Partial Derivatives

1 Limits

With single variable functions, we had a few methods of evaluating limits. For the limit to exist, we required that:

$$\lim_{x \to a^+} f(x) = \lim_{x \to a^-} f(x) = L$$

With single variable functions, there are two directions to approach a limit. You can approach is from the left and from the right. However, with multivariable functions, you can also approach it from the vertical directions as well. With this in mind, we can create the definition for a limit of a multivariable function:

$$\lim_{(x,y) \to (a,b)} f(x,y) = f(a,b)$$

We simply require that the function is continuous at this point. The same things that created discontinuity with single variable functions create discontinuity in multivariable functions: dividing by zero, square roots of negatives, logarithms of numbers ≤ 0.

2 Partial Derivatives

To understand partial derivatives, we must first understand how derivatives actually work. If you take a derivative of a function, say $f(x)$, then you get the resulting change in the value of $f(x)$ as x changes. The idea with partial derivatives, is that we are isolating each variable such that only one of the variables is changing while the other is constant. If you have more than one variable, and they are both changing at the same time, there are technically an infinite amount of ways that they could be changing.

Let's say that we have a multivariable function $f(x,y)$. To find a partial derivative, we must hold one of these variables constant. For instance, to find the partial derivative with respect to x, we must hold y constant. We would write this $f_x(x,y)$. Using the same logic, if we want to find the partial derivative with respect to y of $f(x,y)$, we could write $f_y(x,y)$. Once you determine which variable you are holding constant, you can then treat that variable like a constant and use the same rules we used for single variable calculus. For example, if we have the function:

$$f(x,y) = 3x^2y^3$$

and we want to find $f_x(x,y)$ (the partial derivative of the function with respect to x), we could do:

$$f_x(x,y) = 6xy^3$$

If we wanted to find $f_y(x,y)$, we would have:

$$f_y(x,y) = 9x^2y^2$$

These are very simple cases, of course, and are merely the **first order partial derivatives**. We can have higher order partial derivatives in the same way that we can have higher order single variable derivatives.

Definition of a Partial Derivative If the concept of partial derivatives doesn't make sense yet, the formal definition of a partial derivative may clarify any confusion:

$$f_x(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}$$

$$f_y(x, y) = \lim_{h \to 0} \frac{f(x, y + h) - f(x, y)}{h}$$

The variable that we are partially differentiating with respect to is the variable that has the infinitely small change of h added to it.

Alternative Notation There are also alternative notations to partial derivatives, as seen below:

$$f_x(x, y) = f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(f(x, y))$$

$$f_y(x, y) = f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(f(x, y))$$

Traces A partial derivative is a slope of a **trace**. A **trace** is a two-dimensional part of a three-dimensional graph when you hold one variable constant. For instance, if you are given a function, $f(x, y)$, and asked to find the slopes of the traces at point (x, y), you must use two separate partial derivatives - one for each variable.

3 Higher Order Partial Derivatives

Like single variable functions, we can take the higher order partial derivatives of a multivariable function. However, this time, we have more variables to choose from. If our first order partial derivative is one variable, say x, then we take the derivative again but with respect to x or y. The same applies if our first order partial derivative is with respect to y. This yields four possible second order partial derivatives, represented by the notation below:

$$(f_x)_x = f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$$

$$(f_x)_y = f_{xy} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$$

$$(f_y)_x = f_{yx} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$$

$$(f_y)_y = f_{yy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$$

In the second and third example, we often call these **mixed partial derivatives** since we are taking derivatives with respect to multiple variables. You must always differentiate with respect to the variables left to right. The second example will sometimes give you a different result than the third will despite their similar notation.

Higher-Higher Order Partial Derivatives After taking the second order partial derivative, you can continue doing this with respect to the same variables. It works the exact same way, but becomes increasingly more tedious and complicated the higher you go.

4 Chain Rule

The notation for chain rule with single variable functions is pretty standard, and generally not that complicated:

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

Despite the confusing notation, the chain rule with multivariable functions is, fundamentally, similar to single variable functions. We will look at a few specific cases for the chain rule.

Case 1 In this case, we will have three functions: $z = f(x, y)$, $x = g(t)$, and $y = h(t)$. In this case, the chain rule is as follows:

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

Case 2 In this case, both x and y are multivariable functions as well: $z = f(x, y)$, $x = g(s, t)$, $y = h(s, t)$. The chain rule is as follows:

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

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

General Rule for Chain Rule Suppose that z is a function of n variables $x_1, x_2, ..., x_n$ and that each of these variables are functions of m variables $t_1, t_2, ...t_m$. Then for any variable, t_i, $i = 1, 2, ..., m$ we have the following:

$$\frac{\partial z}{\partial t_i} = \frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial z}{\partial x_2}\frac{\partial x_2}{\partial t_i} + ... + \frac{\partial z}{\partial x_n}\frac{\partial x_n}{\partial t_i}$$

Implicit Differentiation Let's start with a function in the form $F(x, y) = 0$. If a function isn't in this form, we can get it there by moving everything to one side of the equals sign. In single variable calculus, implicit differentiation could be tedious. Let's try to simplify it using partial derivatives. Let's start by differentiation with respect to x:

$$F_x + F_y\frac{dy}{dx} = 0 \implies \frac{dy}{dx} = -\frac{F_x}{F_y}$$

This makes implicit differentiation much easier in single variable calculus! We can do the same thing with functions of three variables, $F(x, y, z) = 0$ and $z = f(x, y)$. We will start by trying to find $\frac{\partial z}{\partial x}$:

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0$$

We will simplify this by treating y as constant ($\frac{dy}{dx} = 0$) and simplifying fractions. Then, solving this equation, we get:

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$

You can do the same thing while treating x as a constant to get:

$$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$$

5 Directional Derivatives

The discussion of **directional derivatives** allows us to figure out what happens when you change both x and y simultaneously. The main issue with this, however, is that x and y can be changing in more than one way. They could be changing at different rates and in different directions as well. We can solve this with the following definition.

Directional Derivative Definition The rate of change of $f(x, y)$ in the direction of the unit vector $\vec{u} = \langle a, b\rangle$ is called the **directional derivative** and is denoted by $D_u f(x, y)$. The formal definition is as follows:

$$D_{\vec{u}} f(x, y) = \lim_{h\to 0} \frac{f(x + ah, y + bh) - f(x, y)}{h}$$

Derivation of a Simpler Formula While this is the formal definition, this is quite challenging to work out in most cases. Let's derive a more efficient formula. Let's start with a single variable function:

$$g(z) = f(x_0 + az, y_0 + bz)$$

where x_0,y_0,a, and b are some fixed numbers. With all of these variables fixed, the only variable that is still changing is z. Therefore, by the single variable definition of a derivative, we have:

$$g'(z) = \lim_{h\to 0} \frac{g(z + h) - g(z)}{h}$$

$$g'(0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h} = \lim_{h\to 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} = D_{\vec{u}} f(x_0, y_0)$$

$$g'(0) = D_{\vec{u}} f(x_0, y_0)$$

With this in the back of our minds, let's move onto another idea. With the chain rule, we have:

$$g'(z) = \frac{dg}{dz} = \frac{\partial f}{\partial x}\frac{dx}{dz} + \frac{\partial f}{\partial y}\frac{dy}{dz} = f_x(x, y)a + f_y(x, y)b$$

$$g'(z) = f_x(x, y)a + f_y(x, y)b$$

$$g'(0) = f_x(x_0, y_0) + f_y(x_0, y_0)b$$

Now, all we have to do is equate what we had before to this new expression:

$$g'(0) = D_{\vec{u}} f(x_0, y_0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b$$

$$D_{\vec{u}} f(x, y) = f_x(x, y)a + f_y(x, y)b$$

Using a similar method, you can also derive a similar formula for vectors in $\mathbb{R}^3$:

$$D_{\vec{u}} f(x, y, z) = f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c$$

6 Gradient

The gradient is a more simplistic looking version of the directional derivative. Using the previous definition of the directional derivative, it should be clear that:

$$D_{\vec{u}}f(x, y, z) = \langle f_x, f_y, f_z \rangle \cdot \langle a, b, c \rangle$$

The second vector is the unit vector $\vec{u}$ that gives the direction of change. The first vector is called the **gradient**. It is written:

$$\nabla f = \langle f_x, f_y, f_z \rangle$$

With this new definition, we can rewrite the directional derivative:

$$D_{\vec{u}}f = \nabla f \cdot \vec{u}$$

Maximum Rate of Change To get the maximum rate of change of a multivariable function, we can use the gradient and directional derivative. The maximum rate of change of a function is given by:

$$\|\nabla f(\vec{x})\|$$

and will occur in the direction given by $\nabla f(\vec{x})$

Quick Fact About Gradients The gradient vector $\nabla f(x_0, y_0)$ is orthogonal the the **level curve** $f(x, y) = k$ at (x_0, y_0). Similarly, the gradient vector $\nabla f(x_0, y_0, z_0)$ is orthogonal to the leve surface $f(x, y, z) = k$ at the point (x_0, y_0, z_0)