Functions of Several Variables

Lang has a very specific definition of function. He requires that the output of f is a number. The input can be any subset of n-space.

Example. $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = \sqrt{x^2 + y^2}$. We can interpret f as a function that tells us our distance to the origin when we're standing at a point (x, y).

Example. $f : \mathbb{R}^3 \to \mathbb{R}$ defined by $f(x, y, z) = x^2 - \sin(xyz) + yz^3$.

The graph of a function on defined on $S \subset \mathbb{R}^2$ would have the form

$$\{(x, y, f(x, y)) : (x, y) \in S\}.$$

In this case, the graph sits in $\mathbb{R}^3$.

For a fixed number c, the equation $f(x, y) = c$ describes a curve in $\mathbb{R}^2$. Such a curve is called a **level curve**.

Question. What do the level curves of $f(x, y) = x^2 + y^2$ look like? What about $f(x, y) = \sqrt{x^2 + y^2}$.

If $f(x, y, z)$ is a function of three variables, the equation $f(x, y, z) = c$ describes a surface, called a **level surface**.

Question. What do the level surfaces of $f(x, y, z) = x^2 + y^2 + z^2$ look like? What about $f(x, y, z) = 3x^2 + 2y^2 + z$?

Partial Derivatives

First consider a function of two variables $f(x, y)$. If we hold one of the variables fixed and allow the other to vary, we obtain a function of one variables, and we can take the derivative as we did in Calc I:

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

This is the **partial derivative with respect to the first variable** or the **partial derivative with respect to x**. The second partial derivative would be

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

Notations for this include $D_1 f, D_2 f$; $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$; f_x, f_y. And of course, we can extend these ideas to functions of 3 or more variables.

Example. Let $f(x, y) = x^2 y^3$. To compute $\partial f / \partial x$, we treat y as a constant and differentiate as usual:

$$\frac{\partial f}{\partial x} = 2xy^3.$$

Similarly,

$$\frac{\partial f}{\partial y} = 3x^2 y^2.$$

Geometrically, for functions of two variables, taking a partial derivative corresponds to slicing the graph at $x = a$ or $y = a$ for a constant a and then looking at the slope of the tangent.

Note that $D_i f$ is itself a function that we can evaluate at points.

Example. Let $f(x, y) = \sin(xy)$. Compute $D_2 f(1, \pi)$.

$$D_2 f(x, y) = \cos(xy)x.$$

So then

$$D_2 f(1, \pi) = \cos(\pi) \cdot 1 = -1.$$

Notice that we can use vector notation and write the partial derivative with respect to x_i as

$$(D_i f)(X) = \lim_{h \to 0} \frac{f(X + hE_i) - f(X)}{h}.$$

The **gradient** of a function is the vector-valued function

$$\operatorname{grad} f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right).$$

One can easily generalize this definition to higher dimensions.

Example. Let $f(x, y, z) = x^2 y \sin(yz)$. Find $\operatorname{grad} f(1, 1, \pi)$.

We note that for functions f, g and any constant c

$$\operatorname{grad}(f + g) = \operatorname{grad} f + \operatorname{grad} g, \ \operatorname{grad}(cf) = c \operatorname{grad} f.$$

Continuity

(For this section, I'll be borrowing from both Lang and Courant)

Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be points. The distance between them is

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

In the case where $n = 1$ (i.e. when x and y are just numbers), we see the distance is the absolute value of the difference $x - y$ (recall that $\sqrt{x^2} = |x|$).

If a and b are *nonnegative* and $a \leq b$, then $\sqrt{a} \leq \sqrt{b}$. Conversely, if $\sqrt{a} \leq \sqrt{b}$, then $a \leq b$. In particular, we have $|x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2}$. In the same way, $|y| \leq \sqrt{x^2 + y^2}$, and of course, this works for more than just two variables.

Recall the idea of a function being continuous at $x_0 \in X$. Intuitively, this means that as x gets closer to x_0, $f(x)$ gets closer $f(x_0)$.

Definition. A function $f(x, y)$ defined on a region R is **continuous** at a point $(x_0, y_0) \in R$ if, for every positive number ϵ, one can find a corresponding positive number δ_ϵ such that $d((x_0, y_0), (x, y)) < \delta_\epsilon$ (where $(x, y) \in R$) guarantees $d(f(x, y), f(x_0, y_0)) < \epsilon$.

Example. Let's look at a single variable example first. Let $f(x) = 10x$. Let us show that f is continuous at 0. So let $x_0 = 0$. Then $f(x_0) = f(0) = 0$. Now, somebody hands us some $\epsilon > 0$, and we need to respond by finding a $\delta > 0$ such that x being within δ of x_0 guarantees that $f(x)$ is within ϵ of $f(x_0)$. Now,

$$|f(x) - f(x_0)| = |10x| = 10|x|,$$

while

$$|x - x_0| = |x|.$$

Let $\delta = \epsilon/10$. Then if $d(x_0, x) = |x - x_0| < \delta$, we have

$$|f(x) - f(x_0)| = 10|x| = 10|x - x_0| < 10(\epsilon/10) = \epsilon.$$

This shows that f is continuous at $x_0 = 0$. In this case, the same argument works at any x_0, not just 0.

A function of several variables can have discontinuties of a more complicated type.

Example. Let $f(x, y) = \frac{2xy}{x^2+y^2}$, $f(0, 0) = 0$. If we approach the origin along the x-axis (i.e. along $y = 0$), we have $f(x, 0) = 0$. Similarly, $f(0, y) = 0$. On the other hand, if we approach the origin along the line $y = x$, we have $f(x, x) = 2x^2/(2x^2) = 1$. Thus f cannot be continuous at $(0, 0)$.

Definition. We say that the **limit** of $f(x, y)$ as (x, y) approaches (x_0, y_0) equals L if, for every positive number ϵ, one can find a corresponding positive number δ_ϵ such that $0 < d((x_0, y_0), (x, y)) < \delta_\epsilon$ guarantees $d(f(x, y), L) < \epsilon$. One writes

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L.$$

Differentiability

In initial attempt to define differentiability of functions of several variables, one might be tempted to write down something like

$$\frac{f(X + H) - f(X)}{H}.$$

This doesn't make sense, though, because we have no notion of dividing by vectors.

Let us look back to the single variable case first. The derivative in this case was defined to be

$$f'(x) = \lim_{h\to 0} \frac{f(x + h) - f(x)}{h}.$$

Let

$$\varphi(h) = \frac{f(x + h) - f(x)}{h} - f'(x).$$

$\varphi(h)$ is not defined for $h = 0$, but

$$\lim_{h\to 0} \varphi(h) = 0.$$

We can go ahead and just set $\varphi(0) = 0$. We can also write

$$f(x + h) - f(x) = f'(x)h + h\varphi(h),$$

and this is true for $h = 0$ now as well. We will now do a silly thing: let $g(h) = \varphi(h)$ for $h > 0$ and $g(h) = -\varphi(h)$ for $h < 0$. This is simply so that we now have

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

where $\lim_{h\to 0} g(h) = 0$.

Conversely, suppose that there exists some number a and a function $g(h)$ with $\lim_{h\to 0} g(h) =$

0 such that

$$f(x+h) - f(x) = ah + |h|g(h).$$

Then when $h = 0$, we have

$$\frac{f(x+h) - f(x)}{h} = a + \frac{h}{|h|}g(h).$$

As $h \to 0$, the rightmost term goes to 0 and we see that f is differentiable. In fact, $f'(x) = a$.

This dicussion is all to say that we could very well take this to be the definition of differentiability. In words, differentiability at a point x_0 is equivalent to a function having a "good" linear approximation. Here, "good" means that the difference between the approximation and f shrinks faster than $|h|$. This is maybe more easily seen if we write

$$f(x+h) - (f(x) + ah) = |h|g(h)$$

and then dividing by $|h|$.

It is this definition that will be easier to adapt to the higher dimension situation. We will start with two variables. Let $X = (x, y)$, $H = (h, k)$, so that $X + H = (x + h, y + k)$ and

$$f(X + H) - f(X) = f(x + h, y + k) - f(x, y).$$

Definition. f is differentiable at X if the partial derivatives exist and if there exists a function g (defined for small H) with $\lim_{H \to O} g(H) = 0$ such that

$$f(X + H) - f(X) = \text{grad}(f) \cdot H + \|H\|g(H).$$

Theorem. Let f be defined on an open set U. Assume that the partial derivatives exist and are continuous at each point of U. Then f is differentiable.

Most functions we encounter will be differentiable.

Repeated Partial Derivatives

Differentiation gives us a new function, and of course, we can differentiate again.

Example. Let $f(x, y) = \cos(xy)$. Then $D_1 f = -y\sin(xy)$, $D_2 f = -x\sin(xy)$. Differentiating again, we have

$$D_2 D_1 f = D_2(-y\sin(xy)) = -\sin(xy) + xy\cos(xy).$$

We also have

$$D_1 D_2 f = D_1(-x\sin(xy)) = -\sin(xy) + xy\cos(xy).$$

We see they agree!

Theorem. If $D_1 f$, $D_2 f$, $D_1 D_2 f$, and $D_2 D_1 f$ all exist and are continuous, then

$$D_1 D_2 f = D_2 D_1 f.$$

This generalizes to higher derivatives as well (i.e. in practice, we can always swap derivative operators around).

The Chain Rule

Example. Let $f(x, y) = e^x \sin(xy)$. One can imagine this describes the temperature of the plane at each point (x, y). Now imagine a bug moving in the plane with parametrization $C(t) = (t^2, t^3)$. Then the temperature the bug is feeling at time t is

$$f(C(t)) = e^{t^2} \sin\big(t^5\big).$$