MAT 203 LECTURE OUTLINE 9/22

- Today we'll be looking at two topics: the chain rule and directional derivatives.
- We can think of the partial derivative $f_x(x, y)$ as representing the derivative of f in the x-direction (the direction $\mathbf{i}$), and $f_y(x, y)$ as representing the derivative of f in the y-direction (the direction $\mathbf{j}$), we can consider the *directional derivative* in any direction.
- Pick a unit vector $\mathbf{u} = \langle u_1, u_2\rangle$. If you have a non-unit vector, then divide by the magnitude to make it a unit vector. Then the *directional derivative of f in the direction* $\mathbf{u}$ is defined as
$$D_{\mathbf{u}}f(x, y) = \lim_{t\to 0} \frac{f(x + tu_1, y + tu_2) - f(x, y)}{t}.$$
We can also write $\mathbf{u} = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j}$ for some angle θ, and the above formula becomes
$$D_{\mathbf{u}}f(x, y) = \lim_{t\to 0} \frac{f(x + t\cos(\theta), y + t\sin(\theta)) - f(x, y)}{t}.$$
- Example. Let $f(x, y) = x^2 \sin(2y)$. Find the directional derivative in the direction $\mathbf{v} = 3\mathbf{i} - 4\mathbf{j}$ at the point $(1, \pi/2)$.
 The answer is $D_{\mathbf{v}}f(1, \pi/2) = 8/5$
- In our last lecture, we introduced the idea of a *differentiable* function $f(x, y)$ as one for which the linearization of f at every point gives a good approximation of f near that point. If f is differentiable at a point (x, y) (for example, if the partial derivatives of f are continuous), then we compute the directional derivative from the formula
$$D_{\mathbf{u}}f(x, y) = f_x(x, y)\cos(\theta) + f_y(x, y)\sin(\theta).$$
- Next, the *gradient* of a function $f(x, y)$ is defined as
$$\nabla f(x, y) = f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j} = \langle f_x(x, y), f_y(x, y)\rangle$$
provided these partial derivatives exist.
- In terms of the gradient, the directional derivative of a vector $\mathbf{u}$ can be written as $D_{\mathbf{u}}f(x, y) = \nabla f(x, y) \cdot \mathbf{u}$.
- The significance of the gradient is that it gives the <u>direction of maximum increase</u> of the function. To justify this, note that $\nabla f(x, y) \cdot \mathbf{u}$ is greatest when $\mathbf{u}$ and $\nabla f(x, y)$ point in the same direction. Try to convince yourself geometrically as well that this should be true.
 In addition, say we have a function f and gradient $\mathbf{u} = \nabla f(x, y)$ at the point (x, y), so that $\mathbf{u}$ is the direction of maximum increase. Then
$$D_{\mathbf{u}}f(x, y) = \|\nabla f(x, y)\|.$$
- Example. The temperature in Celcius on a metal plate is given by $T(x, y) = 20 - 4x^2 - y^2$. In what direction from $(2, -3)$ does the temperature increase most rapidly? Try sketching a contour plot of this function and drawing the gradient vector.
- One feature you should notice in the previous example is that the gradient is orthogonal to the level curve passing through the base point $(2, -3)$. This is always the case:
Proposition. If f is differentiable at a point (x, y) and $\nabla f(x, y) \neq \mathbf{0}$, then $\nabla f(x, y)$ is normal to the level curve through (x, y).
- Now let's move on to the chain rule. First we recall the chain rule for one variable:
 If $h(x) = g(f(x))$, then $h'(x) = g'(f(x))f'(x)$. Alternatively, if we write $y = f(x)$, then
$$\frac{dh}{dx}(x) = \frac{dg}{dy}(y)\frac{df}{dx}(x).$$
We can write this more compactly (and identifying the variable y with the function $f(x)$) as
$$\frac{dh}{dx} = \frac{dh}{dy}\frac{dy}{dx}.$$
- For multivariable functions, the concept is the same but more complicated. The big idea is to account for all possible dependencies between the variables (see Figures 13.39 and 13.41 in the book for a diagram).

- In the first case, suppose we have a differentiable function $w = f(x,y)$, and that x, y are differentiable functions of another variable t: $x = g(t)$ and $y = h(t)$. Then the chain rule states that
$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}.$$
- In the second case, suppose we have the same situation, except that x, y are now functions of two variables: $x = g(s,t)$ and $y = h(s,t)$. Assume that all partial derivatives of g and h exist. Then
$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s},$$
$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t}.$$
- The same idea extends to more variables in the expected way. For example, for functions $w = f(x,y,z)$ and $x = g_1(s,t)$. $y = g_2(s,t)$, $z = g_3(s,t)$, we have
$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t}.$$

- Example. Take $w = xy + yz + xz$, with $x = s\cos(t)$, $y = s\sin(t)$, $z = t$. Find $\frac{\partial w}{\partial s}$ and $\frac{\partial w}{\partial t}$ for the values $s = 1$ and $t = 2\pi$.

  You should get $\frac{\partial w}{\partial s} = (y+z)\cos(t) + (x+z)\sin(t)$ and $\frac{\partial w}{\partial s}(1, 2\pi) = y + z = 0 + 2\pi = 2\pi$, and $\frac{\partial w}{\partial t} = (y+z)(-s\sin(t)) + (x+z)(s\cos(t)) + (y+x)(1)$ and $\frac{\partial w}{\partial s}(1, 2\pi) = x + z + y + x = 2 + 2\pi$.
- In Calculus 1, you likely learned about *implicit differentiation*. We are now able to understand it in a more precise way.

  Recall that, in implicity differentiation, you're given some relation between x and y such that $\sin(xy) = x^5 + y^2$ and asked to find $\frac{dy}{dx}$ at some point. By subtracting one side if needed, this can be written in the form $F(x,y) = 0$. So what we really have is the level set of a function of two variables $w = F(x,y)$, with y locally a function of x: $w = F(x,y) = F(x, f(x))$.

  The chain rule gives $\frac{dw}{dx} = F_x(x,y)\frac{dx}{dx} + F_y(x,y)\frac{dy}{dx}$.

  But $\frac{dw}{dx} = 0$ since $w = F(x, f(x)) = 0$ for all x. So $F_x(x,y) + F_y(x,y)\frac{dy}{dx} = 0$.
- We conclude that
$$\frac{dy}{dx} = -\frac{F_x(x,y)}{F_y(x,y)}$$
provided that $F_y(x,y) \neq 0$. This is exactly the rule that is taught for implicit differentiation.