MA213
Sathaye
Notes on Multivariate Functions.

1 Functions of many variables.

1.1 Plotting.

We consider functions like $z = f(x, y)$. Unlike functions of one variable, the graph of such a function has to be in three space and that is usually hard to draw or read! So, one uses the notion of **level curves.** These are plane curves obtained by fixing various values for z and the result is a sequence of mutually non intersecting plane curves covering the plane (or at least the domain of the function).

Thus, for example, for the function $z = x^2 + y^2$, every non negative constant z produces a circle with center at the origin and radius equal to $\sqrt{z}$. This fills the whole plane by these concentric circles. Negative values of z produce no level curves.

We note that we have already studied the surface as an elliptic paraboloid.

We also note that in this particular case, polar coordinates make the equation $z = r^2$ and this can indeed be plotted in the plane. However, each point of this "curve" matches a whole circle in the original graph. Thus, the reduction is not "faithful"!

A similar analysis can be extended to functions of many more variables, however, for a function $w = f(x, y, z)$ the level curves themselves are level surfaces in three space and are not easy to draw!

The important lesson to learn is that we study the resulting geometric object by taking suitable sections by fixing some variables and imagine the object as composed as layers of such sections or pieces. With the aid of computers, both the process of sectioning and reassembling the sections into ordinal objects can be efficiently executed.

1.2 Limits.

Like the functions of one variable, we study limits of functions of two (or more) variables. In a one variable limit $\lim_{x \to 0} f(x)$ it is enough to study the cases as x approaches from the right or left. If both the limits exist and are equal, then the desired limit is said to exist.

For functions of two variables, the number of possible directions increases to infinity and moreover, it is necessary to even allow approaches along other curves which may not be straight lines. We briefly summarize the situation with a few illustrations.

1. First note that limits of polynomials are obtained by evaluation and limits of rational functions are determined similarly, unless we have an indeterminate form $\frac{0}{0}$. Of course, we first make sure that all common factors are canceled. In general, such indeterminate forms lead to no limit (described as limit DNE!) However, in some cases, it is possible to deduce a limit, provided we are working over reals.

   Over complex numbers the limits would fail to exist, unless simplification removes the indeterminate form. The main reason is that along a curve factor of the denominator, the evaluation is zero, but the numerator stays non zero near the point under consideration, leading to non existence of a limit. In real case, the denominator "curve" may simply reduce to a point and we need other considerations.

2. $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2 + xy + y^2}{x^2 - y^2}$. Note that $y - x$ is a factor of the denominator and $y - x = 0$ is a real curve (line). Since it does not cancel out, the answer is DNE!

3. $\lim_{(x,y)\to(0,0)} \dfrac{x^2+xy+y^2}{x^2+y^2}$. Now, the denominator curve reduces to a single point. If we try a line substitution $x = at, y = bt$ then we get $\dfrac{a^2+ab+b^2}{a^2+b^2}$ as the limit! Since this depends on (a, b) the limit DNE!

4. $\lim_{(x,y)\to(0,0)} \dfrac{xy^2}{x^2+y^4}$. Again, the denominator curve reduces to a point. Our line substitution leads to $\dfrac{ab^2}{(a^2+t^2b^4)}t$ which has a limit 0 as t goes to 0. Thus, the line substitution does not contradict existence of a limit. However, a substitution $x = t^2, y = mt$ produces $\dfrac{m^2}{1+m^4}$ and this varies with m. So, the limit DNE.

5. This may lead us to believe that the answer is always DNE! However, we also have: $\lim_{(x,y)\to(0,0)} \dfrac{x^2y}{x^2+y^2}$.

   Here, a polar substitution produces $\dfrac{r^3\cos^2(\theta)\sin(\theta)}{r^2} = r\cos^2(\theta)\sin(\theta)$. Since, we must have $r \to 0$ and since the rest of the factors are bounded, we see that the limit is zero!

   It is instructive to review the $\lim_{(x,y)\to(0,0)} \dfrac{xy^2}{x^2+y^4}$ with a polar substitution.

   We see that after simplification, we get:
   $\dfrac{r^3\cos(\theta)\sin^2(\theta)}{r^2(\cos^2(\theta)+r^2\sin^4(\theta))}$.

   This does simplify to $r\dfrac{\cos(\theta)\sin^2(\theta)}{\cos^2(\theta)+r^2\sin^4(\theta)}$ and for different θ it leads to different limits as r goes to 0.

   For $\theta = 0$ it obviously gives 0, but if we set $r = m\cos(\theta)$ and let θ approach $\pi/2$ we again get a variable limit $\frac{m^2}{1+m^2}$.

   Thus, the limit DNE!

1.3 Continuity

Just as before, a function $z = f(x, y)$ is said to be continuous at a point (a, b) if $\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b)$. Typically, our functions are composed from the usual functions: polynomials, trigonometric, exponential, logarithmic, n-th roots etc. The continuity properties of these are well understood, except when we step out of their natural domains.

Thus, $\sqrt{f(x, y)}$ is continuous wherever f is continuous and also positive. For points where $f(x, y) = 0$, the continuity is a matter of convention. For example the point $(0, 0)$ as a point of the domain of $\sqrt{x+y}$ can be only approached in the half plane $x + y \geq 0$ and some people may not count it in the points of continuity. Some do allow it, but have to recognize the special behavior. However, the function $\sqrt{x^2+y^2}$ is everywhere continuous!

2 Derivatives.

2.1 introduction.

The introductory material on derivatives of functions discusses specific functions $y = f(x)$ and defines **derivative of y with respect to x**. Then it introduces chain rule and teaches that if x is in turn a function $g(t)$ of t, then we can calculate $\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$.

Since we need to handle several functions of several variables, we take a more efficient approach.

We use the symbol D to denote an unspecified derivative operator. Thus, we can write $D(y)$ or simply Dy etc. This derivative is required to satisfy the two basic rules:

- If a is a constant, then $D(a) = 0$,
- $D(ay + bz) = aD(y) + bD(z)$ if a, b are constants **linearity** and
- $D(yz) = yD(z) + zD(y)$ **Product rule.**

These rules alone will let us work out the formula for the derivative of all polynomials in x and a little extra thought will also give the derivatives of all rational functions.

We need a special calculation to get the derivative $D(\sin(\theta)) = \cos(\theta)$ and then derivatives of other trigonometric functions are easily deduced.

Some people define the exponential function by a formula for its derivative $D(\exp(x)) = \exp(x)$ and then derivatives of all exponential and logarithmic functions are corollaries of the above laws.

Thus, the main theory of derivatives consists of interrelating derivatives of related functions.

2.2 Partial derivatives.

Consider a function of two variables $f(x, y) = 2x^2 + 3xy - y^2$. Let us take its derivative using our symbol D which stands for an unspecified derivative. To fix ideas, you may think that both x, y are functions of some unknown t but we don't know the formulas. Then we calculate:

$$D(f(x, y)) = D(2x^2) + D(3xy) - D(y^2) = (2)(2)xD(x) + 3(xD(y) + yD(x)) - (2)yD(y).$$

Collecting terms, we see that

$$D(f(x, y)) = (4x + 3y)D(x) + (3x - 2y)D(y).$$

Now we learn more efficient ways to calculate this answer. For a moment, we assume that $x = t$ and y is a constant. Then we see that $D(y) = 0$ and $D(x) = 1$, so that our answer reduces to $(4x+3y)$.

We shall **assign the term $\mathbf{D_x f(x, y)}$** for the derivative that we get when we take $D(x) = 1$ and $D(y) = 0$.

Similarly, we shall **assign the term $\mathbf{D_y f(x, y)}$** for the derivative that we get when we take $D(x) = 0$ and $D(y) = 1$.

This gives the formula for the universal derivative:

$$D(f(x, y)) = D_x(f(x, y))D(x) + D_y(f(x, y))D(y).$$

Clearly, this can be extended to any number of variables in a similar manner. For example:

$$D(f(x, y, z)) = D_x(f(x, y, z))D(x) + D_y(f(x, y, z))D(y) + D_z(f(x, y, z))D(z).$$

With little practice, these can be calculated in head. Thus, to find D_x we act as if x is the only variable and everybody else is a constant.

We also define a directional derivative $D_v(x,y,z)$ where $v =< a,b,c >$ is any non zero vector. This is done by setting [1]

$$D(x) = \frac{a}{|v|}, D(y) = \frac{b}{|v|}, D(z) = \frac{c}{|v|}.$$

We make a simpler formula as follows.

For a function $f(x,y,z)$ of three variables, we define its gradient denoted as

$$\nabla(f) =< D_x(f(x,y,z)), D_y(f(x,y,z)), D_z(f(x,y,z)) > .$$

then the directional derivative can be calculated as

$$D_v(f) = \nabla(f) \cdot \frac{v}{|v|}.$$

Note that we have dropped the explicit list of variables and thus our formula is valid for any number of variables. As an interesting case, note that in case of a one variable function, we get $D_v(f) = \pm f'(x)$, since for a one dimensional vector v, we have $\frac{v}{|v|} = \pm 1$.

2.3 Higher derivatives.

It is easy to view the partial derivatives as operators (functions acting on functions) so that it is easy to compose them. We shall use algebraic notation and define $D_x^2(f) = D_x(D_x(f)), D_xD_y(f) = D_x(D_y(f))$ etc.

It is possible to create complicated functions for which $D_xD_y(f) \neq D_yD_x(f)$, but under some mild assumptions, this kind of commutativity holds. Clairaut's theorem says that if the two partials $D_xD_y(f), D_yD_x(f)$ are continuous in a disc around a point, then $D_xD_y(f) = D_yD_x(f)$ at the point.

Under such continuity assumptions, we have:

$$\begin{aligned} D^2(f(x,y)) &= D(D_x(f)D(x) + D_y(f)D(y)) \\ &= D_x(D_x(f)D(x) + D_y(f)D(y))D(x) + D_y(D_x(f)D(x) + D_y(f)D(y))D(y) \\ &= D_x^2(f)D(x)^2 + D_y^2(f)D(y)^2 + 2D_x(f)D_y(f)D(x)D(y) + D_x(f)D^2(x) + D_y(f)D^2(y) \end{aligned}$$

It is easy to work out higher order or higher variable rules.

2.4 Alternate notations.

Different people like different notations for the above. We record them for reference. We give only some samples and others can be guessed.

$$D_x(f) = f_x = \frac{\partial f}{\partial x}, D_x^2(f) = f_{xx} = \frac{\partial^2 f}{\partial x^2}.$$

$$D_x(D_y(f)) = D_xD_y(f) = f_{xy} = \frac{\partial^2 f}{\partial x\,\partial y}.$$

$$D(f(x,y)) = f_xD(x) + f_yD(y).$$

[1] The reason to divide by $|v|$ is to make sure that the derivative depends only on the chosen direction and not the chosen vector in the direction.

2.5 Chain Rule.

Our "D" notation, makes short work of many types of chain rules by converting them to simple substitutions.

For example, suppose that we have:

$$z = x^3 + xy + 2y^2, x = 2u - 3v, y = u^2 + v^2.$$

What is $\frac{\partial z}{\partial u}$?

We write basic derivative equations:

$$D(z) = (3x^2 + y)D(x) + (x + 4y)D(y), D(x) = 2D(u) - 3D(v), D(y) = 2uD(u) + 2vD(v)$$

.

Then we get:

$$D(z) = (3x^2 + y)(2D(u) - 3D(v)) + (x + 4y)(2uD(u) + 2vD(v))$$

and collecting coefficients of $D(u)$ we get the answer:

$$\frac{\partial z}{\partial u} = 2(3x^2 + y) + 2u(x + 4y).$$

The answer may be put in terms of u, v alone, if desired, using the given relations.

Another example is this. Suppose that:

$$z^3 + 2zxy + x^2 + xy + yz = 0.$$

Thus, the three variables are related on this surface and at a chosen point, we can express one of them as a function of other two (provided certain partial derivatives are not zero). Suppose, we wish to find $\frac{\partial z}{\partial x}$ when z is a function of x, y.

We actually take the D derivative of the given equation to write

$$(3z^2 + 2xy + y)D(z) + (2zy + 2x + y)D(x) + (2zx + x + z)D(y) = 0.$$

To answer the question, we solve for $D(z)$ and then pick up the coefficient of $D(x)$ in the answer. We quickly see that the answer is:

$$-\frac{(2zy + 2x + y)}{(3z^2 + 2xy + y)}.$$

It also says that z can be thought of a function of x, y exactly when this denominator is non zero!

3 Uses of the Derivative.

3.1 Linear Approximation.

Often, given a function $f(x, y)$ we wish to study its behavior near a point, i.e. we wish to study the behavior of $f(a + h, b + k)$ in comparison with $f(a, b)$ for small values of h, k. It can be shown that

$$f(a+h, b+k) - f(a, b) \approx hD_x(f)(a, b) + kD_y(f)(a, b) \textbf{provided } \mathbf{D_x(f), D_y(f)} \textbf{ are continuous near } \mathbf{(a, b).}$$

Using a natural suggestive notation $\Delta(f) = f(a+h, b+k) - f(a,b)$ and $\Delta(x) = h, \Delta(y) = k$, we rewrite this as:

$$\Delta(f) \approx \Delta(x)D_x(f) + \Delta(y)D_y(f).$$

Indeed, the function $z = f(x,y)$ is said to be **differentiable** at (a,b), if the difference

$$\Delta(f) - (\Delta(x)D_x(f) + \Delta(y)D_y(f)) = \epsilon_1\Delta(x) + \epsilon_2\Delta(y)$$

where ϵ_1, ϵ_2 have limit 0 as $(\Delta(x), \Delta(y))$ approaches 0.

In other words, differentiability implies that our approximation is valid!

In reality, it is difficult to verify the condition directly, but a theorem guarantees that continuity of $D_x(f), D_y(f)$ near (a,b) is enough.

For example, using the function $f(x,y) = \sqrt{x^2+y^2}$ and $(a,b) = (3,4)$; we see that

$$D_x(f)(a,b) = \frac{a}{\sqrt{a^2+b^2}} = \frac{3}{5} \text{ and } D_y(f)(a,b) = \frac{b}{\sqrt{a^2+b^2}} = \frac{4}{5}.$$

Thus,

$$\sqrt{(3+h)^2+(4+k)^2} - \sqrt{3^2+4^2} \approx h\frac{3}{5} + k\frac{4}{5} \text{ for small values of } h, k.$$

So,

$$\sqrt{3.05^2+3.95^2} - 5 \approx (0.05)\frac{3}{5} - (0.05)\frac{4}{5}.$$

Note that this gives an estimate for $\sqrt{3.05^2+3.95^2}$ as $5+0.05(\frac{3}{5}-\frac{4}{5}) = 4.99$, whereas a more precise value is 4.990490958.

We record a general formula for a multivariate function $f(X)$ where $X = (x_1, x_2, \cdots, x_n)$ as follows. Define $\nabla(f)(a) =< D_{x_1}(f)(a), D_{x_2}(f)(a), \cdots, D_{x_n}(a) >$ where $a = (a_1, a_2, \cdots, a_n)$ is a chosen point.

Let $h =< h_1, h_2, \cdots, h_n >$ be a vector of **small numbers.**

Then

$$f(a+h) \approx f(a) + h \cdot \nabla(f)(a).$$

Of course, we assume that the vector of functions $\nabla(f)$ is continuous near a.

Another way to describe the situation is that the linear function:

$$L(X) = f(a) + < X - a > \cdot \nabla(f)(a)$$

has the same value as f when $X = a$ and approximates $f(X)$ when X is near a. It is, therefore called a linear approximation for f near a.

3.2 Tangent Plane.

The above approximation replaces f by a linear function near the chosen point a. The resulting (hyper) surface is called the tangent (hyper)plane of f near a.

We get the equation of the tangent (hyperplane) by setting

$$z = L(X) = f(a) + (X-a) \cdot \nabla(f)(a)$$

For example, for the above example of $f = \sqrt{x^2+y^2}$ at $x = 3, y = 4$, we get

$$z = \sqrt{3^2+4^2} + < x-3, y-4 > \cdot \nabla(f)(3,4) = 5 + \frac{3}{5}(x-3) + \frac{4}{5}(y-4).$$

Note that we have a plane passing through the point $(3, 4, f(3, 4))$ and having the same normal as $\nabla(f)(3, 4)$. These two properties characterize the tangent plane.

It is easy to check that this gives the same old tangent line equation in one variable.

The equation of a tangent plane can also be similarly derived for a surface of the form

$$f(x, y, z) = \lambda \text{ where } \lambda \text{ is a constant.}$$

Taking universal derivatives, we get

$$D(f) = f_x D(x) + f_y(D(y) + f_z D(z) = 0.$$

As before, if we fix a point (a, b, c) and if we replace $< D(x), D(y), D(z0 >$ by $< \Delta(x), \delta(y), \Delta(z) >$, we get an estimated relation between small changes in x, y, z along the surface, namely:

$$f_x(a, b, c)\Delta(x) + f_y(a, b, c)\Delta(y) + f_z(a, b, c)\Delta(z) = 0.$$

We also get an equation of the tangent plane:

$$f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c) = 0.$$

If you prefer the multivariate notation and write the hypersurface as $f(X) = \lambda$, then you could write the general equation as $\nabla(f) \cdot (X - a) = 0$.

To be continued ...