Lecture 15

14.5 - The Chain Rule

Consider a function $\vec{F}(x_1, \cdots, x_n) = \langle f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n) \rangle$

The <u>total derivative</u> of $\vec{F}$ is the matrix (an $m \times n$ matrix)

$$D\vec{F}(x_1, \cdots, x_n) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

We can also take partial derivatives of $\vec{F}$:

$$\frac{\partial \vec{F}}{\partial x_i} = \left\langle \frac{\partial f_1}{\partial x_i}, \frac{\partial f_2}{\partial x_i}, \frac{\partial f_3}{\partial x_i}, \cdots, \frac{\partial f_m}{\partial x_i} \right\rangle$$

<u>Chain Rule</u>: Suppose $\vec{F}(x_1, \cdots, x_n) = \langle f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n) \rangle$ and $\vec{G}(y_1, \cdots, y_p) = \langle g_1(y_1, \cdots, y_p), \cdots, g_n(y_1, \cdots, y_p) \rangle$ and that the range of $\vec{G}$ is inside the domain of $\vec{G}$. If $\vec{G}$ is differentiable at a point $(a_1, \cdots, a_p)$ and $\vec{F}$ is differentiable at $\vec{G}(a_1, \cdots, a_p)$, then

$$D(\vec{F} \circ \vec{G})(a_1, \cdots, a_p) = D\vec{F}(\vec{G}(a_1, \cdots, a_p)) D\vec{G}(a_1, \cdots, a_p)$$

<u>Def</u>: Let $\vec{F}(x_1,\cdots,x_n)=\langle f_1(x_1,\cdots,x_n),\cdots,f_m(x_1,\cdots,x_n)\rangle$ be a function which is defined at $(a_1,\cdots,a_n)$ and at all points arbitrarily close to $(a_1,\cdots,a_n)$. Then we say $\vec{F}$ is <u>differentiable</u> at $\underbrace{(a_1,\cdots,a_n)}_{=\vec{a}}$ if $D\vec{F}(a_1,\cdots,a_n)$ exists and

$$\lim_{\vec{h}\to\vec{0}} \frac{|\vec{F}(\vec{a}+\vec{h})-\vec{F}(\vec{a})-D\vec{F}(\vec{a})\vec{h}|}{|\vec{h}|}=0.$$

Let us reduce the scope of the chain rule a bit by only considering "outside functions" whose output is a real number, i.e., $f=f(x_1,\cdots,x_n)$.

<u>Chain Rule</u>: Let $z=f(x_1,\cdots,x_n)$ and

$\vec{G}(y_1,\cdots,y_p)=\langle g_1(y_1,\cdots,y_p),\cdots,g_n(y_1,\cdots,y_p)\rangle=\langle x_1(y_1,\cdots,y_p),\cdots,x_n(y_1,\cdots,y_p)\rangle$

where the range of $\vec{G}$ is in the domain of f. If $\vec{G}$ is differentiable at $(a_1,\cdots,a_p)$ and f is differentiable at $\vec{G}(a_1,\cdots,a_p)$ then for any y_i $(1\le i\le p)$ we have:

$$\frac{\partial z}{\partial y_i}(a_1,\cdots,a_p)=\frac{\partial}{\partial y_i}(f\circ\vec{G})(a_1,\cdots,a_p)=\nabla f(\vec{G}(a_1,\cdots,a_p))\cdot\frac{\partial\vec{G}}{\partial y_i}(a_1,\cdots,a_p)$$

$$=\frac{\partial f}{\partial x_1}(\vec{G}(a_1,\cdots,a_p))\frac{\partial x_1}{\partial y_i}(a_1,\cdots,a_p)+\cdots+\frac{\partial f}{\partial x_n}(\vec{G}(a_1,\cdots,a_p))\frac{\partial x_n}{\partial y_i}(a_1,\cdots,a_p)$$

$$=\frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial y_i}+\cdots+\frac{\partial z}{\partial x_n}\frac{\partial x_n}{\partial y_i}$$

If we write $\vec{G}(y_1, \cdots, y_p) = \langle x_1(y_1, \cdots, y_p), \cdots, x_n(y_1, \cdots, y_p) \rangle$, then

$$\frac{\partial z}{\partial y_i} = \nabla f(\vec{G}(y_1, \cdots, y_p)) \cdot \frac{\partial \vec{G}}{\partial y_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y_i} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial y_i}$$

Alright, so this looks a bit complicated, but ~~is~~ it isn't. Let's actually compute some stuff.

<u>Ex</u>: Let $z = f(x,y) = x^2 + y^2 + xy$ and suppose $x = \sin t, y = e^t$. Find $\frac{\partial z}{\partial t}$ (we actually can write $\frac{dz}{dt}$ here since, after plugging in for x and y, we get a function of one variable).

<u>Sol</u>: In our case, $\vec{G}(t) = \langle x(t), y(t) \rangle = \langle \sin t, e^t \rangle$, so $\frac{d\vec{G}}{dt} = \langle \cos t, e^t \rangle$. $\nabla f = \langle 2x + y, 2y + x \rangle$, so $\nabla f(\vec{G}(t)) = \langle 2\sin t + e^t, 2e^t + \sin t \rangle$. Finally,

$$\frac{dz}{dt} = \nabla f(\vec{G}(t)) \cdot \frac{d\vec{G}}{dt} = 2\sin^2 t + e^t \sin t + 2e^{2t} + e^t \sin t$$

$$= 2(\sin^2 t + e^t \sin t + e^{2t})$$

$\diamond$

Ex: Let $z = f(x,y) = x^2 - 2xy + y^2$, $x = r\cos\theta$, $y = r\sin\theta$

Find $\frac{\partial z}{\partial r}$ and $\frac{\partial z}{\partial \theta}$.

Sol: Sometimes it's more convienient to leave ∇f in terms of x and y, compute the required dot product, then plug in for x and y. We do this here: First, $\vec{G}(r,\theta) = \langle x(r,\theta), y(r,\theta)\rangle = \langle r\cos\theta, r\sin\theta\rangle$

and $\nabla f = \langle 2x-2y, -2x+2y\rangle$

So, $$\frac{\partial z}{\partial r} = \nabla f(\vec{G}(r,\theta)) \cdot \frac{\partial \vec{G}}{\partial r} = \langle 2x-2y, -2x+2y\rangle \cdot \langle \cos\theta, \sin\theta\rangle$$

$$= (2x-2y)\cos\theta + (-2x+2y)\sin\theta$$

$$= 2r(\cos\theta - \sin\theta)\cos\theta - 2r(\cos\theta - \sin\theta)\sin\theta$$

$$= 2r(\cos\theta - \sin\theta)^2 = 2r(1 - 2\cos\theta\sin\theta)$$

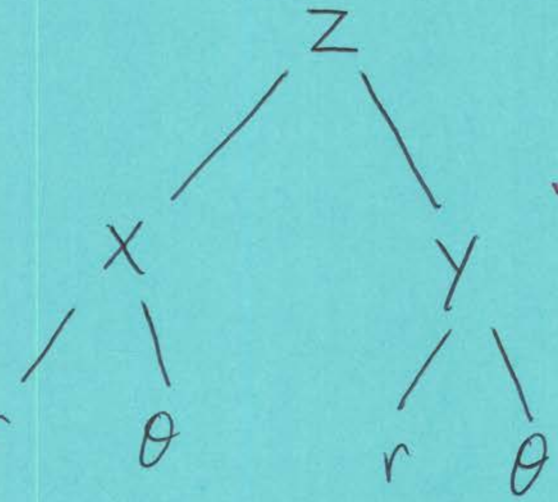
$$= 2r(1 - \sin 2\theta)$$

$$\frac{\partial z}{\partial \theta} = \nabla f(\vec{G}(r,\theta)) \cdot \frac{\partial \vec{G}}{\partial \theta} = \langle 2x-2y, -2x+2y\rangle \cdot \langle -r\sin\theta, r\cos\theta\rangle$$

$$= (2x-2y)(-r\sin\theta) + (-2x+2y)(r\cos\theta)$$

$$= -2r^2(\cos\theta - \sin\theta)\sin\theta - 2r^2(\cos\theta - \sin\theta)\cos\theta = 2r^2(\sin^2\theta - \cos^2\theta) = -2r^2\cos 2\theta$$

There is a useful bookkeeping method we can use for finding derivatives via the chain rule, This is using dependency trees: For the last example we have

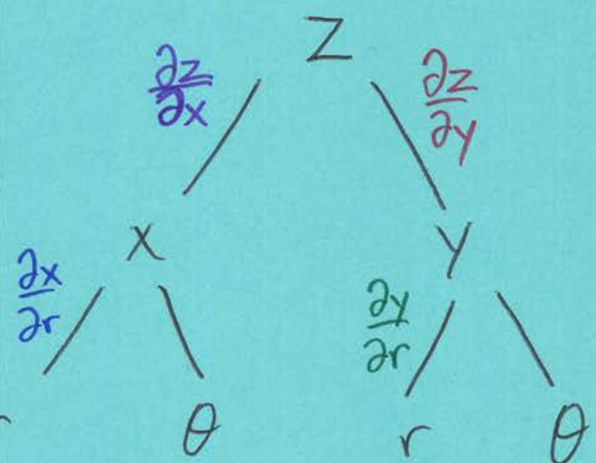


"z depends on x and y"

"x depends on r & θ"

"y depends on r and θ"

Then, to find, say $\frac{\partial z}{\partial r}$, we follow the paths from z to r, each edge corresponding to a derivative to be taken, then we add up paths:

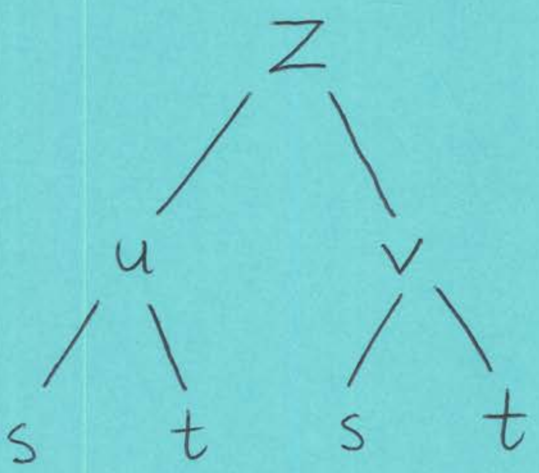


$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r}$$

Let's see an example using this:

<u>Ex</u>: Find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ where $z = \tan\left(\frac{u}{v}\right)$ and $u = 2s+3t$, $v = 3s-2t$.

<u>Sol</u>:



So, $\frac{\partial z}{\partial s} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial s} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial s}$ $\cancel{= \left(\frac{1}{v}\sec^2\left(\frac{u}{v}\right)\right)}$

$$= \left(\frac{1}{v}\sec^2\left(\frac{u}{v}\right)\right)(2) + \left(\frac{-u}{v^2}\sec^2\left(\frac{u}{v}\right)\right)(3)$$

Plug in for u & v $= \frac{2}{3s-2t}\sec^2\left(\frac{2s+3t}{3s-2t}\right) - \frac{3(2s+3t)}{(3s-2t)^2}\sec^2\left(\frac{2s+3t}{3s-2t}\right)$

& $\frac{\partial z}{\partial t} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial t} = \left(\frac{1}{v}\sec^2\left(\frac{u}{v}\right)\right)(3) + \left(\frac{-u}{v^2}\sec^2\left(\frac{u}{v}\right)\right)(-2)$

$$= \left(\frac{3}{v} + \frac{2u}{v^2}\right)\sec^2\left(\frac{u}{v}\right) = \frac{3v+2u}{v^2}\sec^2\left(\frac{u}{v}\right)$$

$$= \frac{3(3s-2t)+2(2s+3t)}{(3s-2t)^2}\sec^2\left(\frac{2s+3t}{3s-2t}\right) = \frac{13s}{(3s-2t)^2}\sec^2\left(\frac{2s+3t}{3s-2t}\right) \diamond$$