Today: Chain Rule

$\boxed{\text{Lecture } 14}$

Last time: $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{x}_0$

if
$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + [Df(\vec{x}_0)]\vec{h} + E(\vec{h}) \quad \text{where} \lim_{\vec{h}\to\vec{0}} \frac{E(\vec{h})}{\|\vec{h}\|} = \vec{0}.$$

where

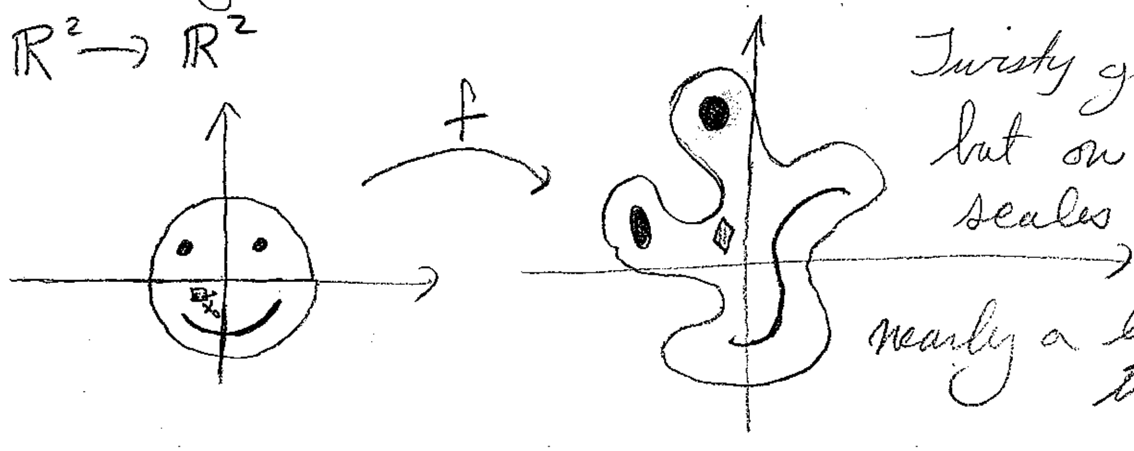
$$Df(\vec{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \frac{\partial f_1}{\partial x_2}(\vec{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}_0) & \cdots & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{x}_0) \end{pmatrix}$$

HW: Section 2.6 # 21, 25, 26

Next time:

[Last time, in the case of a param curve, saw how the derivative could be thought of as the velocity. Here's a way to think about more complicated derivatives.]

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



Twisty globally, but on small scales its nearly a linear transformation.

Ex: $f: \mathbb{R}^2 \to \mathbb{R}^2$ $\quad f(x,y) = (\underbrace{\cos y + x^2}_{= f_1}, \underbrace{e^{x+y}}_{= f_2})$

$$Df(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & -\sin y \\ e^{x+y} & e^{x+y} \end{pmatrix}$$

$$Df(1,0) = \begin{pmatrix} 2 & 0 \\ e & e \end{pmatrix}$$

$g: \mathbb{R}^2 \to \mathbb{R}^2$ $\quad g(u,v) = (e^{u^2}, u - \sin v) = (g_1, g_2)$

$$Dg(u,v) = \begin{pmatrix} 2ue^{u^2} & 0 \\ 1 & -\cos v \end{pmatrix} \qquad Dg(0,0) = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}$$

Consider $f \circ g : \mathbb{R}^2 \to \mathbb{R}^2$

$$f \circ g(u,v) = f(g(u,v)) = f(e^{u^2}, u - \sin v)$$

$$= \left(\cos(u - \sin v) + e^{2u^2}, e^{e^{u^2} + u - \sin v} \right)$$

Compute $D(f \circ g)(\vec{0})$ ...

there has got to be a better way.

Where have you seen this before ???

Recall: $f, g : \mathbb{R} \to \mathbb{R}$. Define $h : \mathbb{R} \to \mathbb{R}$ by $h(x) = f(g(x))$

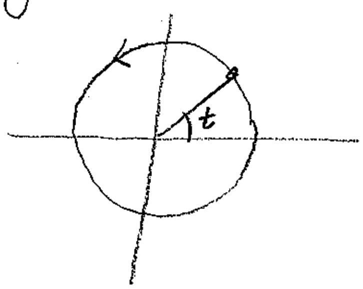
Chain Rule: $h'(x) = f'(g(x))g'(x)$

Ex: $\dfrac{d}{dx} \sin(x^3) = \cos(x^3)\, 3x^2$

temperature, say.

Ex: $g : \mathbb{R} \to \mathbb{R}^2$, $f : \mathbb{R}^2 \to \mathbb{R}$ $\quad f(x,y) = x^2 y$

$g(t) = (\cos t, \sin t)$



$h = f \circ g : \mathbb{R} \to \mathbb{R}$ "temperature observed at time t"

$h(t) = f(\cos t, \sin t) = \cos^2 t \sin t$

$$Dh(t) = \frac{dh}{dt}(t) = -2\cos t \sin^2 t + \cos^3 t.$$

Key: Think in terms of linear approximation.

Want to approx $h(t) = f(g(t))$ near t_0.

Near t_0 we have

$$\begin{pmatrix} g_1'(t_0) \\ g_2'(t_0) \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

$$g(t_0 + h) \approx g(t_0) + [Dg(t_0)]h$$

Near $g(t_0)$ we have

$$f(g(t_0)+\vec{k}) \approx f(g(t_0)) + \left[Df(g(t_0))\right]\vec{k}$$

As

$$Df = \left(\frac{\partial f}{\partial x} \;\; \frac{\partial f}{\partial y}\right) = (2xy \;\; x^2) \text{ we have } (2\cos t_0 \sin t_0, \cos^2 t_0)$$

Combining we get

$$h(t_0+h) = f(g(t_0+h)) \approx f(g(t_0)) + \left[Df(g(t_0))\right]\left[Dg(t_0)\right]h$$

Thus

$$Dh(t_0) = \left[Df(g(t_0))\right]\overset{\text{maxtrix mult.}}{\downarrow}\left[Df(g(t_0))\right]$$

{ Moral: The linear approx of the composite $f \circ g$ is the composit of the linear approximations }

Here

$$(2\cos t_0 \sin t_0 \;\; \cos^2 t_0)\begin{pmatrix} -\sin t_0 \\ \cos t_0 \end{pmatrix} =$$

$$-2\cos t_0 \sin^2 t_0 + \cos^3 t_0$$

which agrees with our formula for Dh from before.

Annother way of thinking about this is

$$g(t) = (x(t), y(t)) \qquad x, y : \mathbb{R} \to \mathbb{R}$$
$$h(t) = f(x(t), y(t))$$

$$\frac{dh}{dt} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{pmatrix} =$$

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

also sometimes write $\frac{\partial h}{\partial x}\frac{dx}{dt} + \frac{\partial h}{\partial y}\frac{dy}{dt}$

<u>General Chain Rule</u>: $g : \mathbb{R}^n \to \mathbb{R}^m$, $f : \mathbb{R}^m \to \mathbb{R}^k$

Suppose g is diff at $\vec{x}_0$ and f is diff at $g(\vec{x}_0)$.
Then $h = f \circ g$ is diff at $\vec{x}_0$ with

$$[Dh(\vec{x}_0)] = [Df(g(\vec{x}_0))][Dg(\vec{x}_0)]$$

Return to the 1^{st} example, $f, g: \mathbb{R}^2 \to \mathbb{R}^2$

Compute $D(f \circ g)$ at $\vec{x}_0 = (0,0)$.

$$Dg(0,0) = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}$$

$$g(0,0) = (1,0) \qquad Df(1,0) = \begin{pmatrix} 2 & 0 \\ e & e \end{pmatrix}$$

$$D(f \circ g) = [Df(g(0,0))][Dg(0,0)] =$$

$$\begin{pmatrix} 2 & 0 \\ e & e \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ e & -e \end{pmatrix}$$