Partial derivatives and the chain rule

Partial derivative is extensively used in physics and mathematics. It is not a hard concept, if you know what derivative is.

Recall what ordinary derivative is. If $f(x, y) = x^2y + y^2$, we have:

$$\frac{df(x,y)}{dx} = 2xy + (x^2 + 2y)\frac{dy}{dx} \tag{1}$$

On the other hand, partial derivative of $f(x, y)$ with respect to x is given by:

$$\frac{\partial f(x,y)}{\partial x} = 2xy \tag{2}$$

In other words, we have set y as a constant that does not depend on x. (i.e. $\frac{dy}{dx} = 0$) When taking partial derivative with respect to x, we treat all other variables as constants. Now, let's take (2) once more, this time with respect to y. We get:

$$\frac{\partial^2 f(x,y)}{\partial y \partial x} = \frac{\partial(2xy)}{\partial y} = 2x \tag{3}$$

To step further, let's calculate the following quantity:

$$\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial(x^2 + 2y)}{\partial x} = 2x \tag{4}$$

So, we see that

$$\frac{\partial^2 f(x,y)}{\partial y \partial x} = \frac{\partial^2 f(x,y)}{\partial x \partial y} \tag{5}$$

In fact, this is true for any function $f(x, y)$ as long as the partial derivatives exist. In other words, we say partial derivatives "commute." One can actually check the above formula from the definition of partial derivatives as follows:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{\Delta x_0 \to 0} \frac{f(x_0 + \Delta x_0, y_0) - f(x_0, y_0)}{\Delta x_0} \tag{6}$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{\Delta y_0 \to 0} \frac{f(x_0, y_0 + \Delta y_0) - f(x_0, y_0)}{\Delta y_0} \tag{7}$$

$$\frac{\partial^2 f}{\partial y \partial x} = \lim_{\Delta y \to 0} \lim_{\Delta x \to 0} \frac{(f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)) - (f(x_0 + \Delta x, y_0) - f(x_0, y_0))}{\Delta y \Delta x} \tag{8}$$

On the other hand, we have:

$$\frac{\partial^2 f}{\partial x \partial y} = \lim_{\Delta x \to 0} \lim_{\Delta y \to 0} \frac{(f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)) - (f(x_0, y_0 + \Delta y) - f(x_0, y_0))}{\Delta y \Delta x} \tag{9}$$

Therefore, we see that they are indeed equal.

We can also Taylor-expand an arbitrary function $f(x, y)$ in terms of partial derivatives. To this end, let's first regard y as a constant. Then, we have

$$f(x, y) = f(x_0, y) + \left[\frac{\partial f}{\partial x}(x_0, y)\right](x - x_0) + \cdots \tag{10}$$

We also have

$$f(x_0, y) = f(x_0, y_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0) + \cdots \tag{11}$$

$$\left[\frac{\partial f}{\partial x}(x_0, y)\right] = \left[\frac{\partial f}{\partial x}(x_0, y_0)\right] + \left[\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)\right](y - y_0) + \cdots \tag{12}$$

Summarizing, we have

$$f(x, y) = f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0) + \cdots \tag{13}$$

Not sure, what this equation means? First, let's re-write the above formula. Upon substituting $x_0 \to x$, $y_0 \to y$, $(x - x_0) \to \Delta x$, $(y - y_0) \to \Delta y$, we have

$$f(x + \Delta x, y + \Delta y) - f(x, y) = \left[\frac{\partial f}{\partial x}(x, y)\right]\Delta x + \left[\frac{\partial f}{\partial y}(x, y)\right]\Delta y + \cdots \tag{14}$$

The left-hand side is the change in f. When $\cdots$ in the right-handside can be ignored (i.e., considering only the linear order in Δx and Δy), the above formula says that the change in f is the sum of the change in f due to x and the change in f due to y.

Actually, we can visualize what we just said. Before doing so, first recall how the concept of derivative arised. We can approximate any graph $y = g(x)$ as a straight line near a given point as

$$g(x + \Delta x) = g(x) + \frac{dg}{dx}\Delta x + \cdots \tag{15}$$

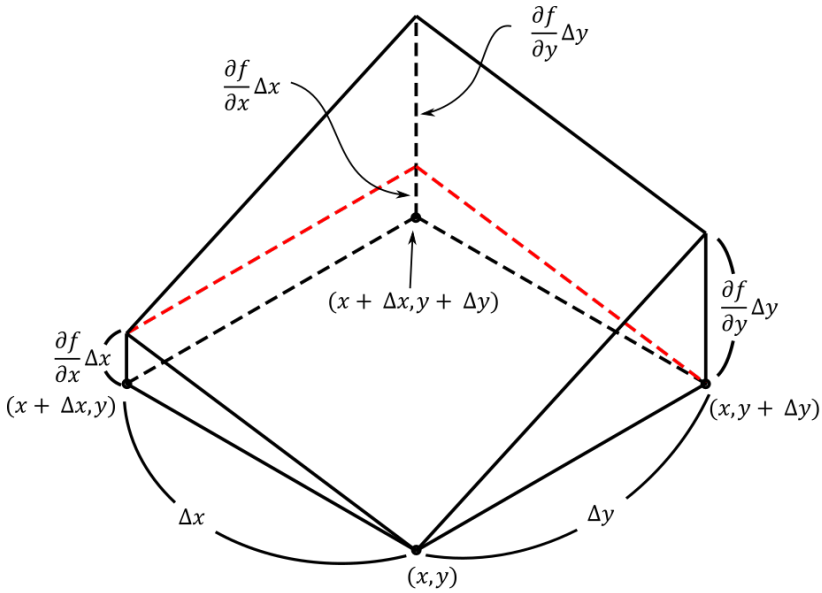

Figure 1: The change of f is approximately, the sum of its change due to the change of x and its change due to the change of y.

when $\cdots$ is negligible i.e., when Δx is small. Here, dg/dx is the slope.

Similarly, we can approximate any graph $z = f(x, y)$ as a flat plane near a given point. See Fig. 1. You see a graph of $z = f(x, y)$, which can be regarded as flat, when Δx and Δy are sufficiently small. In the figure, you see that $f(x + \Delta x, y)$ is greater than $f(x, y)$ by $\dfrac{\partial f}{\partial x}\Delta x$ and $f(x, y + \Delta y)$ is greater than $f(x, y)$ by $\dfrac{\partial f}{\partial y}\Delta y$. Through the aid of two red dotted lines, we see that $f(x + \Delta x, y + \Delta y)$ is greater than $f(x, y)$ by the sum of these two terms.

Now, let's see a further application of the partial derivatives. You are already familiar with the chain rule which is given by following formula:

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx} \tag{16}$$

If f is a function of several variables x, y, z, which are in turn functions of t, we have:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial x}\frac{dz}{dt} \tag{17}$$

You can prove this by using (14). The only difference is that we have three variables (i.e. x,y,z) in our case. We can also have the partial derivative

version of this formula as follows, when x,y,z are functions of t and s:

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial x}\frac{\partial z}{\partial t} \tag{18}$$

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial x}\frac{\partial z}{\partial s} \tag{19}$$

Problem 1. Using (14), convince yourself that (17) is correct.

Problem 2. Let $g(x, y) = \sin(x^2 y)$. By explicit calculations, check that the partial derivatives commute.

Summary

- A partial derivative with respect to x is a derivative assuming all the other variables besides x is a constant. It is denoted by $\dfrac{\partial}{\partial x}$.
- Partial derivatives commute:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

- If f is a function x, y, z, which are in turn functions of t, we have:

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

- If f is a function of variables x, y, z which are in turn functions of t, s, the chain rule says

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial x}\frac{\partial z}{\partial t}$$

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial x}\frac{\partial z}{\partial s}$$