MATH 222: Week 4

Sarah Randall

Last updated: May 25, 2017

Contents

1 §14.5 Chain Rule

The chain rule in 1-dimension is as follows:

For an equation $y = f(x(t))$

$$\frac{dy}{dt} = \frac{df}{dx}\frac{dx}{dt}$$

Example. *If* $y = (x(t))^2$ *and* $x(t) = \ln 1 + t$

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt} = 2x\frac{1}{1+t} = \frac{2\ln 1 + t}{1+t}$$

Example. *Suppose* $f(x, y) = xy + x^2 + y$
$x(t) = \ln 1 + t,\ y(t) = e^{t^2}$
Turn $f(x, y)$ *into a function* $g(t)$ *with only the time parameter.*

$$g(t) = f(x(t), y(t)) = \ln 1 + te^{t^2} + (\ln 1 + t)^2 + e^{t^2}$$

$$\frac{dg}{dt} = \frac{df}{dt} = \frac{e^{t^2}}{1+t} + 2t\ln 1 + te^{t^2} + \frac{2\ln 1 + t}{1+t} + 2te^{t^2}$$

Wherever we can, replace the values of $x(t),\ y(t)$ with $x(t),\ y(t)$.

$$\begin{aligned}\frac{dg}{dt} = \frac{df}{dt} &= \frac{y(t)}{1+t} + 2te^{t^2}x(t) + \frac{2}{1+t} + 2te^{t^2} \\ &= x'(t)y(t) + y'(t)x(t) + 2x'(t)x(t) + y'(t) \\ &= x'(t)(y(t) + 2x(t)) + y'(t)(x(t) + 1) \\ &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\end{aligned}$$

In the second to last line, we use the fact that differentiating f with respect to x gives $y + 2x$ and doing the same for y gives $x + 1$.

There are two possible cases for the chain rule. Suppose in both cases we have $z = f(x, y)$. In the first case we have $x = g(t)$ and $y = h(t)$. In this case z is a differentiable function of t.

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

In the second case, $x = g(s, t)$ and $y = h(s, t)$. Then z is a differentiable function of both s and t.

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

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

Example. $z = e^x\cos(x+y),\ x = s^2t,\ y = st^2$. *Find* $\frac{dz}{ds}$ *and* $\frac{dz}{dt}$.

$$\frac{\partial z}{\partial x} = e^x(\cos(x+y) - \sin(x+y))$$

$$\frac{\partial z}{\partial y} = -e^x\sin(x+y)$$

Then we need $\frac{dx}{ds},\ \frac{dx}{dt},\ \frac{dy}{ds},\ \frac{dy}{dt}$

$$\frac{dx}{ds} = 2st,\ \frac{dx}{dt} = s^2$$

$$\frac{dy}{ds} = t^2,\ \frac{dy}{dt} = 2st$$

So now we can find the general equations for $\frac{dz}{ds}$ *and* $\frac{dz}{dt}$.

$$\frac{dz}{ds} = (e^x)(\cos(x+y) - \sin(x+y))(2st) + (-e^x)(\sin(x+y))(t^2)$$

$$\frac{dz}{dt} = (e^x)(\cos(x+y) - \sin(x+y))(s^2) + (-e^x)(\sin(x+y))(2st)$$

Example. *If* $g(s,t) = f(s^2 - t^2, t^2 - s^2)$ *and* f *is differentiable, show that* $t\frac{dg}{ds} + s\frac{dg}{dt} = 0$.

Based on f, $x(s,t) = s^2 - t^2$ *and* $y(s,t) = t^2 - s^2$. *Therefore we can write that* $g(s,t) = f(x(s,t), y(s,t))$. *We need to find* $\frac{dg}{dt}$ *and* $\frac{dg}{ds}$ *and to do this we need need to differentiate* $x(s,t)$, $y(s,t)$ *by both* s *and* t.

$$\frac{dg}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = \frac{\partial f}{\partial x}(-2t) + \frac{\partial f}{\partial y}(2t)$$

$$\frac{dg}{ds} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} = \frac{\partial f}{\partial x}(2s) + \frac{\partial f}{\partial y}(-2s)$$

If we multiply the first equation all by t *and the second all by* s, *we get*

$$t\frac{dg}{ds} = -2st\frac{\partial f}{\partial x} + 2st\frac{\partial f}{\partial y}$$

$$s\frac{dg}{dt} = 2st\frac{\partial f}{\partial x} - 2st\frac{\partial f}{\partial y}$$

Doing linear combination gives

$$t\frac{dg}{ds} + s\frac{dg}{dt} = 0$$

1.1 Chain rule and implicit functions

In 1 dimension, if we had an implicit function $F(x,y) = 0$ we would do the following to differentiate it.

$$\frac{dF}{dx}(x,y) = 0 \;\Rightarrow\; \frac{dF}{dx}\frac{dx}{dx} + \frac{dF}{dy}\frac{dy}{dx} = 0$$

$\frac{dx}{dx}$ always equals 1, so we get an equation for $\frac{dy}{dx}$.

$$\frac{dy}{dx} = \frac{\frac{-dF}{dx}}{\frac{dF}{dy}}$$

Provided that $\frac{dF}{dy}$ doesn't equal 0.

We can apply this idea to an implicit function like $F(x,y,x)=0$ as well

$$\frac{dF}{dx}=\frac{dF}{dx}\frac{dx}{dx}+\frac{dF}{dy}\frac{dy}{dx}+\frac{dF}{dz}\frac{dz}{dx}=0$$

Like before, $\frac{dx}{dx}$ is 1. In addition, because y is no longer a function of x we can say that $\frac{dy}{dx}=0$. y doesn't depend on x at all. However, z does depend on x so that stays put.

$$0=\frac{dF}{dx}+\frac{dF}{dz}\frac{dz}{dx}$$

$$\frac{dz}{dx}=\frac{\frac{-dF}{dx}}{\frac{dF}{dz}}$$

1.2 Implicit function theorem

Suppose a function $F(x,y,z)$ is defined on a sphere around a point $(a,b,c)\in\mathbb{R}^3$ satisfying $F(a,b,c)=0$.

If $F_x,\ F_y,\ F_z$ are continuous and $\frac{dF}{dz}$ evaluated at (a,b,c) does not equal 0, then in a neighborhood of (a,b,c) we have that the equation $F(x,y,z)=0$ defines z as a function of $x,\ y$ near (a,b,c) in this neighborhood. In addition, this function is differentiable in this area and its partial derivatives $\frac{dz}{dx}=f_x(x,y),\ \frac{dz}{dy}=f_y(x,)$ are given by

$$\frac{dz}{dx}=\frac{\frac{-dF}{dx}}{\frac{dF}{dz}}$$

$$\frac{dz}{dy}=\frac{\frac{-dF}{dy}}{\frac{dF}{dz}}$$

This theorem is really a test that we use to determine if we can get a tangent plane at a given point on the function. If the slope is ∞ then our equations for $\frac{dz}{dx},\ \frac{dz}{dy}$ won't work because $\frac{dF}{dz}$ will be zero.

In clas we used the example of a simple circle $x^2+y^2=1$, a circle of radius 1 centered at the origin. If we choose (a,b) to be somewhere in the middle of the top-right quadrant, in this neighborhood of (a,b) we can talk about the curve as a function. There is no point in this neighborhood where the slope is ∞. However, if we choose (a,b) to be $(1,0)$ or $(-1,0)$ we have a situation where the curve is not a function. We can tell this by looking at the graph of the circle but for more confusing curves we need the Implicit Function Theorem. If $f(a,b)$ has a slope of ∞ then since we're going to be dealing with the neighborhood of (a,b), there will be points in this neighborhood that fail the vertical line test. Then we can't find derivatives and tangent planes in this area since it's not a function here.

Example. *Let $F(x, y, z) = 0$ for $F(x, y, z) = x^3 + y^3 + z^3 + 6zyz - 9$. Show that around $(1, 1, 1)$ we can define z as a function of $x,\ y$. Find the values of $\frac{dz}{dy}$ and $\frac{dz}{dx}$ at $(1, 1, 1)$.*

First check if the point is on the surface. $1 + 1 + 1 + 6 - 9 = 0$, so it's on the surface. Next we need to check that $\frac{dF}{dz}$ at the point isn't 0.

$$\frac{dF}{dz} = 3z^2 + 6xy$$

Evaluating this is the point gives 9, which isn't 0. Since F is a polynomial we can also say that F_x, F_y, F_z are continuous .

By the Implicit Function Theorem, near $(1, 1, 1)$ z is a function of $x,\ y$. Now we need to find $\frac{dz}{dy}$ and $\frac{dz}{dx}$.

$$\frac{dz}{dx} = \frac{-F_x}{F_z} = \frac{-(3x^2 + 6yz)}{3z^2 + 6xy}\bigg|_{1,1,1} = -1$$

$$\frac{dz}{dy} = \frac{-F_y}{F_z} = \frac{-(3y^2 + 6xz)}{3z^2 + 6xy}\bigg|_{1,1,1} = -1$$

1.3 Directional Derivative

Definition. *The directional derivative of $f(x, y)$ at (x_0, y_0) in the direction of a unit vector $\vec{u} =< a, b >$ is*

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

Let $g(h) = f(x + ah, y + bh)$ so that $g(0) = f(x, y)$.

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

We can rewrite this using what we know about the chain rule

$$g'(h) = \frac{\partial f}{\partial x}\frac{dx}{dh} + \frac{\partial f}{\partial y}\frac{dy}{dh}$$

We can say that $x(h) = x_0 + ah$ and $y(h) = y_0 + bh$. From this we can also say that $\frac{dx}{dh} = a$ and $\frac{dy}{dh} = b$. Substitute these into this equation.

$$g'(h) = af_x(x + ah, y + bh) + bf_y(x + ah, y + bh)$$

Set $h = 0$

$$g'(0) = af_x(x, y) + bf_y(x, y)$$

Since we had previously that $g'(0) = D_{\vec{u}}f(x,y)$, we can set these equal and get

$$D_{\vec{u}}f(x,y) = af_x(x,y) + bf_y(x,y)$$

Another common way of writing this is

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

$D_{\vec{u}}f$ is the rate of change of f in direction $\vec{u}$. $\vec{u}$ must be a unit vector because we use its coordinates in our equation. If $\vec{u}$ isn't a unit vector, we won't get the correct answer.

Example. *Compute the directional derivative of $f(x,y) = x^2 + 2y^2 + y$ at $(1,1)$ in direction $< 1,2 >$*

To find $a,\ b$ we need to divide the coordinates of $\vec{u}$ by the vector's length.

$$a = \frac{1}{\| \vec{u} \|} = \frac{\sqrt{5}}{5}$$

$$b = \frac{2}{\| \vec{u} \|} = \frac{2\sqrt{5}}{5}$$

Now if we just find f_x and f_y, we can find the directional derivative.

$$D_{\vec{u}} = < \frac{\sqrt{5}}{5}, \frac{2\sqrt{5}}{5} > \cdot < 2x, 4y+1 >$$

Plug in the point we were given.

$$< \frac{\sqrt{5}}{5}, \frac{2\sqrt{5}}{5} > \cdot < 2,5 > = \frac{12\sqrt{5}}{5}$$

This is the directional derivative at $(1,1)$ in direction $< 1,2 >$.

1.4 The Gradient Vector

A question that naturally follows from this is for what $\vec{u} = < a,b >$ do we get the largest directional derivative? If we want to maximize $< a,b > \cdot < f_x(x,y), f_y(x,y) >$ we should look at an identity we learned earlier in the course.

$$< a,b > \cdot < f_x(x,y), f_y(x,y) > = \| < a,b > \| \ \| < f_x, f_y > \| \cos(\theta)$$

We already know that the length of $< a, b >$ is 1. $< f_x, f_y >$ has a fixed sized. $\cos(\theta)$ is largest at $\theta = 0$. Therefore the directional derivative is largest when $< a, b >$ is parallel to $f_x, f_y >$ (when the angle between them is 0). Therefore $D_{\vec{u}}f$ is maximized when $< a, b >= \frac{< f_x, f_y >}{\| < f_x, f_y > \|}$.

We call this vector in the direction of the maximum change if you start at some point x_0, y_0 the gradient vector:

$$\nabla f(x_0, y_0) =< f_x(x_0, y_0), f_y(x_0, y_0) >$$

We don't necessarily need this to be a unit vector, so we won't divide the vector by its length here like it would be in the directional derivative.

Example. *Find the direction of the maximum derivative for $f(x, y) = x^2 + 2y^2 + y$ and find the value of that max derivative.*

The direction of the max derivative will be the gradient vector. Find f_x, f_y and plug them into our gradient vector equation.

$$\nabla f =< 2x, 4y + 1 > \Rightarrow \ \nabla f(1, 1) =< 2, 5 >$$

This vector gives the direction (we don't really care that it's not a unit vector). However, to find the max value of the directional derivative we'll need to do:

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

In this case, $< a, b >$ is $< \frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}} >$ since we want the directional derivative in this direction. We already know that $< f_x, f_y >$ is $< 2, 5 >$ because we already calculated these values.

$$D_{\nabla f}f(1, 1) = \frac{< 2, 5 >}{\| < 2, 5 > \|} \cdot < 2, 5 >= \frac{\| < 2, 5 > \|^2}{\| < 2, 5 > \|} = \sqrt{29}$$

This gives us valuable information. The maximum directional derivative (directional derivative calculated at ∇f) will be

$$D_{\nabla f}f = \frac{< f_x, f_y >}{\| < f_x, f_y > \|} \cdot < f_x, f_y >= \frac{\| < f_x, f_y > \|^2}{\| < f_x, f_y > \|} =\| < f_x, f_y > \|$$

We can do this because of the rule we learned early in the course that the dot product of a vector and itself equals the magnitude squared of that vector.

Here is an example of a type of problem that might make more sense. As a brief note, the directional derivative in 3D is as follows:

$$D_{\vec{u}}f =< u_1, u_2, u_3 > \cdot \nabla f(x, y, z)$$

Example. *Suppose the temperature at a point* (x, y, z) *in space is given by* $T(x, y, z) = \frac{80}{1+x^2+2y^2+3z^2}$ *in degrees Celsius. In what direction does the temperature increase fastest when starting from* $(1, 1, -2)$*? What is the rate of increase?*

$\nabla T =< T_x, T_y, T_z >$*, so we need* T_x, T_y, T_z

$$T_x = \frac{-160x}{(1 + x^2 + 2y^2 + 3z^2)^2}$$

$$T_y = \frac{-320y}{(1 + x^2 + 2y^2 + 3z^2)^2}$$

$$T_z = \frac{-480y}{(1 + x^2 + 2y^2 + 3z^2)^2}$$

We can factor out a lot of stuff from these, so we end up with the gradient vector

$$\nabla T = \frac{-160x}{(1 + x^2 + 2y^2 + 3z^2)^2} < -x, -2y, -3z >$$

Now plug in $(1, 1, -2)$

$$\nabla T(1, 1, -2) = \frac{5}{8} < -1, -2, 6 >$$

This is the direction for the greatest change. Now let's find the actual rate of that change by using what we learned previously, that the rate of change in the direction of the gradient vector is the magnitude of the gradient vector.

$$\| \nabla T \| = \frac{5}{8}\sqrt{1 + 4 + 36} = \frac{5}{8}\sqrt{41}$$

This is the greatest rate of change possible from $(1, 1 - 2)$ *in degrees Celsius per meter.*

1.5 Significance of gradient vector

Suppose we have a level curve of $z = f(x, y)$ that is $k = f(x, y),\ k \in \mathbb{R}$. If we parametrize $z = f(x, y)$ we can find x and y as a function of t, $x(t),\ y(t)$. Then we can rewrite our level curve as $k = f(x(t), y(t))$. Suppose we then took the derivative of this, using the chain rule.

$$\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = 0$$

Rewriting this as a dot product gives

$$< \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} > \cdot < x'(t), y'(t) >= 0$$

$$\nabla f \cdot < x'(t), y'(t) >= 0$$

This implies that the gradient vector is perpendicular to $\vec{r'}(t)$ (which equals $< x'(t), y'(t) >$ and represents the tangent to the level curve). Therefore the gradient vector at some point on a level curve is perpendicular to the tangent vector along the level curve at that point.

We obtain the same result when dealing with level surfaces. Let $\vec{r}(t)$ be a curve in space with $\vec{r}(t) =< x(t), y(t), z(t) >$ and let $F(x(t), y(t), z(t)) = k$ be a level curve of $\vec{r}(t)$. Let's do the same thing we didn last time and differentiate the level curve.

$$\frac{dF}{dt} = \frac{\partial F}{\partial x}x'(t) + \frac{\partial F}{\partial f}y'(t) + \frac{\partial F}{\partial z}z'(t) = 0$$

We can turn this into a dot product as well.

$$\nabla F \cdot < x'(t), y'(t), z'(t) >= 0$$

$$\nabla F \cdot \vec{r'}(t) = 0$$

Therefore the gradient vector is perpendicular to the tangent of the level surface. The gradient vector is normal to the tangent plane at the point where the gradient vector is evaluated.

So at some point $p(x_0, y_0, z_0)$ we have that the tangent plane to p is

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0$$

Therefore the normal to this plane is $\nabla F =< F_x, F_y, F_z >$.

Suppose we have a surface that can be written in the form $z = f(x, y)$. We can rewrite the equation as $f(x, y) - z = 0$, so that it is technically a function $F(x, y, z)$. Then we can create an equation for the tangent plane at any point (x_0, y_0, z_0) by differentiating the equation.

$$f_x(x - x_0) + f_y(y - y_0) - (z - z_0) = 0$$

z_0 is equals to $f(x_0, y_0)$.

$$z = f_x(x - x_0) + f_y(y - y_0) + f(x_0, y_0)$$

2 §14.7 Maximum and Minimum Values

Given a function $z = f(x, y)$ we would like to find local maximum and minimum of $f(x, y)$ on its domain.

Definition. *A function of two variables has a local max at* (a, b) *if* $f(x, y) \leq f(a, b)$ *for all* (x, y) *near* (a, b)*.* $f(a, b)$ *is the local maximum value. A function of two variables has a local min at* (a, b) *if* $f(x, y) \geq f(a, b)$ *for all* (x, y) *near* (a, b)*.* $f(a, b)$ *is the local minimum value.*

Theorem. *If* f *has a local max/min at point* (a, b) *and* $f_x(a, b)$*,* $f_y(a, b)$ *both exist, then* $f_x(a, b) = f_y(a, b) = 0$*. This implies that* $\nabla f(a, b) = 0$ *and that the tangent plane is horizontal to the x-y plane.*

Definition. *A point* (a, b) *is called a critical point (or stationary point) of* f *if* $f_x(a, b) = f_y(a, b) = 0$*.*

3 §14.8 Lagrange Multipliers