Chain Rule Let $f:\mathbb{R}^n \to \mathbb{R}^m$ and $g:\mathbb{R}^m \to \mathbb{R}^p$ and form the composite $g\circ f:\mathbb{R}^n \to \mathbb{R}^p$. Suppose f is diff. at $x_0 \in \mathbb{R}^n$ and g is diff at $y_0 = f(x_0) \in \mathbb{R}^m$. Then $g\circ f$ is diff. at x_0 and the Jacobian is

$$\underbrace{D(g\circ f)(x_0)}_{p\times n} = \underbrace{Dg(y_0)}_{p\times m} \circ \underbrace{Df(x_0)}_{m\times n}.$$

$\blacksquare$

Special Cases

(a) (Derivative along a path) Let $f:\mathbb{R}^n \to \mathbb{R}$ and $c:\mathbb{R}\to\mathbb{R}^n$. Write $c(t) = (x_1(t), \ldots, x_n(t))$, $x_i:\mathbb{R}\to\mathbb{R}$. Form $h(t) = f\circ c(t)$.

By chain rule:

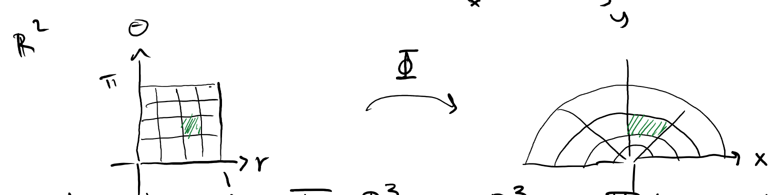
$$\begin{aligned} \frac{dh}{dt} &= D(f\circ c)(t) = Df(c(t))\cdot Dc(t) \\ &= \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial_2} & \cdots & \frac{\partial f}{\partial x_n}\end{bmatrix} \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_n}{dt}\end{bmatrix} \\ &= \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt} \\ &= \nabla f \cdot c' \end{aligned}$$

(b) (Change of Variables/Coordinates) Let $f:\mathbb{R}^2 \to \mathbb{R}$ and $g:\mathbb{R}^2\to\mathbb{R}^2$ w/ $g(x,y) = (u(x,y), v(x,y))$. Form the composite $h(x,y) = f\circ g(x,y)$.

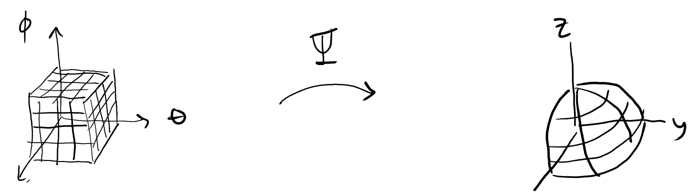
The map g "changes the coordinates" from (u,v) to (x,y).

For example,

(i) (Polar coordinate) $\Phi(r,\theta) = (\underbrace{r\cos\theta}_{x}, \underbrace{r\sin\theta}_{y})$



(ii) (Spherical Coordinates) $\Psi:\mathbb{R}^3\to\mathbb{R}^3$, $\Psi(\rho,\theta,\phi) = (\rho\sin\phi\sin\theta, \rho\sin\phi\cos\theta, \rho\cos\phi)$



$h(x,y) = f \circ g(x,y)$

In this case, the chain rule says

$f: \mathbb{R}^2 \longrightarrow \mathbb{R}$

$g: \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \quad g(x,y) = (u(x,y), v(x,y))$

$$\begin{bmatrix} \underline{\frac{\partial h}{\partial x}} & \underline{\frac{\partial h}{\partial y}} \end{bmatrix} = Dh = Df \cdot \overbrace{Dg}^{2\times 2}$$

$$= \underbrace{\begin{bmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{bmatrix}}_{1\times 2} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} \end{bmatrix}$$

Comparing entries:

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}, \quad \frac{\partial h}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y}.$$

Back to ex (i): Consider $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$, $\Phi: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $\Phi(r,\theta) = (\underbrace{r\cos\theta}_{x}, \underbrace{r\sin\theta}_{y})$.

By chain rule,

$$\begin{bmatrix} \frac{\partial f}{\partial r} & \frac{\partial f}{\partial \theta} \end{bmatrix} = D f \circ \Phi$$

$$= \underline{Df} \cdot D\Phi$$

$$= \begin{bmatrix} \frac{\partial f}{\partial x}, & \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

Compare entries:

$$\frac{\partial f}{\partial r} = \cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y}$$

$$\frac{\partial f}{\partial \theta} = -r\sin\theta \frac{\partial f}{\partial x} + r\cos\theta \frac{\partial f}{\partial y}.$$

■

<u>Ex</u> The chain rule gives a geometric interperetation of the Jacobian as a linear map.

Let $f: \mathbb{R}^2 \to \mathbb{R}^2$ and $c: \mathbb{R} \to \mathbb{R}^2$ be differentiable. The composition $p = f \circ c: \mathbb{R} \to \mathbb{R}^2$ is a new path, the restriction of f to c. the chain rule says

$$p'(t) = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = D(f \circ c)(t)$$

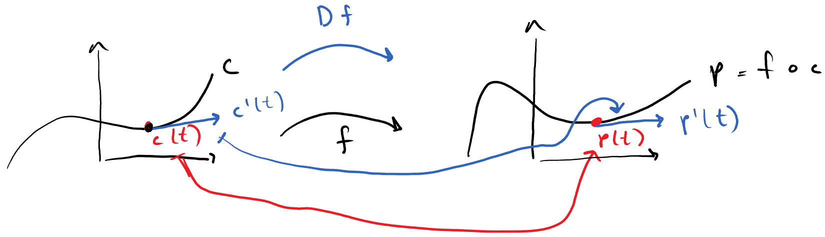
$$p'(t) = \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = D(f\circ c)(t)$$

$$= Df(c(t)) \cdot Dc(t)$$

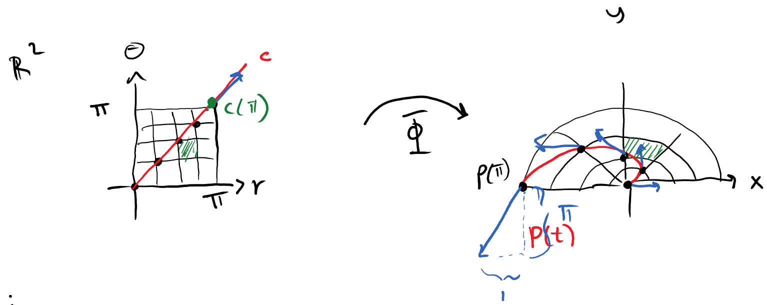
$$= Df(c(t))\ c'(t)$$

tangent to p — tangent to c

This eq. just says that the Jacobian $Df(c(t))$ sends tangent vectors of c to tangent vectors of p.



(Specific Example) Consider $c:\mathbb{R}\to\mathbb{R}^2$, $c(t) = (t,t)$ and $\Phi(r,\theta) = (r\cos\theta, r\sin\theta)$. Form $p(t) = \Phi \circ c(t) = (t\cos t, t\sin t)$



By chain rule:

$$p'(t) = D\Phi(c(t))\ c'(t)$$

$$= \begin{bmatrix} \cos t & -t\sin t \\ \sin t & t\cos t \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

For instance, $p'(\pi) = \begin{bmatrix} -1 & 0 \\ 0 & -\pi \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -\pi \end{bmatrix}$ ■

<u>Gradient</u> Let $f:\mathbb{R}^n \to \mathbb{R}$ be differentiable. The gradient of f is the vector valued function $\nabla f:\mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots\ldots, \frac{\partial f}{\partial x_n}\right)$$

This is just Df, viewed as a vector valued function. ■

<u>Directional Derivative</u> Let $f:\mathbb{R}^3 \to \mathbb{R}$ be differentiable function, and fix $x, v \in \mathbb{R}^n$.

Consider the line $c:\mathbb{R}\to\mathbb{R}^3$ defined by $c(t) = x + tv$. The directional derivative of f at x in the direction of v is derivative of f at x in the direction of v is

$$\frac{d}{dt} f \circ c(t)\Big|_{t=0} = \frac{d}{dt} f(x+tv)\Big|_{t=0}.$$

By special case (1),

$$\begin{aligned}\frac{d}{dt} f \circ c(t)\Big|_{t=0} &= \nabla f(c(t)) \cdot c'(t)\Big|_{t=0}\\ &= \nabla f(c(0)) \cdot c'(0)\\ &= \boxed{\nabla f(x) \cdot v}.\end{aligned}$$

Usually we take v to be a unit vector. ∎

Some Applications

Thm Assume $\nabla f(x) \neq 0$. The $\nabla f(x)$ points in the direction of steepest ascent of f.

Proof Let n be a unit vector. Look at the direction derivative of f in the direction of n:

$$\begin{aligned}\nabla f(x) \cdot n &= \|\nabla f(x)\|\,\|n\| \cos\theta\\ &= \|\nabla f(x)\| \cos\theta.\end{aligned}$$

This is the rate of change of f in the direction of n. This is maximized when $\theta = 0$, i.e., when $\nabla f(x)$ and n are parallel.

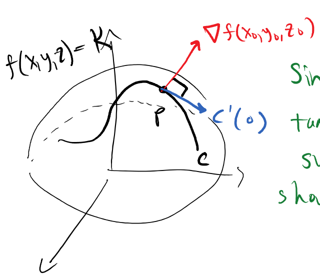
∎

★ Tangent Planes to Level Surfaces

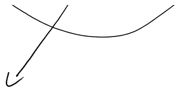
Thm Let $f: \mathbb{R}^3 \to \mathbb{R}$ be differentiable and consider the level surface S defined by $f(x,y,z) = K$, $K \in \mathbb{R}$. If $p = (x_0, y_0, z_0)$ lies on S, then $\nabla f(x_0, y_0, z_0)$ is a vector normal to S at p.

Proof Let $c(t) = (x(t), y(t), z(t))$ be any differentiable curve in S satisfying $c(0) = (x_0, y_0, z_0)$ (c passes through $p = (x_0, y_0, z_0)$ when $t=0$). We need to show that $\nabla f(x_0, y_0, z_0) \cdot c'(0) = 0$. We have

$$\begin{aligned}\nabla f(x_0, y_0, z_0) \cdot c'(0) &= \nabla f(c(0)) \cdot c'(0)\\ &= \frac{d}{dt} f \circ c(t)\Big|_{t=0} && \left(\text{chain rule: special case \#1}\right)\\ &= \frac{d}{dt} K\Big|_{t=0} && \left(f \circ c(t) = K \text{ since } c \text{ lies in the level surface}\right)\\ &= 0. && (K \text{ is a constant})\end{aligned}$$



Since $c'(0)$ is tangent to the surface, $\nabla f(x_0, y_0, z_0)$ should be $\perp$ to $c'(0)$.



should be $\perp$ to $c'(0)$. ` level surface ,

$= 0$. (K is a constant)

Now, we can define the tangent plane to a level surface $f(x,y,z) = K$. A normal vector is $\nabla f(x_0,y_0,z_0)$ if (x_0,y_0,z_0) lies on the surface. So an eq. of the tangent plane is

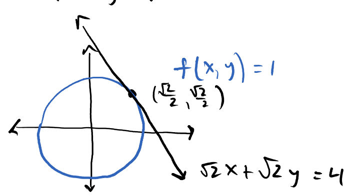
$$\nabla f(x_0,y_0,z_0)\cdot(x-x_0, y-y_0, z-z_0) = 0$$

<u>Note</u> This generalizes to functions $f: \mathbb{R}^n \to \mathbb{R}$: $\nabla f(x_0)$ is normal to level sets $f(x_0) = K$, $K \in \mathbb{R}$ (n-dimensional hypersurfaces). The eq.

$$\nabla f(x_0)\cdot(x - x_0) = 0$$

defines an $(n-1)$-dimensional affine tangent space to the surface. If $n=2$, then this is just the tangent line to the level curves.

<u>Ex</u> $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2+y^2$. The level curves $f(x,y) = K^2$ are circle of radius K:



$\nabla f(x,y) = (2x, 2y)$. The tangent line at $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ is given by

$$\nabla f\left(\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}\right)\cdot\left(x - \tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}\right) = 0$$

$$\Rightarrow \quad \sqrt{2}x + \sqrt{2}y = 4 \iff \boxed{y = 4/\sqrt{2} - x}$$