# *AI Essentials: Derivatives*

- Author: Christian M.M. Frey (web) - (email)
- Math Basics - Derivatives

## Content

## Intuition of the first derivative

The first derivative of a function $f(x)$, which we write as $f'(x)$ or as $\frac{df}{dx}$ is the slope of the tangent line to the function at the point $x$, i.e., the first derivative tells us to which amount a function is increasing or decreasing. A positive (negative) slope tells as that, as $x$ increases, $f(x)$ increases (decreases). A slope of zero indicates that we are at a saddle point (also known as minimax point) or at a local maximum, resp., local minimum at that point. To summarize, we have:

- $\frac{df}{dx}(p) > 0$, then $f(x)$ is an increasing function at $x = p$. This can also be written as $\left.\frac{df}{dx}\right|_{x=p} > 0$
- $\frac{df}{dx}(p) < 0$, then $f(x)$ is an decreasing function at $x = p$. This can also be written as $\left.\frac{df}{dx}\right|_{x=p} < 0$
- $\frac{df}{dx}(p) = 0$, then $x = p$ is called a critical point of $f(x)$. This can also be written as $\left.\frac{df}{dx}\right|_{x=p} = 0$

## Intuition of the second derivative

The second derivative of a function is the derivative of the derivative of that function. The notation for it is $f''(x)$ or $\frac{d^2f}{dx^2}$. The second derivative tells us whether the first derivative is increasing or decreasing. If the second derivative is positive, then the first derivative is increasing. Graphically, this means that the curve of the graph is concave up. Likewise, if the second derivative is negative, the first derivative is decreasing, i.e., the curve of the graph is concave down. At a critical point where the second derivative is zero the graph may be concave up or concave down or it may change from one to the other (saddle point). To summarize:

- $\frac{d^2f}{dx^2}(p) > 0$, then at $x = p$ the function $f(x)$ is concave up
- $\frac{d^2f}{dx^2}(p) < 0$, then at $x = p$ the function $f(x)$ is concave down
- $\frac{d^2f}{dx^2}(p) = 0$, then there is a critical point of $f(x)$ at $x = p$.



## Notation for Differentiation

As differential calculus is applied in a plethora of research fields there have been proposed various notations for differentiation by various mathematicians. Here it is listed the most common notations for differentiation:

- **Leibniz's notation**: Considering the equation $y = f(x)$, where $x$ is the independent variable and $y$ being the dependent variable, the derivative of this functional relationship is written as:

  $$\frac{dy}{dx} = \frac{df(x)}{dx} = \frac{df}{dx}(x) = \frac{d}{dx}f(x)$$

  To express higher derivative exponents for $d$ and the independent variable is used, i.e., $\frac{d^ny}{dx^n}$ denotes the $n$-th derivative. If the independent variable $x$ takes a specific value $a$ we can write it as:

  $$\left.\frac{dy}{dx}\right|_{x=a} = \frac{dy}{dx}(a)$$

- **Lagrange's notation**: The probably most widely used notation to express differentiation is the notation by Lagrange using a prime mark to denote a derivative. For $f(x)$ being a function evaluated at $x$, the first derivative is written as:

$$f'(x)$$

  For Higher derivatives we write multiple prime marks, i.e., $f''(x)$ denotes the second derivative, $f'''(x)$ for the third derivative and so on. For a better reading of higher derivatives some authors use Roman numerals, such as $f^{IV}(x), f^{V}, \ldots$ to denote the $4$-th and $5$-th derivative, or we can simply use Arabic numerals, e.g., $f^{(4)}(x), f^{(5)}(x), \ldots, f^{(n)}(x)$.

- **Euler's notation**: Euler's notation uses the $D$ (D operator). For $f(x)$ being a function, the derivative is defined as:

$$Df = \frac{df}{dx}$$

  For higher derivatives it is used an exponent for $D$, i.e., $D^n f$ for the $n$-th derivative. In order to express explicitly the variable with respect to which the differentiation is done, it is used a subscript, i.e., $D_x^n f$ denotes the $n$-th derivative w.r.t variable $x$. In the scope of 2 or more variables, the symbol $\partial$ is used rather than $D$. Hence, e.g., for the first derivative of a function $f(x)$ we can write:

$$\frac{\partial f}{\partial x} = \partial_x f$$

- **Newton's notation**: In Newton's notation, it is used the dot notation for differentiation, i.e., it is place a dot over the dependent variable. For $y$ being a function of $t$ (generally Newton's method is used where $t$ denotes the variable 'time'), the derivative is defined as:

$$\dot{y}$$

  For higher derivatives, multiple dots are used, e.g., $\ddot{y}$ for the second derivative, $\dddot{y}$ for the third derivative, .... And Arabic numeral on top of a single dot indicate the order of the derivative, i.e., $\overset{n}{\dot{y}}$ denotes the $n$-th derivative.

## Critical Points and the second derivative test

If we face a critical point of the function $f(x)$, we can use the second derivative of the function to decide whether the function has a local maximum or a local minimum at the evaluated point. A critical point of a function means that the slope of the function is zero at that point. If the second derivative at point $x$ is positive, it tells us that the derivative of $f(x)$ is increasing at that point. Hence, the curve of the graph is concave up at that point. Therefore, if $x$ is a critical point of $f(x)$ and the second derivative of $f(x)$ is positive, then $x$ is a local minimum of $f(x)$. Analogously, if $x$ is a critical point of $f(x)$ and the second derivative of $f(x)$ is negative, then the slope of the graph of the function is zero at that point and the curve of the graph is concave down. Hence, the point $x$ is a local maximum of the function. If $x$ is a critical point of $f(x)$ and the second derivative of $f(x)$ is zero, then we do not learn new information about that point. The point $x$ may be a local maximum or a local minimum, and the function may also be increasing or decreasing at that point. To summarize the second derivative test :

- if $\frac{df}{dx}(p) = 0$ and $\frac{d^2f}{dx^2}(p) > 0$, then $f(x)$ has a local minimum at $x = p$

- if $\frac{df}{dx}(p) = 0$ and $\frac{d^2 f}{dx^2}(p) < 0$, then $f(x)$ has a local maximum at $x = p$
- if $\frac{df}{dx}(p) = 0$ and $\frac{d^2 f}{dx^2}(p) = 0$, then nothing can be said about the behavior of $f(x)$ at $x = p$

## Functions of 2 or more variables

Recall for functions with two variables, we use the notation $z = f(x, y)$. This notation applies also for functions with more than 2 variables. In the case of 2 variables, we call $x$ and $y$ independent variables, whereas $z$ is a dependent variable. When we want to differentiate this function, we have to decide whether we are differentiating with respect to $x$ or w.r.t. $y$. Therefore, we use the notation by using the symbol $\partial$ instead of $d$ and introduce the partial derivatives of a function:

- $\frac{\partial z}{\partial x}$ meaning "partial derivative of $z$ with respect to $x$". We differentiate w.r.t x holding $y$ constant
- $\frac{\partial z}{\partial y}$ means to differentiate with respect to $y$ holding $x$ constant

## List of Derivative Rules

- Constant Rule: $f(x) = c$, then $f'(x) = 0$
- Constant Multiple Rule: $g(x) = c \cdot f(x)$, then $g'(x) = c \cdot f'(x)$
- Power Rule: $f(x) = x^n$, then $f'(x) = n \cdot x^{n-1}$
- Sum and Difference Rule: $h(x) = f(x) \pm g(x)$, then $h'(x) = f'(x) \pm g'(x)$
- Product Rule: $h(x) = f(x)g(x)$, then $h'(x) = f'(x)g(x) + f(x)g'(x)$
- Quotient Rule: $h(x) = \frac{f(x)}{g(x)}$, then $h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$
- Chain Rule: $h(x) = f(g(x))$, then $h'(x) = f'(g(x))g'(x)$
- Trigonometry Derivatives:
  - $f(x) = sin(x)$, then $f'(x) = cos(x)$
  - $f(x) = cos(x)$, then $f'(x) = -sin(x)$
- Exponential Derivatives:
  - $f(x) = a^x$, then $f'(x) = ln(a)a^x$
  - $f(x) = e^x$, then $f'(x) = e^x$
- Logarithm Derivatives:
  - $f(x) = log_a(x)$, then $f'(x) = \frac{1}{ln(a)x}$
  - $f(x) = ln(x)$, then $f'(x) = \frac{1}{x}$