Derivative

Johan Boissard

October 16, 2010

1 What 's a derivative

We define a function $f(x) : \mathbb{R} \to \mathbb{R}$, taking x as argument; for every x there is one (and only one) value $f(x)$ associated. $f'(x)$ denotes the **derivative** of $f(x)$.

1.1 Intuitive definitions

Basically, the derivatives says **how steep is** $f(x)$ **at** x.

Or a little bit more mathematically:

$$f'(x) = \frac{\Delta f(x)}{\Delta x} \tag{1}$$

... but when Δx is very small.

1.2 Three cases

- $f'(x) > 0$ "at x "goes up"
- $f'(x) < 0$ at x "goes down"
- $f'(x) = 0$ at x "remains stationary" $\to$ local optimum

See `http://en.wikipedia.org/wiki/Derivative` for illustration.

2 Very small $\to$ Infinitesimal

$$\frac{\Delta f(x)}{\Delta x} \tag{2}$$

$$\lim_{x \to 0} \Delta x = dx \tag{3}$$

2.1 Official Definition

$$f'(x) = \frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{4}$$

2.1.1 Notations

Note that $\frac{df}{dx}$ (Leibniz's notation) is strictly equal to $f'(x)$ (Lagrange notation). However the notation $\frac{df}{dx}$ should be preferred since it really represents what differentiation is: the measure of how a function change (df) as its input changes (dx).

$\frac{d}{dx}$ is the operator for differentiation, thus one can simply write

$$\frac{d(7x^3)}{dx}$$

instead of

$$f'(x), \quad \text{where } f(x) = 7x^3$$

Note that some people (depending on the context) write it like this: $\dot{f}$ (Newton's notation) or $D_x f$ (Euler's notation)

2.2 Some examples

for $f(x) = ax$

$$\begin{aligned}
\frac{df}{dx} &= \lim_{h \to 0} \frac{a(x+h) - ax}{x+h-x} \\
&= \lim_{h \to 0} \frac{ah}{h} \\
&= \lim_{h \to 0} ah \\
&= a
\end{aligned}$$

So for a linear function, $f(x) = ax + b$ (the previous example can be easily generalized), the derivative is a **constant**. This is only valid for linear function.

for $f(x) = ax^2$

$$\begin{aligned}
\frac{df}{dx} &= \lim_{h \to 0} \frac{a(x+h)^2 - ax^2}{x+h-x} \\
&= \lim_{h \to 0} a \frac{x^2 + 2hx + h^2 - x^2}{h} \\
&= \lim_{h \to 0} a \frac{2hx + h^2}{h} \\
&= \lim_{h \to 0} a(2x + h) \\
&= 2ax
\end{aligned}$$

for $f(x) = e^x$

$$\begin{aligned}
\frac{df}{dx} &= \lim_{h \to 0} \frac{e^{x+h} - e^h}{x+h-x} \\
&= \lim_{h \to 0} \frac{e^x(e^h - 1)}{h} \\
&= e^x \underbrace{\lim_{h \to 0} \frac{e^h - 1}{h}}_{1} \\
&= e^x
\end{aligned}$$

2.3 Relation between angle and slope

One can intuitively notice that there is a tight relation between "slope" and angle, in fact this relation is very precise and is the following:

$$\theta = \arctan\left(f'(x)\right) \tag{5}$$

where θ is the angle (in radians) between the function f and the horizontal axis at point x. Once again, if $f'(x) = c$ the angle is also constant for very x and this is only the case for linear functions.

If $f'(x) = 1$, then $\theta = \frac{\pi}{4} = 45°$, this is the case when the variation of f is the same as the variation of x.

If $f'(x) = 0$, then $\theta = 0$: there is no variation in f $(df = 0)$.

If $f'(x) = \infty$, then $\theta = 0$: there is a huge variation in f, the curve is vertical at x this is a limitation of the derivative (there are methods to overcome this problem, but they are beyond the scope of that tutorial).

3 Rules for calculating derivatives

Using limits is time consuming and not very practical, fortunately there exist some properties that allows to calculate derivatives in a much simpler way.

3.1 Properties

- Linearity

$$f(x) = ag(x)$$
$$f'(x) = ag'(x)$$

$$f(x) = g(x) + h(x)$$
$$f'(x) = g'(x) + h'(x)$$

- Constants disappear

$$f(x) = 1$$
$$f'(x) = 0$$

$$f(x) = a$$
$$f'(x) = 0$$

- Polynomes

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

- Products

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

- Division

$$f(x) = \frac{g(x)}{h(x)}$$
$$f'(x) = \frac{g'(x) \cdot h(x) - g(x) \cdot h'(x)}{h^2(x)}$$

$$f(x) = \frac{1}{g(x)}$$
$$f'(x) = -\frac{g'(x)}{g^2(x)}$$

- Exponentials

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

$$f(x) = e^{h(x)}$$
$$f'(x) = h'(x)e^{h(x)}$$

- Logarithms

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

$$f(x) = \ln|g(x)|$$
$$f'(x) = \frac{g'(x)}{g(x)}$$

3.2 Examples

4 2^{nd} order derivatives

Until now, we only discussed about 1^{st} order derivatives.

2^{nd} order derivatives are defined as follows

$$f''(x) = \frac{d^2 f}{dx^2} = \frac{d}{dx}\left(\frac{df}{dx}\right) \tag{6}$$

In order to find, the 2^{nd} order derivative of a function $f(x)$, one just has to take **the derivative of the derivative**.

4.1 Examples

for $f(x) = x^2$, we have

$$\begin{aligned} f'(x) &= \frac{df}{dx} = 2x \\ f''(x) &= \frac{d}{dx}\left(\frac{df}{dx}\right) \\ f''(x) &= \frac{d}{dx}\left(f'(x)\right) \\ f''(x) &= \frac{d}{dx}\left(2x\right) \\ f''(x) &= 2 \end{aligned}$$

for $f(x) = e^{-x}$ we have

$$\begin{aligned} f''(x) &= \frac{d}{dx}\left(\frac{df}{dx}\right) \\ f''(x) &= \frac{d}{dx}\left(\frac{d(e^{-x})}{dx}\right) \\ f''(x) &= \frac{d}{dx}\left(f'(x)\right) \\ f''(x) &= \frac{d}{dx}\left(-e^{-x}\right) \\ f''(x) &= e^{x} \end{aligned}$$

5 Practical Uses of the Derivatives

5.1 Optimum

In order to find an optimum of f, the optimum x^* has to satisfy the following relation

$$f(x^*) = 0 \tag{7}$$

There are 3 kinds of optimum

- local maximum
- local minimum
- saddle point

To differentiate those different category (it can be a good thing to be able to tell the difference between the x^* that <u>maximizes</u> the cost and the x^* that <u>minimizes</u> the cost...), we have the following relation.

x^* must satisfy $f'(x^*) = 0$ <u>and</u>

- local maximum $\Leftarrow f''(x) < 0$
- local minimum $\Leftarrow f''(x) > 0$
- saddle point $\Leftarrow f''(x) = 0$

6 Partial Derivative

Suppose a function of more than one variable, $f(x_1, x_2, \ldots, x_n)$.

The partial derivative of f with respect to x_k is

$$\frac{\partial f}{\partial x_k}. \tag{8}$$

6.1 Example

If we set

$$f(x, y) = x^2 + xy + y^2$$

we have the following partial derivatives

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x + y \\ \frac{\partial f}{\partial y} &= x + 2y \end{aligned}$$

7 Total derivative

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} \tag{9}$$

or

$$df = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \tag{10}$$

7.1 Differential operator

$$\frac{d}{dx} = \frac{\partial f}{\partial x} + \sum_{i=1}^{k} \frac{dy_i}{dx}\frac{\partial}{\partial y_i} \tag{11}$$