CalcStudio Math Documentation

Aayush Koora

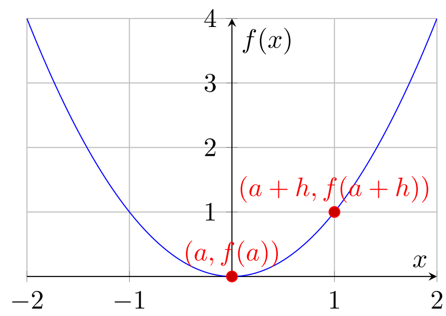
June 23, 2025

1 Derivative: Tangent Line at Point

A derivative is the instantaneous rate of change of a function. It answers how fast a function is changing at a specific point $x = a$.

This is the formal definition of a limit:

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{1}$$

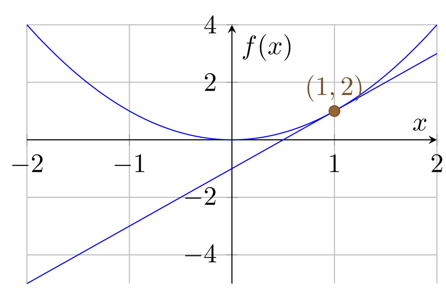


We can decrease h until it is infinitely small. As the distance between the two points (h) approaches 0, we get the instantaneous slope of $f(x)$ at point a.

A tangent line is a line that touches a function at one point and has the same slope as that instantaneous point.

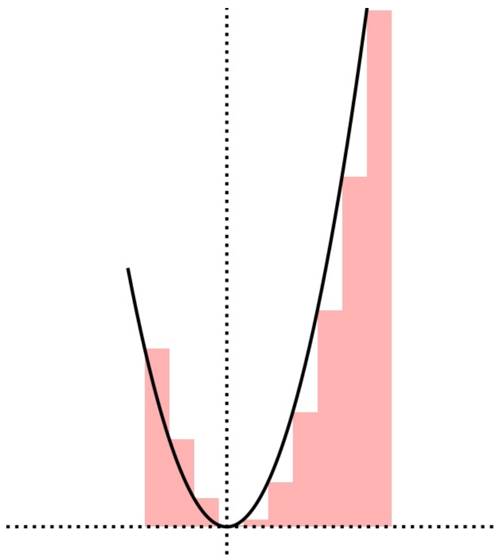
We can model a tangent line like this:

$$y = f(a) + f'(a)(x - a) \tag{2}$$

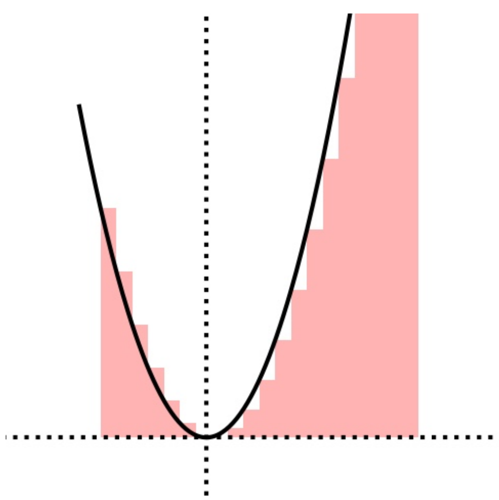


2 Integral: Area Under a Curve/Riemann Sums

An integral is the area under a curve. We can approximate the area under a curve using riemann sums.



As we use increasingly smaller rectangles, we can get closer to the actual value of the are under the curve.



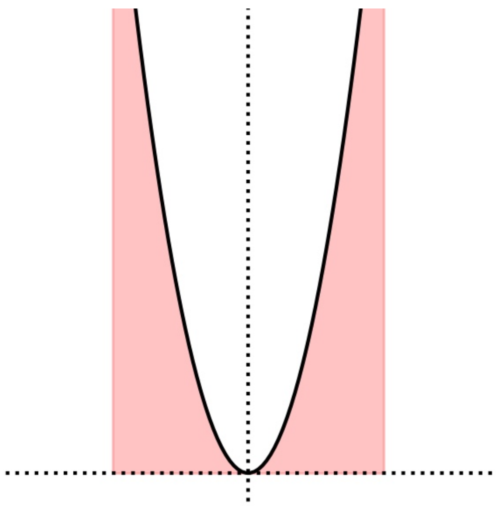
We can use a limit statement to find the area under curve using an infinite amount of rectangles.

This is the formal definition of an integral:

$$Area = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_i^*) \Delta x \tag{3}$$

This definition can be rewritten in this notation:

$$Area = \int_a^b f(x)\,dx \tag{4}$$



The integral finds the exact area under the curve.

3 Multivariable: Gradient Descent

Gradient descent is an interactive algorithm used to minimize the function by following the steepest descent. This algorithm is used in machine learning and can work with multiple parameters.

We can model this algorithm using the function $x^2 = y$

Let's start with a random x value: 5. The first step is to take the derivative of the function at x=5. This value is 10.

$$f'(x) = 2x \implies f'(5) = 10$$

The next step is to calculate the step size of the next value. By multiplying the learning rate times the derivative of the value.

$$\Delta x = \eta \times f'(x)$$

For example, if $\eta = 0.1$, then $\Delta x = 0.1 \times 10 = 1$.

The learning rate η determines how much of a step each iteration of the algorithm takes.

The last step is to update the current value by subtracting the step size:

$$x_{\text{new}} = x_{\text{old}} - \Delta x$$

Using our example:

$$x_{\text{new}} = 5 - 1 = 4$$

Now we repeat this process: calculate the derivative at the new x, multiply by the learning rate, update x, and so on, until the step size Δx is smaller than a threshold (e.g., 0.001) or a maximum number of iterations is reached.

This iterative procedure allows the algorithm to converge to the function's minimum.