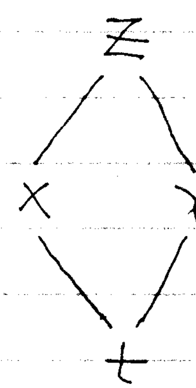
Math 17C
Kouba
Chain Rule, Gradient Vectors, and Directional Derivatives

<u>Chain Rule</u>: assume $z = f(x, y)$, $x = g(t)$, and $y = h(t)$:
Then



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

<u>proof</u>: Consider the linearization $L(x, y)$ for $f(x, y)$ at the point (x_0, y_0):

$$L(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y - y_0).$$

Let $(x_0 + \Delta x, y_0 + \Delta y)$ be a point near (x_0, y_0). The difference of the z-values is

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$$

since $L(x,y) \approx f(x,y)$

$$\approx L(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$$

$$= \cancel{f(x_0,y_0)} + \frac{\partial f}{\partial x}(x_0,y_0)((x_0 + \Delta x) - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)((y_0 + \Delta y) - y_0) - \cancel{f(x_0,y_0)}$$

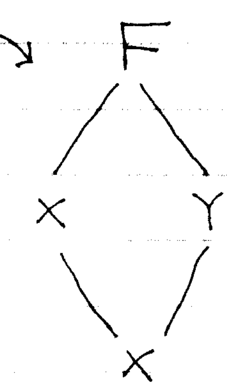
$$\rightarrow \quad \Delta z \approx \frac{\partial f}{\partial x}(x_0, y_0) \cdot \Delta x + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \Delta y$$

$$\rightarrow \quad \frac{\Delta z}{\Delta t} \approx \frac{\partial f}{\partial x}(x_0, y_0) \cdot \frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{\Delta y}{\Delta t}$$

(now let $\Delta t \to 0$.)

$$\rightarrow \quad \frac{dz}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt}. \qquad \text{QED}$$

<u>Note</u>: Suppose that $F(x, y) = c$, a constant, and $y = g(x)$. Then so that



$$\frac{dF}{dx} = F_x \cdot \frac{dx}{dx} + F_y \cdot \frac{dy}{dx} = 0$$

$$\rightarrow \boxed{\frac{dy}{dx} = \frac{-F_x}{F_y}}.$$

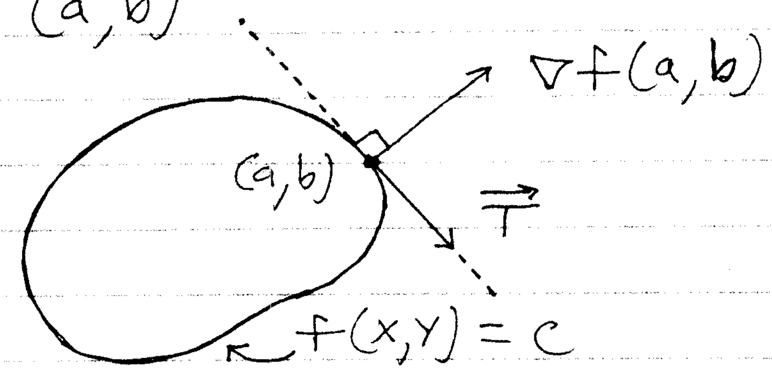
This is an alternative method to implicit differentiation.

<u>Gradient Vector</u>

<u>Def</u>: Let $z = f(x, y)$ be a function. The <u>gradient vector</u> of f at point (x, y) is

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}.$$

Fact: Let $z = f(x,y)$ be a function and let $(x,y) = (a,b)$ with $f(a,b) = c$. Then the gradient vector $\nabla f(a,b)$ is perpendicular to the level curve $f(x,y) = c$ at the point (a,b).



proof: The SLOPE of the tangent line at (a,b) is

$$\frac{dy}{dx} = \frac{-f_x}{f_y} = \frac{\text{rise}}{\text{run}} = \frac{\text{change in } y}{\text{change in } x},$$

so a vector parallel to the tangent line is

$$\vec{T} = \begin{bmatrix} f_y \\ -f_x \end{bmatrix}.$$

(Recall that vectors $\vec{v}$ and $\vec{w}$ are $\perp$ iff $\vec{v} \cdot \vec{w} = 0$.) Then

$$\vec{T} \cdot \nabla f(a,b) = \begin{bmatrix} f_y \\ -f_x \end{bmatrix} \cdot \begin{bmatrix} f_x \\ f_y \end{bmatrix}$$

$$= f_x f_y - f_x f_y = 0.$$

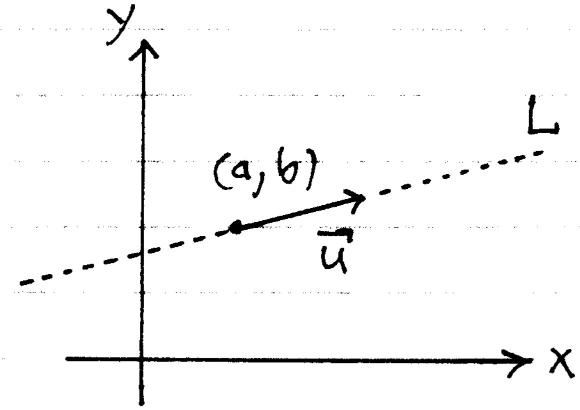
Thus, $\nabla f(a,b)$ is $\perp$ to $f(x,y) = c$.

QED

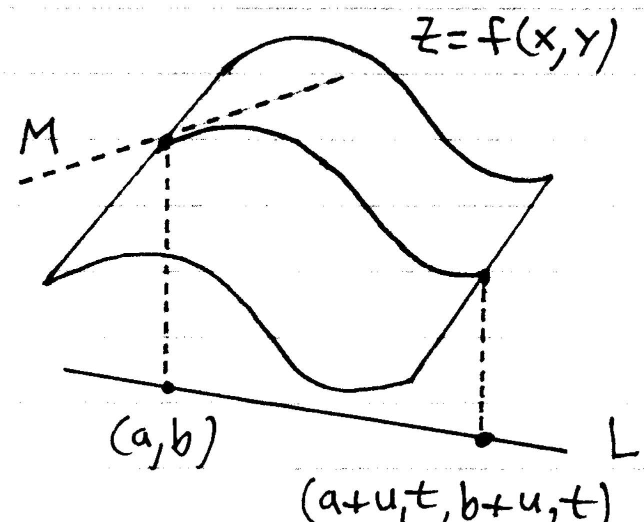
Directional Derivative

Consider a point (a,b) and a unit vector (length = 1) $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$. The line through (a,b) in the direction of $\vec{u}$ is given parametrically by

$$L: \begin{cases} x = a + u_1 t \\ y = b + u_2 t \end{cases} \quad \text{for } -\infty < t < \infty.$$



Let $z = f(x,y)$ be a surface defined on points (x,y) on line L. We want the slope of the line M, which is tangent to surface $z = f(x,y)$ at the point (a,b) in the direction $\vec{u}$.



<u>Note</u>: along line L we have

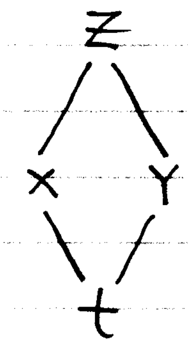
$$z = f(x,y) = f(a+u_1 t, b+u_2 t) = z(t),$$

a function of t only!

Def: The slope of the line tangent to the surface $z = f(x,y)$ at point (a,b) in the direction of unit vector $\vec{u}$ is called the <u>directional derivative</u> of f and is written

$$D_{\vec{u}} f(a,b) \; .$$

<u>Q</u>: How do we compute the value of $D_{\vec{u}} f(a,b)$?



$$D_{\vec{u}} f(x,y) = \frac{d}{dt} z(t)$$

$$= \frac{d}{dt} f(a+u_1 t, b+u_2 t)$$

Chain Rule

$$= f_x(a+u_1 t, b+u_2 t) \cdot \frac{dx}{dt} + f_y(a+u_1 t, b+u_2 t) \cdot \frac{dy}{dt}$$

$$= f_x(a+u_1 t, b+u_2 t) \cdot u_1 + f_y(a+u_1 t, b+u_2 t) \cdot u_2 \; .$$

now let $t=0$ so that $(x,y) = (a,b) \longrightarrow$

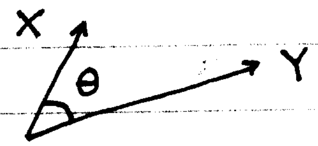
$$D_{\vec{u}} f(a,b) = f_x(a,b) \cdot u_1 + f_y(a,b) \cdot u_2$$

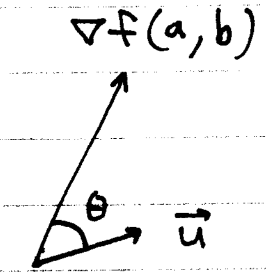
$$= \begin{bmatrix} f_x(a,b) \\ f_y(a,b) \end{bmatrix} \bullet \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} ,$$

i.e.,

$$\boxed{D_{\vec{u}} f(a,b) = \nabla f(a,b) \bullet \vec{u}}$$

<u>Recall</u>: $X \cdot Y = |X||Y| \cos\theta$





Then $D_{\vec{u}} f(a,b) = \nabla f(a,b) \cdot \vec{u}$

$= |\nabla f(a,b)| \, \cancel{|\vec{u}|}^{\,1} \cdot \cos\theta$

$= |\nabla f(a,b)| \cdot \cos\theta$.

This leads to the following facts:

<u>Facts</u>: 1.) $D_{\vec{u}} f(a,b)$ is <u>largest</u> (f is $\uparrow$ most rapidly) when $\theta = 0$ and $\vec{u}$ points in the direction of the gradient vector $\nabla f(a,b)$.

2.) $D_{\vec{u}} f(a,b)$ is <u>smallest</u> (f is $\downarrow$ most rapidly) when $\theta = \pi$ and $\vec{u}$ points in the opposite direction of the gradient vector $\nabla f(a,b)$.