WELCOME TO

TECH I.S.

Calculus

Introduction to Differential Calculus

Differential Calculus is the study of rate of change, and in Data Science it is used for optimization of algorithms.

Differential Calculus Basics

For a function f, the derivative of f is defined as:

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

We can also interpret the derivative $f'(x)$ as the instantaneous rate of change of $f(x)$ with respect to x.

Meaning: for a small change in x, what is the rate of change of $f(x)$

Basic Rules for Differential Calculus

The following rules are used for computing the derivatives of functions:

Derivative of constants. $\frac{d}{dx}c = 0.$

Derivative of linear functions. $\frac{d}{dx}(ax) = a.$

Power rule. $\frac{d}{dx}x^n = nx^{n-1}.$

Derivative of exponentials. $\frac{d}{dx}e^x = e^x.$

Derivative of the logarithm. $\frac{d}{dx}\log(x) = \frac{1}{x}.$

Sum rule. $\frac{d}{dx}\left(g(x) + h(x)\right) = \frac{dg}{dx}(x) + \frac{dh}{dx}(x).$

Product rule. $\frac{d}{dx}\left(g(x) \cdot h(x)\right) = g(x)\frac{dh}{dx}(x) + \frac{dg}{dx}(x)h(x).$

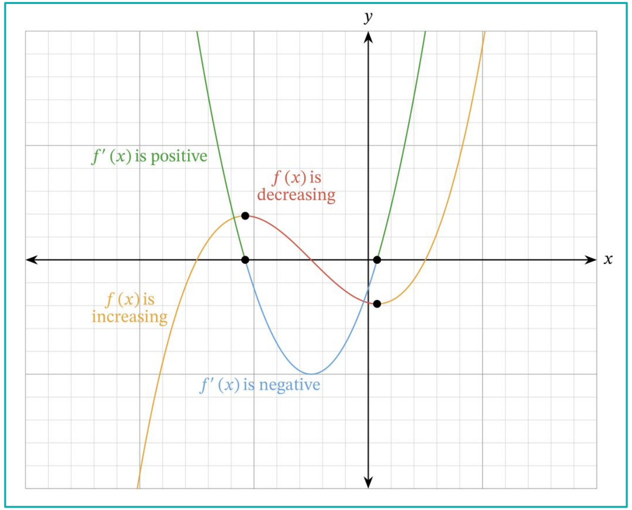
Chain rule. $\frac{d}{dx}g(h(x)) = \frac{dg}{dh}(h(x)) \cdot \frac{dh}{dx}(x).$

Higher Order Derivatives

The derivative of the first derivative of a function $f(x)$ is the second derivative $f''(x)$:

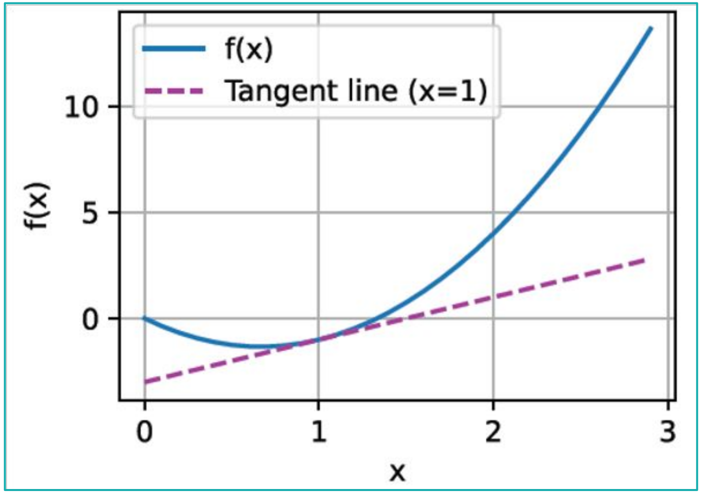
$$\frac{d^2f}{dx^2} = \frac{d}{dx}\left(\frac{df}{dx}\right)$$

Geometrically it means : Second derivative tells us if the rate of change of $f(x)$ is increasing or decreasing.



Geometric Interpretation of Derivatives

The first derivative of a function is also known as the slope of the tangent line.



Partial Derivatives

A Function can be dependent on many variables, these are called multivariate functions.

Let ($\mathbf{y}$) be a multivariate function with (n) variables and input (x),

$$y = f(\mathbf{x}) = f(x_1, x_2, \dots, x_n) \qquad \mathbf{x} = [x_1, x_2, \dots, x_n]^T$$

To calculate partial derivative ("∂ doh"),

we can treat all (x_n) as constants and calculate the derivative of ($\mathbf{y}$) only with respect to (x_i)

The partial derivative of ($\mathbf{y}$) with respect to (x_i) is given b:,

$$\frac{\partial y}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, x_2, \dots, x_i + h, \dots, x_n) - f(x_1, x_2, \dots, x_i, \dots, x_n)}{h}$$

Gradient

We can combine all the partial derivatives of a multivariate function we obtain the gradient vector of the function, denoted by “del” ,

$$\nabla_{\mathbf{x}} f(\mathbf{x}) = \left[\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right]^T$$

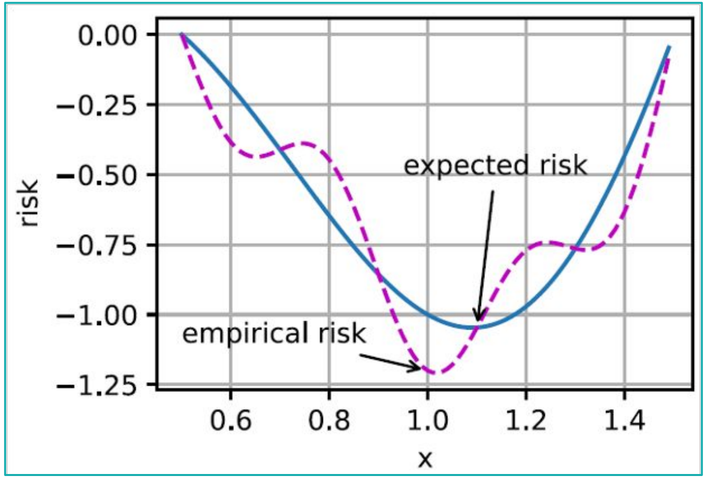
Use Case:

In ML, the gradient descent algorithm uses gradient of the loss function ($\mathscr{L}$) for minimizing the learning loss function.

Optimization

Optimization an objective function means — finding the values that minimizes or maximizes the cost/loss function.

ML algorithms attempt to find the minimum value that can minimize the error in prediction.



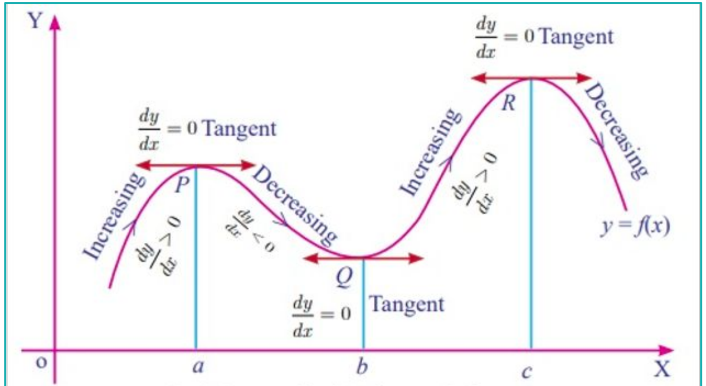
Optima: Minima & Maxima

A point where where the derivative of the function is zero, i.e., $f'(x)=0$ is called an Optima.

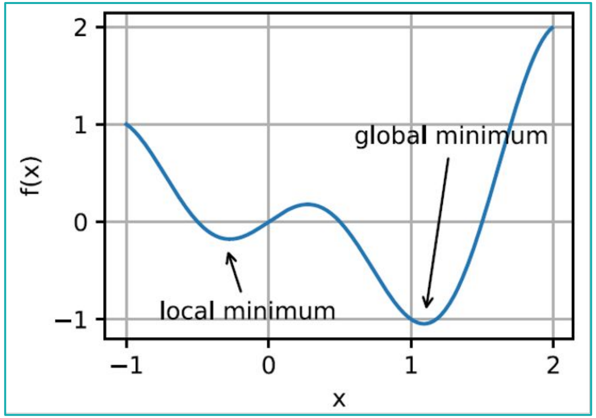
Minima, is an optima value at which function is minimum.

Maxima, is an optima point at which a function is maximum.

Differential calculus is used to reach optima points.



P, Q, R are called Stationary Points



Much obliged.

TECH I.S.