph without lifting pencil on $(-\infty,0)$ and $(0,+\infty)$ but not at 0. That's why we say that it is **continuous** all points except for 0. We say that it is **discontinuous** at 0.

Notice how the graph of $f(x) = \frac{1}{x}$ looks like it consists of two separate graphs put together at x = 0.

You can draw its graph without lifting pencil on $(-\infty,0)$ and $(0,+\infty)$ but not at 0. That's why we say that it is **continuous** all points except for 0. We say that it is **discontinuous** at 0.

Definition

A function f(x) is said to be *continuous* at the point c if it is defined at c and

- $\lim_{x \to c} f(x) \text{ exists,}$
- $\lim_{x\to c} f(x) = f(c) \text{ (i.e. limit} = \text{value)}.$

Notice how the graph of $f(x) = \frac{1}{x}$ looks like it consists of two separate graphs put together at x = 0.

You can draw its graph without lifting pencil on $(-\infty,0)$ and $(0,+\infty)$ but not at 0. That's why we say that it is **continuous** all points except for 0. We say that it is **discontinuous** at 0.

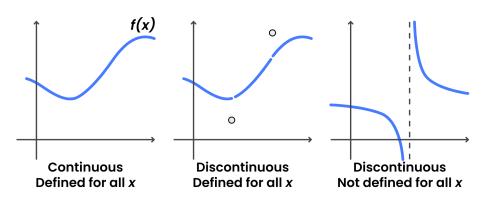
Definition

A function f(x) is said to be *continuous* at the point c if it is defined at c and

- $\lim_{x \to c} f(x) \text{ exists,}$
- $\lim_{x \to c} f(x) = f(c) \text{ (i.e. limit} = \text{value)}.$

If a function is continuous at all points, it is called a **continuous function**.

Aprikyan, Tarkhanyan Lecture 6 19 / 40



Properties

If f and g are continuous at some point a, then

• f + g

Properties

- f + g
- f − g

Properties

- f + g
- f g
- fg

Properties

- \bullet f + g
- f − g
- fg• $\frac{f}{g}$ (if $g(a) \neq 0$)

Properties

- f + g
- f g
- fg
- $\frac{f}{g}$ (if $g(a) \neq 0$)
- cf for any scalar c

Properties

If f and g are continuous at some point a, then

- f + g
- f g
- fg
- $\frac{f}{g}$ (if $g(a) \neq 0$)
- cf for any scalar c

are also continuous at a.

Properties

If f and g are continuous at some point a, then

- f + g
- f g
- fg
- $\frac{f}{g}$ (if $g(a) \neq 0$)
- cf for any scalar c

are also continuous at a.

In fact, most "good" functions are continuous (in their domains!):

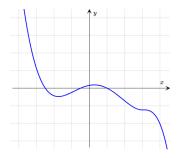
- Polynomials (e.g. $x^2 + 7x 1$, $xy y^4 + z$)
- Root functions (e.g. \sqrt{x} , $\sqrt[5]{x}$)
- Exponential and logarithmic functions (e.g. 2^x , e^{3x} , $\ln x$)
- Trigonometric functions and their inverses (e.g. cos(3x), arcsin x)

Derivative

Continuous functions are better and more preferable to work with (continuous $= \heartsuit$), and many processes in the world are actually described by continuous functions.

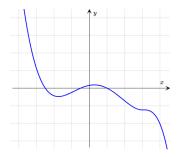
Derivative

Continuous functions are better and more preferable to work with (continuous $= \heartsuit$), and many processes in the world are actually described by continuous functions.



Now assume we want to maximize or minimize this function.

Continuous functions are better and more preferable to work with (continuous $= \heartsuit$), and many processes in the world are actually described by continuous functions.



Now assume we want to maximize or minimize this function. Notice how at some points it changes "faster" than at the others. How can we measure that?

First, we fix a point where we want to measure the "speed" of the function, say a.

First, we fix a point where we want to measure the "speed" of the function, say a. If we move now slightly to the left or right of a, say by some small amount h, the function will change by

$$f(a+h)-f(a)$$

First, we fix a point where we want to measure the "speed" of the function, say a. If we move now slightly to the left or right of a, say by some small amount h, the function will change by

$$f(a+h)-f(a)$$

If we now divide it by the change we made to a, we get

$$\frac{f(a+h)-f(a)}{h}$$

First, we fix a point where we want to measure the "speed" of the function, say a. If we move now slightly to the left or right of a, say by some small amount h, the function will change by

$$f(a+h)-f(a)$$

If we now divide it by the change we made to a, we get

$$\frac{f(a+h)-f(a)}{h}$$

Definition

If the speed of f is bounded by some constant M, i.e.

$$\left|\frac{f(a+h)-f(a)}{h}\right|\leq M$$

in all points a of its domain, then we say that f is Lipschitz continuous.

If this number is large, it means that the function changes a lot when we change a a little, so the function is "fast" at a.

$$\frac{f(a+h)-f(a)}{h}$$

If this number is large, it means that the function changes a lot when we change a a little, so the function is "fast" at a.

$$\frac{f(a+h)-f(a)}{h}$$

Intuitively, this value should express how quickly or slowly the function changes at the point a. We just have to make sure that h is really small enough (so the values of f far from a do not affect it):

Aprikyan, Tarkhanyan Lecture 6 24 / 40

If this number is large, it means that the function changes a lot when we change a a little, so the function is "fast" at a.

$$\frac{f(a+h)-f(a)}{h}$$

Intuitively, this value should express how quickly or slowly the function changes at the point a. We just have to make sure that h is really small enough (so the values of f far from a do not affect it):

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

24 / 40

If this number is large, it means that the function changes a lot when we change a a little, so the function is "fast" at a.

$$\frac{f(a+h)-f(a)}{h}$$

Intuitively, this value should express how quickly or slowly the function changes at the point a. We just have to make sure that h is really small enough (so the values of f far from a do not affect it):

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}=f'(a)$$

Definition

We say f(x) is differentiable at a point a, if the following limit

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

or, equivalently,

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists and is finite. We denote it by f'(a) or $\frac{df}{dx}(a)$ and call the *derivative* of function f(x) at a.

Definition

We say f(x) is differentiable at a point a, if the following limit

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

or, equivalently,

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists and is finite. We denote it by f'(a) or $\frac{df}{dx}(a)$ and call the *derivative* of function f(x) at a.

The derivative of a function at a given point shows the "speed" of the function at that point, i.e. the sensitivity of change of its output with respect to its input.

Definition

We say f(x) is differentiable at a point a, if the following limit

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

or, equivalently,

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists and is finite. We denote it by f'(a) or $\frac{df}{dx}(a)$ and call the *derivative* of function f(x) at a.

The derivative of a function at a given point shows the "speed" of the function at that point, i.e. the sensitivity of change of its output with respect to its input.

Note that f'(x) is a **function** itself and not a fixed number!

Aprikyan, Tarkhanyan Lecture 6 25 / 40

Example

Consider the function $f(x) = x^2$. The derivative of f(x) is:

$$f'(x) =$$

Example

Consider the function $f(x) = x^2$. The derivative of f(x) is:

$$f'(x) = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} =$$

Example

Consider the function $f(x) = x^2$. The derivative of f(x) is:

$$f'(x) = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h}$$

Example

Consider the function $f(x) = x^2$. The derivative of f(x) is:

$$f'(x) = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} = 2x$$

Thus, f(x) is differentiable at any point, and f'(x) = 2x.

Example

Consider the function $f(x) = x^2$. The derivative of f(x) is:

$$f'(x) = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} = 2x$$

Thus, f(x) is differentiable at any point, and f'(x) = 2x.

Remark

If a function is differentiable at some point, then it is also continuous, but the reverse is not always true.

Example

Consider the function $f(x) = x^2$. The derivative of f(x) is:

$$f'(x) = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} = 2x$$

Thus, f(x) is differentiable at any point, and f'(x) = 2x.

Remark

If a function is differentiable at some point, then it is also continuous, but the reverse is not always true.

Example

The function f(x) = |x| is continuous but it is not differentiable at point x = 0.

Aprikyan, Tarkhanyan Lecture 6 26 / 40

Similarly, we can compute the derivative of f'(x) itself (it will show the speed of the speed of f(x), i.e. its *acceleration*).

We denote the derivative of f'(x) by f''(x), that of f''(x) by f'''(x), and so on. The derivative taken of f(x) n times is also denoted by $f^{(n)}(x)$.

Similarly, we can compute the derivative of f'(x) itself (it will show the speed of the speed of f(x), i.e. its acceleration).

We denote the derivative of f'(x) by f''(x), that of f''(x) by f'''(x), and so on. The derivative taken of f(x) n times is also denoted by $f^{(n)}(x)$. For any differentiable functions f,g and a real number c, the following rules hold:

Similarly, we can compute the derivative of f'(x) itself (it will show the speed of the speed of f(x), i.e. its acceleration).

We denote the derivative of f'(x) by f''(x), that of f''(x) by f'''(x), and so on. The derivative taken of f(x) n times is also denoted by $f^{(n)}(x)$. For any differentiable functions f,g and a real number c, the following rules hold:

- $(cf)' = c \cdot f'$
- ② $(f \pm g)' = f' \pm g'$

Similarly, we can compute the derivative of f'(x) itself (it will show the speed of the speed of f(x), i.e. its acceleration).

We denote the derivative of f'(x) by f''(x), that of f''(x) by f'''(x), and so on. The derivative taken of f(x) n times is also denoted by $f^{(n)}(x)$. For any differentiable functions f,g and a real number c, the following rules hold:

- ② $(f \pm g)' = f' \pm g'$
- **3** (fg)' = f'g + fg'

Similarly, we can compute the derivative of f'(x) itself (it will show the speed of the speed of f(x), i.e. its acceleration).

We denote the derivative of f'(x) by f''(x), that of f''(x) by f'''(x), and so on. The derivative taken of f(x) n times is also denoted by $f^{(n)}(x)$. For any differentiable functions f,g and a real number c, the following rules hold:

②
$$(f \pm g)' = f' \pm g'$$

$$(fg)' = f'g + fg'$$

Similarly, we can compute the derivative of f'(x) itself (it will show the speed of the speed of f(x), i.e. its acceleration).

We denote the derivative of f'(x) by f''(x), that of f''(x) by f'''(x), and so on. The derivative taken of f(x) n times is also denoted by $f^{(n)}(x)$. For any differentiable functions f,g and a real number c, the following rules hold:

②
$$(f \pm g)' = f' \pm g'$$

3
$$(fg)' = f'g + fg'$$

Similarly, we can compute the derivative of f'(x) itself (it will show the speed of the speed of f(x), i.e. its acceleration).

We denote the derivative of f'(x) by f''(x), that of f''(x) by f'''(x), and so on. The derivative taken of f(x) n times is also denoted by $f^{(n)}(x)$. For any differentiable functions f,g and a real number c, the following rules hold:

Properties

②
$$(f \pm g)' = f' \pm g'$$

3
$$(fg)' = f'g + fg'$$

$$(f(g(x)))' = f'(g(x)) \cdot g'(x)$$

Aprikyan, Tarkhanyan Lecture 6 27

• The derivative of any constant f(x) = c is:

$$(c)'=0$$

• The derivative of any constant f(x) = c is:

$$(c)'=0$$

• The derivative of f(x) = x is:

$$(x)'=1$$

• The derivative of any constant f(x) = c is:

$$(c)'=0$$

• The derivative of f(x) = x is:

$$(x)' = 1$$

• The derivative of $f(x) = x^2$ is:

$$(x^2)'=2x$$

• The derivative of any constant f(x) = c is:

$$(c)'=0$$

• The derivative of f(x) = x is:

$$(x)' = 1$$

• The derivative of $f(x) = x^2$ is:

$$(x^2)'=2x$$

• For any constant n, the derivative of $f(x) = x^n$ is:

$$(x^n)' = nx^{n-1}$$

• The derivative of $f(x) = e^x$ is:

$$(e^x)'=e^x$$

• The derivative of $f(x) = e^x$ is:

$$(e^x)'=e^x$$

• The derivative of $f(x) = a^x$ is:

$$(a^x)'=a^x\ln a$$

• The derivative of $f(x) = e^x$ is:

$$(e^x)'=e^x$$

• The derivative of $f(x) = a^x$ is:

$$(a^{x})'=a^{x}\ln a$$

• The derivative of $f(x) = \ln x$ is:

$$(\ln x)' = \frac{1}{x}$$

• The derivative of $f(x) = e^x$ is:

$$(e^x)'=e^x$$

• The derivative of $f(x) = a^x$ is:

$$(a^x)'=a^x \ln a$$

• The derivative of $f(x) = \ln x$ is:

$$(\ln x)' = \frac{1}{x}$$

• The derivative of $f(x) = \sin x$ is:

$$(\sin x)' = \cos x$$

• The derivative of $f(x) = e^x$ is:

$$(e^x)'=e^x$$

• The derivative of $f(x) = a^x$ is:

$$(a^x)' = a^x \ln a$$

• The derivative of $f(x) = \ln x$ is:

$$(\ln x)' = \frac{1}{x}$$

• The derivative of $f(x) = \sin x$ is:

$$(\sin x)' = \cos x$$

• The derivative of $f(x) = \cos x$ is:

$$(\cos x)' = -\sin x$$

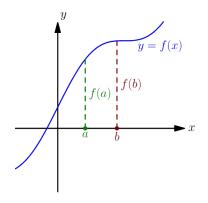
Derivative is All You Need

These formulas might seem to much, but you do not have to memorize them—after using them for a while, one begins to "feel" how fast or slow a given function is.

Derivative is All You Need

These formulas might seem to much, but you do not have to memorize them—after using them for a while, one begins to "feel" how fast or slow a given function is.

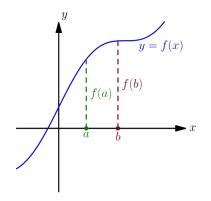
Derivatives tell about amazingly many interesting properties of the function.



We say that a function f(x) is increasing at the point a if

$$f(c) < f(a) < f(b)$$
 when $c < a < b$

Aprikyan, Tarkhanyan



We say that a function f(x) is increasing at the point a if

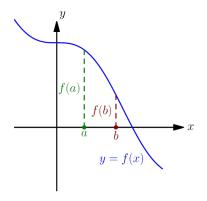
$$f(c) < f(a) < f(b)$$
 when $c < a < b$

and we say that it is just non-decreasing if

$$f(c) \le f(a) \le f(b)$$
 when $c < a < b$

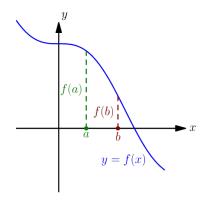
Aprikyan, Tarkhanyan Lecture 6

31 / 40



Similarly, we say that f(x) is decreasing at the point a if

$$f(c) > f(a) > f(b)$$
 when $c < a < b$



Similarly, we say that f(x) is decreasing at the point a if

$$f(c) > f(a) > f(b)$$
 when $c < a < b$

and we say that it is just non-increasing if

$$f(c) \ge f(a) \ge f(b)$$
 when $c < a < b$

If your profit function f(x) is increasing at some point, you might want to consider increasing x. How do you know whether f is increasing?

If your profit function f(x) is increasing at some point, you might want to consider increasing x. How do you know whether f is increasing?

Theorem

• If f'(a) > 0 then f is increasing at a.

If your profit function f(x) is increasing at some point, you might want to consider increasing x. How do you know whether f is increasing?

Theorem

- If f'(a) > 0 then f is increasing at a.
- If $f'(a) \ge 0$ then f is non-decreasing at a.

If your profit function f(x) is increasing at some point, you might want to consider increasing x. How do you know whether f is increasing?

Theorem

- If f'(a) > 0 then f is increasing at a.
- If $f'(a) \ge 0$ then f is non-decreasing at a.
- If f'(a) < 0 then f is decreasing at a.

If your profit function f(x) is increasing at some point, you might want to consider increasing x. How do you know whether f is increasing?

Theorem

- If f'(a) > 0 then f is increasing at a.
- If $f'(a) \ge 0$ then f is non-decreasing at a.
- If f'(a) < 0 then f is decreasing at a.
- If $f'(a) \le 0$ then f is non-increasing at a.

If your profit function f(x) is increasing at some point, you might want to consider increasing x. How do you know whether f is increasing?

Theorem

- If f'(a) > 0 then f is increasing at a.
- If $f'(a) \ge 0$ then f is non-decreasing at a.
- If f'(a) < 0 then f is decreasing at a.
- If $f'(a) \le 0$ then f is non-increasing at a.

That simple.

If your profit function f(x) is increasing at some point, you might want to consider increasing x. How do you know whether f is increasing?

Theorem

- If f'(a) > 0 then f is increasing at a.
- If $f'(a) \ge 0$ then f is non-decreasing at a.
- If f'(a) < 0 then f is decreasing at a.
- If $f'(a) \le 0$ then f is non-increasing at a.

That simple.

Question

What if f'(a) = 0?

Aprikyan, Tarkhanyan

Finally, what about the largest/smallest values of the function?

Finally, what about the largest/smallest values of the function?

Definition

We say x_0 is a *point of local maximum (minimum)* for f, if the largest (smallest) value of f on a small enough interval $(x_0 - \delta, x_0 + \delta)$ is the value $f(x_0)$ itself.

Finally, what about the largest/smallest values of the function?

Definition

We say x_0 is a *point of local maximum (minimum)* for f, if the largest (smallest) value of f on a small enough interval $(x_0 - \delta, x_0 + \delta)$ is the value $f(x_0)$ itself.

More technically, if there exists small enough $\delta > 0$ such that

$$f(x) \leq f(x_0)$$

$$(f(x) \geq f(x_0))$$

for any x from $(x_0 - \delta, x_0 + \delta)$.

Finally, what about the largest/smallest values of the function?

Definition

We say x_0 is a *point of local maximum (minimum)* for f, if the largest (smallest) value of f on a small enough interval $(x_0 - \delta, x_0 + \delta)$ is the value $f(x_0)$ itself.

More technically, if there exists small enough $\delta > 0$ such that

$$f(x) \leq f(x_0)$$

$$(f(x) \geq f(x_0))$$

for any x from $(x_0 - \delta, x_0 + \delta)$.

In that case, the value $f(x_0)$ is called a **local maximum (minimum)** of f.

34 / 40

Finally, what about the largest/smallest values of the function?

Definition

We say x_0 is a *point of local maximum (minimum)* for f, if the largest (smallest) value of f on a small enough interval $(x_0 - \delta, x_0 + \delta)$ is the value $f(x_0)$ itself.

More technically, if there exists small enough $\delta > 0$ such that

$$f(x) \leq f(x_0)$$

$$(f(x) \geq f(x_0))$$

for any x from $(x_0 - \delta, x_0 + \delta)$.

In that case, the value $f(x_0)$ is called a **local maximum (minimum)** of f. Together, local minima and maxima are called local extrema of f.

Definition

We say x_0 is a *point of global maximum (minimum)* for f, if the largest (smallest) value of f (on whole \mathbb{R}) is the value $f(x_0)$ itself.

Technically, x_0 is a global maximum (minimum) point if for any x,

$$f(x) \leq f(x_0)$$

$$(f(x) \geq f(x_0))$$

Definition

We say x_0 is a *point of global maximum (minimum)* for f, if the largest (smallest) value of f (on whole \mathbb{R}) is the value $f(x_0)$ itself.

Technically, x_0 is a global maximum (minimum) point if for any x,

$$f(x) \leq f(x_0)$$

$$(f(x) \geq f(x_0))$$

Together, global minima and maxima are called global extrema of f. Apparently, every global extrema is also a local extrema (and the converse is not true).

Definition

We say x_0 is a *point of global maximum (minimum)* for f, if the largest (smallest) value of f (on whole \mathbb{R}) is the value $f(x_0)$ itself.

Technically, x_0 is a global maximum (minimum) point if for any x,

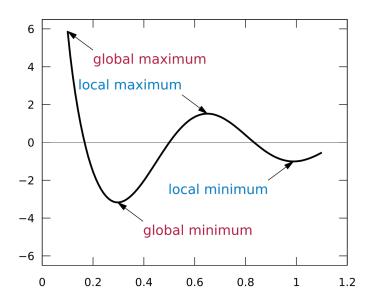
$$f(x) \leq f(x_0)$$

$$(f(x) \geq f(x_0))$$

Together, global minima and maxima are called global extrema of f. Apparently, every global extrema is also a local extrema (and the converse is not true).

Theorem

Every continuous function f has both a global maximum and a global minimum on any **closed** interval [a, b].



Theorem

If x_0 is an extremum point of f and there exists $f'(x_0)$, then $f'(x_0) = 0$.

Theorem

If x_0 is an extremum point of f and there exists $f'(x_0)$, then $f'(x_0) = 0$.

Example

Consider the function $f_1(x) = x^2$ which has a minimum at x = 0.

$$f_1'(x) = 2x, \qquad f_1'(0) = 0$$

Theorem

If x_0 is an extremum point of f and there exists $f'(x_0)$, then $f'(x_0) = 0$.

Example

Consider the function $f_1(x) = x^2$ which has a minimum at x = 0.

$$f_1'(x) = 2x, \qquad f_1'(0) = 0$$

Moreover, the function $f_2(x) = -x^2$ has a maximum at x = 0 and again:

$$f_2'(x) = -2x, \qquad f_2'(0) = 0$$

Theorem

If x_0 is an extremum point of f and there exists $f'(x_0)$, then $f'(x_0) = 0$.

Example

Consider the function $f_1(x) = x^2$ which has a minimum at x = 0.

$$f_1'(x) = 2x, \qquad f_1'(0) = 0$$

Moreover, the function $f_2(x) = -x^2$ has a maximum at x = 0 and again:

$$f_2'(x) = -2x, \qquad f_2'(0) = 0$$

The function $f_3(x) = x^3$ has no local extremum point, yet

$$f_3'(x) = 3x^2, \qquad f_3'(0) = 0$$

37 / 40

Theorem

If x_0 is an extremum point of f and there exists $f'(x_0)$, then $f'(x_0) = 0$.

Example

Consider the function $f_1(x) = x^2$ which has a minimum at x = 0.

$$f_1'(x) = 2x, \qquad f_1'(0) = 0$$

Moreover, the function $f_2(x) = -x^2$ has a maximum at x = 0 and again:

$$f_2'(x) = -2x, \qquad f_2'(0) = 0$$

The function $f_3(x) = x^3$ has no local extremum point, yet

$$f_3'(x) = 3x^2, \qquad f_3'(0) = 0$$

Hence, the condition f'(x) = 0 is necessary but not sufficient.

How to find the actual local extrema?

How to find the actual local extrema?

Definition

If $f'(x_0)$ doesn't exist or $f'(x_0) = 0$, then we call x_0 a *critical point*.

How to find the actual local extrema?

Definition

If $f'(x_0)$ doesn't exist or $f'(x_0) = 0$, then we call x_0 a *critical point*.

How can we tell if a critical point is a local minimum/maximum point?

Theorem 1 (f'' at one point)

If $f'(x_0) = 0$ and there exists finite $f''(x_0)$, then

- If $f''(x_0) > 0$, then x_0 is a local minimum point,
- 2 If $f''(x_0) < 0$, then x_0 is a local maximum point.

Theorem 1 (f'' at one point)

If $f'(x_0) = 0$ and there exists finite $f''(x_0)$, then

- ① If $f''(x_0) > 0$, then x_0 is a local minimum point,
- ② If $f''(x_0) < 0$, then x_0 is a local maximum point.

Theorem 2 (f' at multiple points)

If for some $\delta > 0$, f is differentiable in the intervals $(x_0 - \delta, x_0)$ and $(x_0, x_0 + \delta)$ and continuous at x_0 , then

- If f'(x) > 0 for $x \in (x_0 \delta, x_0)$ and f'(x) < 0 for $x \in (x_0, x_0 + \delta)$, then x_0 is a local maximum point.
- ② If f'(x) < 0 for $x \in (x_0 \delta, x_0)$ and f'(x) > 0 for $x \in (x_0, x_0 + \delta)$, then x_0 is a local minimum point.
- **3** If f'(x) doesn't change its sign, then x_0 is not an extremum point.

Wrapping up, how can we use our knowledge to find the local extrema of a given function f(x)?

Wrapping up, how can we use our knowledge to find the local extrema of a given function f(x)?

Step 0: Make sure that f(x) is continuous in the given set. Otherwise, divide it into intervals where f(x) is continuous. Example: For $f(x) = \frac{1}{x}$ we should consider it separately on $(-\infty,0)$ and $(0,+\infty)$

Wrapping up, how can we use our knowledge to find the local extrema of a given function f(x)?

- Step 0: Make sure that f(x) is continuous in the given set. Otherwise, divide it into intervals where f(x) is continuous. Example: For $f(x) = \frac{1}{x}$ we should consider it separately on $(-\infty,0)$ and $(0,+\infty)$
- Step 1: Find the critical points x_0 of f(x).

Wrapping up, how can we use our knowledge to find the local extrema of a given function f(x)?

- Step 0: Make sure that f(x) is continuous in the given set. Otherwise, divide it into intervals where f(x) is continuous. Example: For $f(x) = \frac{1}{x}$ we should consider it separately on $(-\infty,0)$ and $(0,+\infty)$
- Step 1: Find the critical points x_0 of f(x).
- Step 2: If f(x) is given on a closed interval [a, b], also check its values on endpoints a and b.

Example: When defined on a closed interval [a, b], f(x) = 2x has its minimum at x = a and maximum at x = b

Wrapping up, how can we use our knowledge to find the local extrema of a given function f(x)?

- Step 0: Make sure that f(x) is continuous in the given set. Otherwise, divide it into intervals where f(x) is continuous. Example: For $f(x) = \frac{1}{x}$ we should consider it separately on $(-\infty,0)$ and $(0,+\infty)$
- Step 1: Find the critical points x_0 of f(x).
- Step 2: If f(x) is given on a closed interval [a, b], also check its values on endpoints a and b.

 Example: When defined on a closed interval [a, b], f(x) = 2x has its minimum at x = a and maximum at x = b
- Step 3: a) If there exists finite $f''(x_0) \neq 0$, use Theorem 1. b) If you find the sign of f'(x) on left and right "sides" of x_0 , use Theorem 2.