Derivatives

Contents

1 Overview

The *derivative* of a function measures the rate of variation of the function with respect to one of the function's independent variables, i.e., how much does the value of the function change given a small change in the value of the function's independent variables. Here, the small change is taken when its limit tends to zero.

While the concept of derivative as a measure of rate of variation remains the same for all functions, the form of the derivative depends on the type of the function in which we are calculating. The function types to consider are the following:

1. **A scalar function of a single scalar variable**, i.e., $f(x) \in \mathbb{R}$ and $x \in \mathbb{R}$.

2. **A scalar function of multiple scalar variables (or a scalar function of a vector variable)**. In this case, $f(\mathbf{x}) \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$.

3. **A vector function of a single scalar variable**, i.e., $\mathbf{f}(x) \in \mathbb{R}^n$ and $x \in \mathbb{R}$.

4. **A vector function of a vector variable**, i.e., $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$.

2 Derivative of a scalar function of a single scalar variable

Let f be a scalar function of a single variable x. The derivative of the function w.r.t. x is:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \tag{1}$$

It describes the slope (i.e., rate of change) of the function at a point x. For example, the derivative of the function:

$$f(x) = x \sin x^2 + 1 \tag{2}$$

is

$$\frac{df}{dx}(x) = 2x^2 \cos x^2 + \sin x^2. \tag{3}$$

The values of the derivative at $x = -1$, $x = -0.5$, $x = 1$, $x = 0.5$, and $x = 1$ are shown in Figure 1. Each derivative is represented by a tangent line indicating the rate of change or slope of the curve at each point.

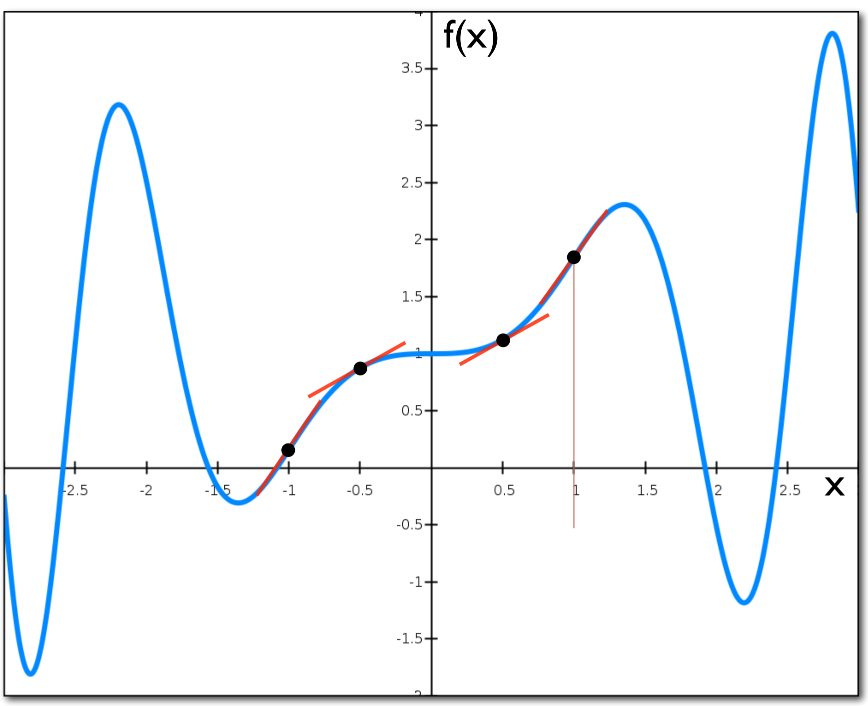

Figure 1: The derivative of the function $f(x) = x \sin x^2 + 1$ at $x = -1$, -0.5, 1, 0.5, and 1.

The derivative is positive when the value of the function is increasing and negative otherwise.

3 Derivative of a scalar function of multiple scalar variables (i.e., vector variable)

Let f be a scalar function of a vector variable. The vector variable is $\mathbf{x} = (x_1, x_2, \ldots, x_N)^\mathsf{T}$. This type of function is also called a scalar function of multiple variables or multi-variate function. The value of the function at a point $\mathbf{x}$ is given by $f(\mathbf{x})$ or $f(x_1, x_2, \ldots, x_N)$, and its derivative w.r.t. x is:

$$\frac{\mathrm{d}f}{\mathrm{d}\mathbf{x}} = \frac{\partial f}{\partial (x_1, x_2, \ldots, x_N)} = \left[\frac{\partial f}{\partial x_1}\ \frac{\partial f}{\partial x_2}\ \cdots\ \frac{\partial f}{\partial x_N}\right]^\mathsf{T} = \boldsymbol{\nabla} f, \tag{4}$$

which is called the *Gradient* of f at x, and denoted by $\boldsymbol{\nabla} f$. Note that the components of $\boldsymbol{\nabla} f$ are simply the derivatives of function f w.r.t. each independent variable x_i (in isolation) calculated as described by Equation 1. The gradient is a vector field because it yields a vector at each location $\mathbf{x} = (x_1, x_2, \ldots, x_N)^\mathsf{T}$ of the domain of the function. In the gradient, vectors point towards the growth of the function at a point (i.e., vectors point uphill). An example of a gradient vector field of a function $f(x_1, x_2)$ is shown in Figure 2.

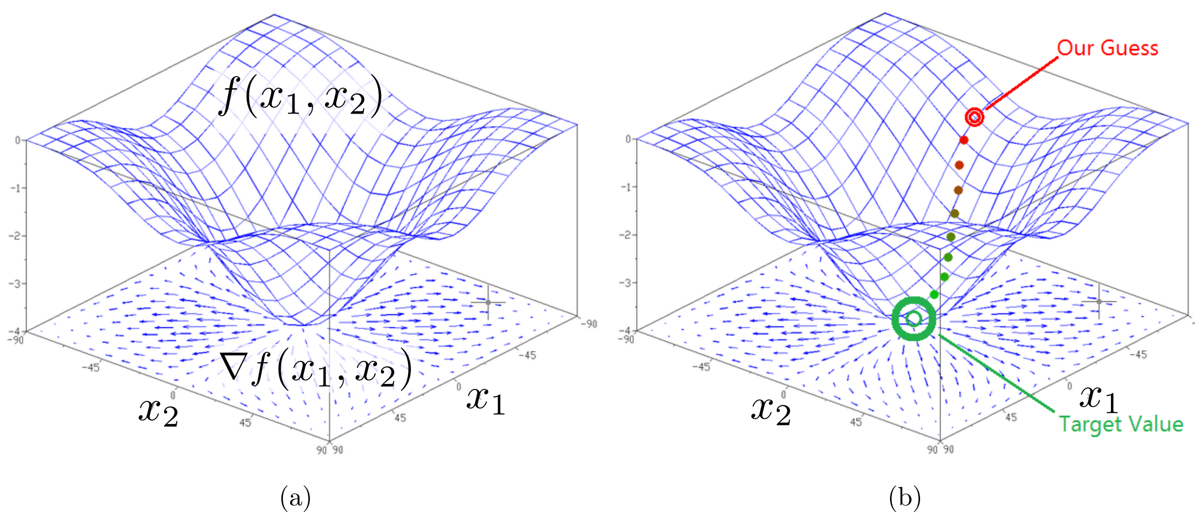

Figure 2: (a) Gradient vector field of two-variate function. (b) Gradient-descent method. In the gradient-descent method, we want to take the direction inverse to the gradient as the gradient (slope) points “uphill”. Plots adapted from `https://goo.gl/zejgBw`.

4 Derivative of a vector function of a single scalar variable

Let $\mathbf{r}$ be a vector function representing the 3-D motion of a particle as a function of time t:

$$\mathbf{r} = [r_x \; r_y \; r_z]^\mathsf{T} . \tag{5}$$

Its derivative is given by:

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \left[\frac{\mathrm{d}r_x}{\mathrm{d}t} \; \frac{\mathrm{d}r_y}{\mathrm{d}t} \; \frac{\mathrm{d}r_z}{\mathrm{d}t}\right]^\mathsf{T} , \tag{6}$$

which is also a vector field representing the velocity of the particle at time t. This vector field runs along the curve that describes the motion. Similar to the gradient, each component of the vector field in Equation 6 are calculated by Equation 1.

5 Derivative of a vector function of a vector variable

Some applications require us to calculate derivatives of vector quantifies with respect to other vector quantities. For example, if $\mathbf{f}$ is a vector-valued function of a vector of variables, $\mathbf{x}$. Here, $\mathbf{f}(\mathbf{x}) = (f_1, f_2, \ldots, f_M)^\mathsf{T}$ and $\mathbf{x} = (x_1, x_2, \ldots, x_N)^\mathsf{T}$.

The derivative of $\mathbf{f}$ w.r.t. $\mathbf{x}$ is:

$$\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}\mathbf{x}} = J(\mathbf{f}, \mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_N} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial x_1} & \frac{\partial f_M}{\partial x_2} & \cdots & \frac{\partial f_M}{\partial x_N} , \end{bmatrix} \tag{7}$$

and is called the *Jacobian*. By inspecting the Jacobian matrix, we notice that it can be written as a stack of gradients $\boldsymbol{\nabla} f_i$ as rows, i.e.:

$$J(\mathbf{f}, \mathbf{x}) = \begin{bmatrix} \boldsymbol{\nabla} f_1 \\ \boldsymbol{\nabla} f_2 \\ \vdots \\ \boldsymbol{\nabla} f_M \end{bmatrix} \tag{8}$$

or as matrix of columns where each column is derivative of the vector function with respect to a component of the vector variable, i.e.:

$$J(\mathbf{f}, \mathbf{x}) = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \frac{\partial \mathbf{f}}{\partial x_2} & \cdots & \frac{\partial \mathbf{f}}{\partial x_N} \end{bmatrix} \tag{9}$$

The Jacobian gives us all we need to know about the variations of function with respect to all variables. It is hard to visualize the Jacobian. However, we can think of it as a volume where each plane is a gradient vector field. Alternatively, maybe we can think of the Jacobian as a stack of tangent vectors of several particle motions (I am not so sure about this interpretation!).

References