Differentiation of Functions
(C001)

Paul Valentine

17 March 2020

Abstract

This paper develops the concept of differentiation of functions as expressing the rate of change of a function at a point. Firstly the rate of change of single variable functions is considered leading to differentiation of multi variable functions and ending in the concept of a direction derivative. In considering directional derivatives the concept of a vector at a point naturally occurs.

1 Single Variable

Let $x \in \mathbb{R}$ and $f(x) \in \mathbb{R} \mid f : x \to f(x)$.

Consider the linear approximation of the function $f(x)$ at a point displaced from x by δ. Define the tangent to a point on $f(x)$ as being the line that locally best approximates the rate of change of the function not bisecting $f(x)$ evaluated at that point. Define this rate of change to be $f'(x)$

$$f(x+\delta) = f(x) + f'(x)\delta \ \therefore \ f'(x) \approx \frac{f(x+\delta) - f(x)}{\delta} \tag{1}$$

This approximation gets more and more accurate as $\delta \to 0$. The rate of change of the function $f(x)$ (w.r.t x) is therefore.

$$f'(x) = \lim_{\delta \to 0} \frac{f(x+\delta) - f(x)}{\delta} \tag{2}$$

The process of differentiation can be viewed as a function g acting on the function f such that $g : f \to f'$. When expressed in this form the differentiation becomes an operator with the notation $\frac{\mathrm{d}}{\mathrm{d}x}$.

$$\frac{\mathrm{d}}{\mathrm{d}x} : f \to f' \tag{3}$$

2 Multi-variable

A function $f(x,y) \mid x, y \in \mathbb{R} \mid f(x,y) : \mathbb{R}^2 \to \mathbb{R}$ can be viewed as a family of functions $f_x(y) : \mathbb{R} \to \mathbb{R}$. For example let $f(x,y) = x + xy + y$. Then $f_a(y) = a + ay + y$ and $f'_a(y) = a + 1$. The partial derivative is the whole set of such functions $f'_x(y) = x + 1$. This is, in effect the derivative of the function holding x constant. This is given the following general notation.

$$\frac{\partial}{\partial y} f(x,y) \tag{4}$$

This is a partial derivative. Given in this form, the partial derivative is a function that acts on a function to produce a function. The above is effectively the rate of change of the function w.r.t to y only.

Consider $f(x,y) = \cos(x) + \sin(y)$ as shown in figure 1. It is clear from this that the function at any point on its surface has distinct rates of change (gradients) in each direction. Therefore to pick out a single rate of change we need to impose a direction on the surface. If we go in any direction we can calculate the change in the function relative to its value at point p from equation 1.

$$\delta f = f_y(x + \delta x) - f_y(x) + f_x(y + \delta y) - f_x(y) = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y \tag{5}$$

In which δx and δy are components of a displacement vector from point p. We parametrise the travel along the vector with a parameter $\lambda \in \mathbb{R}$ such that we can ask what is the rate of change of f w.r.t λ as $\lambda \to 0$.

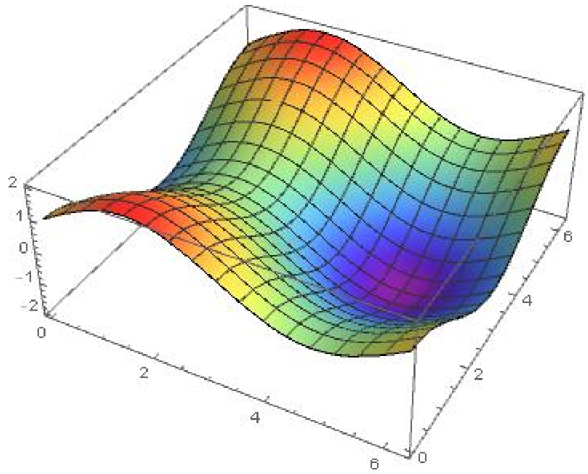

Figure 1: $\cos(x) + \sin(y)$

This becomes the rate of change at p in the direction of the displacement vector $(\delta z, \delta y)$. Therefore equation 5 becomes.

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = \frac{\mathrm{d}x}{\mathrm{d}\lambda}\frac{\partial f}{\partial x} + \frac{\mathrm{d}y}{\mathrm{d}\lambda}\frac{\partial f}{\partial y} \tag{6}$$

Equation 6 is illuminating if once more we consider differentiation as a function acting on functions.

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}f = \frac{\mathrm{d}x}{\mathrm{d}\lambda}\frac{\partial}{\partial x}f + \frac{\mathrm{d}y}{\mathrm{d}\lambda}\frac{\partial}{\partial y}f \tag{7}$$

Equation 7 exposes how a vector (in this specific case a displacement vector) $\frac{\mathrm{d}x}{\mathrm{d}\lambda}\frac{\partial}{\partial x} + \frac{\mathrm{d}y}{\mathrm{d}\lambda}\frac{\partial}{\partial y}$ acts on a function f to produce $\frac{\mathrm{d}f}{\mathrm{d}\lambda}$. Seen through this lens it is clear that $\frac{\mathrm{d}}{\mathrm{d}\lambda}$ is a vector v in the most abstract of senses arising out of consideration of analytical differentiation. A remarkable result! Here, equation 7 is a **Directional Derivative**.

$$\left.\frac{\mathrm{d}}{\mathrm{d}\lambda}\right|_p : \mathbb{R}^n \to \mathbb{R} \tag{8}$$

Thus we can view a vector at a point as forming a linear map of a function to $\mathbb{R}$

$$v_p : f \to \mathbb{R} \tag{9}$$

The family of linear maps (vectors) at a point for a tangent space T_p at that point of the same dimension as the function.