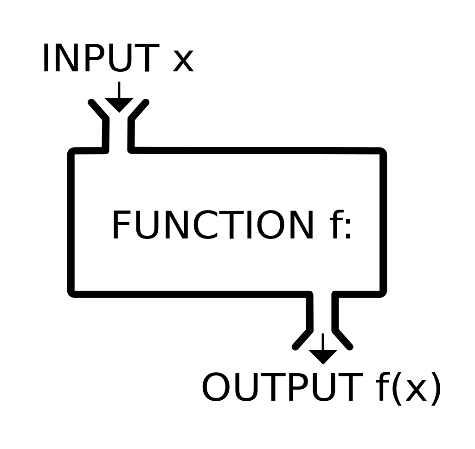
1. What is function?

It is a binary relation over two sets



1. Derivatives?

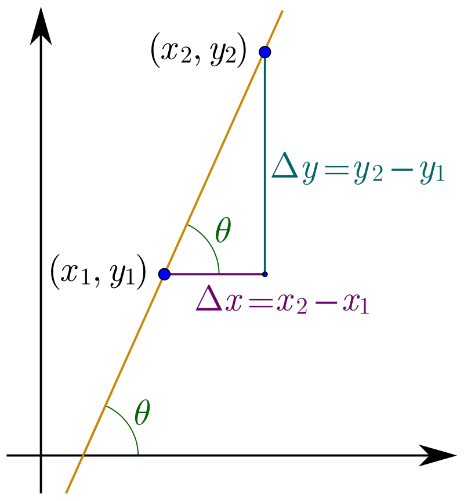
Derivative: measures the sensitivity to change of the function (output) with respect to a change in its argument (input)

1. What does positive/negative sign of the derivative show?

Positive sign of derivative shows that our function is increasing and negative sign of derivative shows that our function is decreasing

1. How to calculate it?

Using trigonometric way



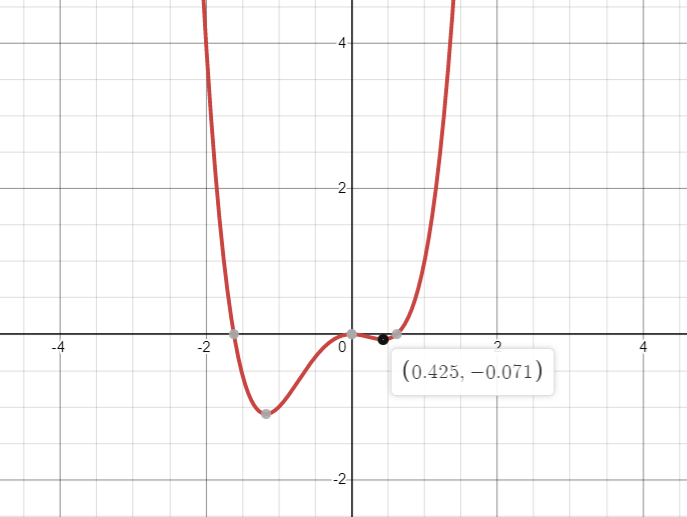
We calculate ∆x = x2 – x1

After we calculate ∆y = y2 – y1

And then we do a relation between them

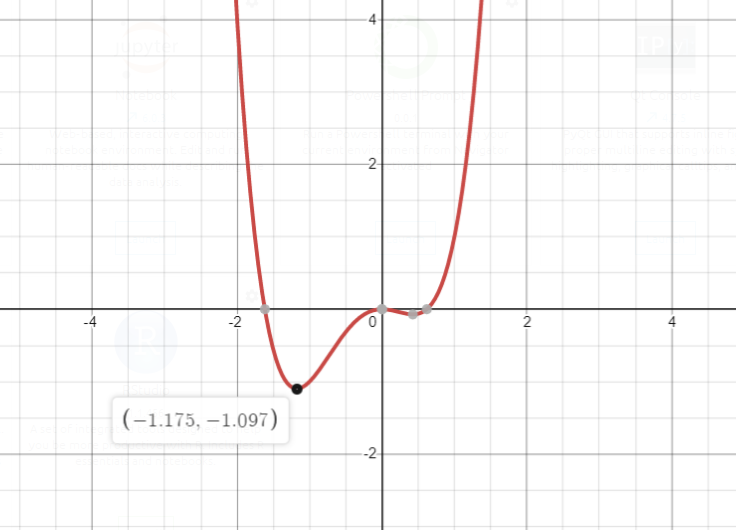
And we found the angle between then, it calls tangent(тангенс)

1. How many local mins the y=x^4 + x^3 - x^2 has?



This function has only 1 local minimum at (0.425; -0.071)

1. Find the global minimum



Global minimum is at (-1.175; -1.097)

1. What is the learning rate

Learning rate is a constant little number (for example 0.0001)

1. how to calculate the step

step = learning rate \* derivative

if our function changes fast we can change our step as well

1. When do we need to stop the function?

We need to stop gradient descent when our derivative of x is close to 0

1. What is y\_pred and y\_true?

Y\_pred is our prediction that we calculate using futures

Y\_true is always a given number, we can take it in our dataset

1. How to calculate it?

We can calculate our y\_pred for example like this:

A \* x + b \* y + c \* z

1. Linear regreassion

Linear regression is a model of the dependence of the variable x on one or several other variables (factors, independent variables) with a linear dependence function.

1. Df/Da and df/dx

Df/da means calculating derivative using parameters

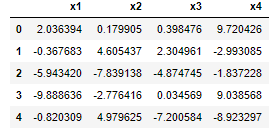
df/dx means calculating derivative using variables

we can calculate derivative using and variables, and parameters.

1. What are parameters and variables of linear regression

Parameters it is what we are looking for. We can find them using linear regreassion and gradient descent

Variables are given numbers that we get in our dataset



1. Error functions

Error function helps us to calculate how many errors does our prediction makes and how big are this errors, we need it in our gradient descent.

We have many error function, but this two are the most used:

* MSE(Mean squared error)

MSE = (Y\_pred - Y\_true) \*\* 2   
 squaring allows us to always have positive values and avoid minus infinity problem

We use mean squared error when our features are normally distributed

* MAE(Mean absolute error)

MAE = abs(y\_pred - y\_true)

\* abs(absolute) change the negative sign of number to positive

We use absolute error when our features are not normally distributed