

Math 2010 Week 8

Recall: Matrix multiplication

Let A, B be matrices,

A is $m \times n$, B is $n \times k \Rightarrow AB$ is $m \times k$
 m rows, n columns

$$\text{Let } A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} -\vec{a}_1- \\ \vdots \\ -\vec{a}_m- \end{bmatrix}$$

$$\text{If } b = \begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix} \in \mathbb{R}^n \text{ (written as a column vector)}$$

then

$$Ab = \begin{bmatrix} -\vec{a}_1- \\ \vdots \\ -\vec{a}_m- \end{bmatrix} \begin{bmatrix} | \\ \vec{b} \\ | \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \cdot \vec{b} \\ \vdots \\ \vec{a}_m \cdot \vec{b} \end{bmatrix} \in \mathbb{R}^m$$

Similarly,

$$\begin{bmatrix} -\vec{a}- \end{bmatrix} \begin{bmatrix} | \\ \vec{b}_1 \cdots \vec{b}_k \\ | \end{bmatrix} = [\vec{a} \cdot \vec{b}_1 \cdots \vec{a} \cdot \vec{b}_k]$$

$1 \times n$ matrix (row vector) $n \times k$ B $1 \times k$

$$AB = A \begin{bmatrix} | \\ \vec{b}_1 \cdots \vec{b}_k \\ | \end{bmatrix} = \begin{bmatrix} | \\ A\vec{b}_1 \cdots A\vec{b}_k \\ | \end{bmatrix}$$

$$= \begin{bmatrix} -\vec{a}_1- \\ \vdots \\ -\vec{a}_m- \end{bmatrix} B = \begin{bmatrix} -\vec{a}_1 B- \\ \vdots \\ -\vec{a}_m B- \end{bmatrix}$$

eg

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix} = \begin{bmatrix} 21 & 24 & 27 \\ 47 & 54 & 61 \end{bmatrix}$$

$$A \begin{bmatrix} 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 21 \\ 47 \end{bmatrix} \quad A \begin{bmatrix} 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 24 \\ 54 \end{bmatrix} \quad A \begin{bmatrix} 7 \\ 10 \end{bmatrix} = \begin{bmatrix} 27 \\ 61 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \end{bmatrix} B = \begin{bmatrix} 21 & 24 & 27 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 4 \end{bmatrix} B = \begin{bmatrix} 47 & 54 & 61 \end{bmatrix}$$

Vector-valued Functions

Let $\vec{f}: \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$

$$\vec{f}(\vec{x}) = (f_1(\vec{x}), \dots, f_m(\vec{x})) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$$

vector-valued

Suppose $\frac{\partial f_i}{\partial x_j}(\vec{a})$ exists for each i, j

For each $1 \leq i \leq m$,

$$f_i(\vec{x}) = f_i(\vec{a}) + \nabla f_i(\vec{a}) \cdot (\vec{x} - \vec{a}) + \varepsilon_i(\vec{x}) \quad *$$

$1 \times 1 \quad 1 \times 1 \quad \underbrace{1 \times n \quad n \times 1}_{\text{matrix multiplication}} \quad 1 \times 1$

Here, regard $\nabla f_i(\vec{a})$ as a row vector
 $\vec{x} - \vec{a}$ as a column vector
 in order to use matrix multiplication

Writing $*$ for $1 \leq i \leq m$ in a matrix:

$$\begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix} = \begin{bmatrix} f_1(\vec{a}) \\ \vdots \\ f_m(\vec{a}) \end{bmatrix} + \underbrace{\begin{bmatrix} -\nabla f_1(\vec{a}) - \\ \vdots \\ -\nabla f_m(\vec{a}) - \end{bmatrix}}_{\text{m} \times \text{n matrix of } \frac{\partial f_i}{\partial x_j}} \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon_1(\vec{x}) \\ \vdots \\ \varepsilon_m(\vec{x}) \end{bmatrix}}_{\text{Errors}}$$

Defn Define the Jacobian matrix of \vec{f} at \vec{a} to be

$$D\vec{f}(\vec{a}) = \begin{bmatrix} -\nabla f_1(\vec{a}) - \\ \vdots \\ -\nabla f_m(\vec{a}) - \end{bmatrix} \quad (\text{m} \times \text{n matrix})$$

and the linearization of \vec{f} at \vec{a} to be

$$\vec{L}(\vec{x}) = \vec{f}(\vec{a}) + D\vec{f}(\vec{a})(\vec{x} - \vec{a})$$

\vec{f} is said to be differentiable at \vec{a} if

error term $\vec{\varepsilon}(\vec{x}) = \vec{f}(\vec{x}) - \vec{L}(\vec{x})$

of linearization satisfies $\lim_{\vec{x} \rightarrow \vec{a}} \frac{\|\vec{\varepsilon}(\vec{x})\|}{\|\vec{x} - \vec{a}\|} = 0$

Rmk

$$\textcircled{1} [D\vec{f}(\vec{a})]_{ij} = \frac{\partial f_i}{\partial x_j}(\vec{a})$$

$$\textcircled{2} \begin{matrix} \vec{f}(\vec{x}) & = & \vec{f}(\vec{a}) & + & D\vec{f}(\vec{a})(\vec{x}-\vec{a}) & + & \vec{\varepsilon}(\vec{x}) \\ m \times 1 & & m \times 1 & & m \times n & & n \times 1 & & m \times 1 \end{matrix}$$

$\textcircled{3}$ If f is real-valued ($m=1$), then

$$Df(\vec{a}) = \vec{\nabla} f(\vec{a})$$

$\textcircled{4}$ $\|\vec{\varepsilon}(\vec{x})\|, \|\vec{x}-\vec{a}\|$ is length in $\mathbb{R}^m, \mathbb{R}^n$ resp.

$$\textcircled{5} \lim_{\vec{x} \rightarrow \vec{a}} \frac{\|\vec{\varepsilon}(\vec{x})\|}{\|\vec{x}-\vec{a}\|} = 0 \Leftrightarrow \lim_{\vec{x} \rightarrow \vec{a}} \frac{\varepsilon_i(\vec{x})}{\|\vec{x}-\vec{a}\|} = 0 \quad \forall i=1, \dots, m$$

Hence

\vec{f} is differentiable at \vec{a}	\Leftrightarrow	f_i is differentiable at $\vec{a}, \forall i=1, \dots, m$
--	-------------------	---

Approximation:

$$\vec{f}(\vec{x}) \approx \mathcal{L}(\vec{x}) = \vec{f}(\vec{a}) + D\vec{f}(\vec{a})(\vec{x}-\vec{a})$$

$$\Rightarrow \underbrace{\vec{f}(\vec{x}) - \vec{f}(\vec{a})}_{\Delta \vec{F} = \text{change in } f} \approx \underbrace{D\vec{f}(\vec{a})}_{\text{Jacobian Matrix}} \underbrace{(\vec{x}-\vec{a})}_{\Delta \vec{x} = \text{change in } \vec{x}}$$

Can consider $D\vec{f}(\vec{a})$ as a linear map:

$$D\vec{f}(\vec{a}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

$$\Delta \vec{x} \longmapsto D\vec{f}(\vec{a}) \Delta \vec{x} = d\vec{f}$$

approximated change in f

$$\Delta \vec{F} \approx d\vec{f} = D\vec{f}(\vec{a}) d\vec{x}$$

\uparrow vector \uparrow matrix \uparrow vector

Rmk Compare with $f: \mathbb{R} \rightarrow \mathbb{R}$

$$\Delta y \approx df = f'(a) \Delta x$$

\uparrow number \uparrow \uparrow number

eg $\vec{F}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$

$$\vec{F}(x,y) = [(y+1) \ln x, x^2 - \sin y + 1]$$

$$= \begin{bmatrix} (y+1) \ln x \\ x^2 - \sin y + 1 \end{bmatrix} \quad \begin{array}{l} \text{(Rewrite as)} \\ \text{column vector} \end{array}$$

① Find $D\vec{F}(1,0)$

② Approximate $\vec{F}(0.9, 0.1)$

Sol $f_1(x,y) = (y+1) \ln x$

$$f_2(x,y) = x^2 - \sin y + 1$$

$$\vec{\nabla} f_1 = \left[\frac{y+1}{x} \quad \ln x \right]$$

$$\vec{\nabla} f_2 = [2x \quad -\cos y]$$

$$\Rightarrow D\vec{F}(x,y) = \begin{bmatrix} \frac{y+1}{x} & \ln x \\ 2x & -\cos y \end{bmatrix}$$

$$\therefore D\vec{F}(1,0) = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$$

Linearization of \vec{F} at $(1,0)$

$$\begin{aligned} \vec{L}(x,y) &= \vec{F}(1,0) + D\vec{F}(1,0) \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \\ &= \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x-1 \\ y \end{bmatrix} \end{aligned}$$

$$\vec{F}(0.9, 0.1) \approx \vec{L}(0.9, 0.1)$$

$$= \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 0.9-1 \\ 0.1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} -0.1 \\ -0.3 \end{bmatrix} \quad \Delta \vec{x} = d\vec{x} = \text{change in } \vec{x}$$

$$= \begin{bmatrix} -0.1 \\ 1.7 \end{bmatrix} \quad d\vec{F} = \text{approximated change of } \vec{F}$$

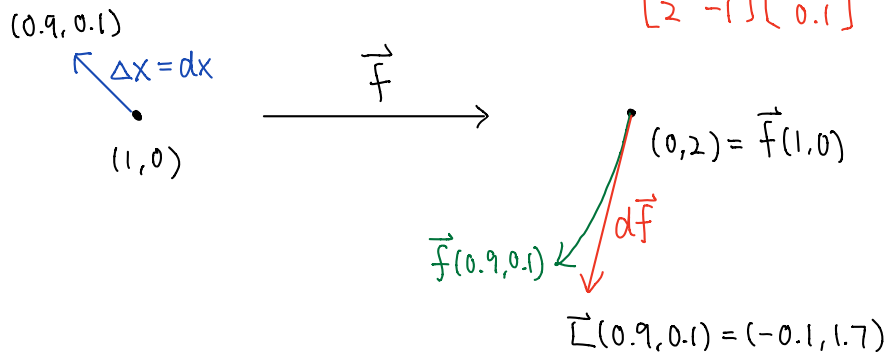
Rmk Actual change in \vec{F} :

$$\Delta \vec{F} = \vec{F}(0.9, 0.1) - \vec{F}(1,0) = \begin{bmatrix} -0.1159... \\ -0.2898... \end{bmatrix}$$

Picture $\vec{f}(1,0) = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ $D\vec{f}(1,0) = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$. Approximate $\vec{f}(0.9, 0.1)$.

$$\Delta \vec{x} = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -0.1 \\ 0.1 \end{bmatrix}$$

$$\begin{aligned} \Delta \vec{f} &\approx d\vec{f} = D\vec{f}(1,0) \Delta \vec{x} \\ &= \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} -0.1 \\ 0.1 \end{bmatrix} = \begin{bmatrix} -0.1 \\ -0.3 \end{bmatrix} \end{aligned}$$



Rmk Total differential can also be written in matrix form

$$f: \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\vec{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} \quad d\vec{f} = \begin{bmatrix} df_1 \\ \vdots \\ df_m \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix} = Df(\vec{a}) d\vec{x}$$

Chain Rule (Au 4.1, 4.2) Thomas 14.4)

Recall One variable:

$$w = g(u) = 2u + 1$$

$$u = f(x) = x^2$$

$$(g \circ f)'(x) = g'(f(x)) f'(x) \quad \text{or}$$

$$\frac{dw}{dx} = \frac{dw}{du} \cdot \frac{du}{dx}$$

$$= 2 \cdot 2x = 4x$$

Rmk We may write

$$g \circ f(x) = g(f(x)) = g(x) \quad (\text{Abuse of notation})$$

$$\frac{dg}{dx} = \frac{dg}{du} \cdot \frac{du}{dx}$$

For multivariable functions:

Thm (Chain Rule)

$$\text{Let } \vec{f}: \Omega_1 \subseteq \mathbb{R}^k \longrightarrow \mathbb{R}^n$$

$$\vec{g}: \Omega_2 \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

Suppose that \vec{f} is differentiable at \vec{a}

\vec{g} is differentiable at $\vec{b} = \vec{f}(\vec{a})$

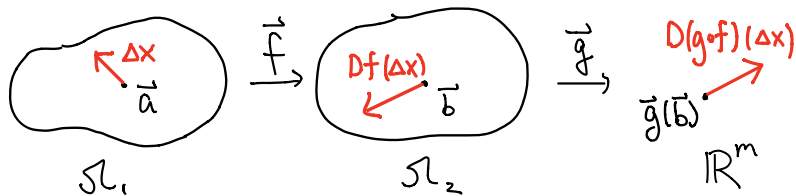
Then $\vec{g} \circ \vec{f}$ is differentiable at \vec{a} ,

$$D(\vec{g} \circ \vec{f})(\vec{a}) = D\vec{g}(f(\vec{a})) D\vec{f}(\vec{a})$$

$m \times k$

$m \times n$

$n \times k$



Rmk For simplicity, we may omit \rightarrow for vectors
from now on: $\vec{f} = f$, $\vec{x} = x$

$$\text{eg } f: \mathbb{R} \rightarrow \mathbb{R}^2, \quad g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$f(\theta) = (\cos\theta, \sin\theta) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

$$g(u, v) = (2uv, u^2 - v^2) = \begin{bmatrix} 2uv \\ u^2 - v^2 \end{bmatrix}$$

Find $D(g \circ f)(\theta)$.

Sol Method 1: Find composition explicitly

$$(g \circ f)(\theta) = g(\cos\theta, \sin\theta)$$

$$= \begin{bmatrix} 2\cos\theta \sin\theta \\ \cos^2\theta - \sin^2\theta \end{bmatrix}$$

$$= \begin{bmatrix} \sin 2\theta \\ \cos 2\theta \end{bmatrix}$$

$$\therefore D(g \circ f)(\theta) = \begin{bmatrix} \frac{d\sin 2\theta}{d\theta} \\ \frac{d\cos 2\theta}{d\theta} \end{bmatrix} = \begin{bmatrix} 2\cos 2\theta \\ -2\sin 2\theta \end{bmatrix}$$

Method 2: Chain Rule

$$Df(\theta) = \begin{bmatrix} f_1' \\ f_2' \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$Dg(u, v) = \begin{bmatrix} \nabla g_1 \\ \nabla g_2 \end{bmatrix} = \begin{bmatrix} 2v & 2u \\ 2u & -2v \end{bmatrix}$$

$$Dg(f(\theta)) = Dg(\cos\theta, \sin\theta) = \begin{bmatrix} 2\sin\theta & 2\cos\theta \\ 2\cos\theta & -2\sin\theta \end{bmatrix}$$

By Chain Rule,

$$D(g \circ f)(\theta) = Dg(f(\theta)) Df(\theta)$$

$$= \begin{bmatrix} 2\sin\theta & 2\cos\theta \\ 2\cos\theta & -2\sin\theta \end{bmatrix} \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} -2\sin^2\theta + 2\cos^2\theta \\ -4\cos\theta \sin\theta \end{bmatrix}$$

$$= \begin{bmatrix} 2\cos 2\theta \\ -2\sin 2\theta \end{bmatrix} \quad \text{(same answer)}$$

eg 2 $f(x,y) = (x^2, 3xy, x+y^2)$

$$g(u,v,w) = \frac{uw}{v}$$

Consider $g \circ f$:

$$\begin{array}{ccc} x & \xrightarrow{f} & \begin{array}{l} f_1 = u \\ f_2 = v \\ f_3 = w \end{array} \\ y & & \xrightarrow{g} g \end{array}$$

Find $\frac{\partial g}{\partial x}(1,1)$.

Rmk \uparrow Regard g as a function of x,y

Sol $Dg = \nabla g = \left[\frac{w}{v} \quad -\frac{uw}{v^2} \quad \frac{u}{v} \right]$

$$Dg(f(1,1)) = Dg(1,3,2)$$

$$= \left[\frac{2}{3} \quad -\frac{2}{9} \quad \frac{1}{3} \right]$$

$$Df = \begin{bmatrix} \nabla f_1 \\ \nabla f_2 \\ \nabla f_3 \end{bmatrix} = \begin{bmatrix} 2x & 0 \\ 3y & 3x \\ 1 & 2y \end{bmatrix}$$

$$Df(1,1) = \begin{bmatrix} 2 & 0 \\ 3 & 3 \\ 1 & 2 \end{bmatrix}$$

$$\therefore D(g \circ f)(1,1)$$

$$= Dg(f(1,1)) Df(1,1)$$

$$= \begin{bmatrix} \frac{2}{3} & -\frac{2}{9} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 3 & 3 \\ 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Note $D(g \circ f) = \begin{bmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}$

$$\therefore \frac{\partial g}{\partial x}(1,1) = 1$$

In last example, $D(g \circ f) = Dg \cdot Df$

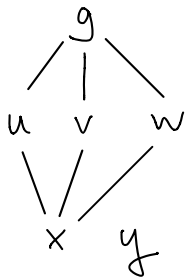
$$\begin{matrix} \frac{\partial g}{\partial x} \\ \downarrow \\ [1 \ 0] \end{matrix} = \begin{bmatrix} \frac{2}{3} & -\frac{2}{9} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{\partial f_1}{\partial x} & & \\ 2 & 0 & \\ 3 & 3 & \\ 1 & 2 & \end{bmatrix} \begin{matrix} f_1 = u \\ f_2 = v \\ f_3 = w \end{matrix}$$

From matrix multiplication, we get another form of chain rule (in classical notation)

$$\begin{aligned} \frac{\partial g}{\partial x} &= \frac{\partial g}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial g}{\partial v} \cdot \frac{\partial v}{\partial x} + \frac{\partial g}{\partial w} \cdot \frac{\partial w}{\partial x} \\ \frac{\partial g}{\partial y} &= \frac{\partial g}{\partial u} \cdot \frac{\partial u}{\partial y} + \frac{\partial g}{\partial v} \cdot \frac{\partial v}{\partial y} + \frac{\partial g}{\partial w} \cdot \frac{\partial w}{\partial y} \end{aligned}$$

Branch Diagram

for $\frac{\partial g}{\partial x}$



Output

Intermediate variables

input

eg $w(x, y, z) = \sqrt{x^2 + y^2 + z^2}$, where

$$\begin{cases} x = 3e^t \sin s \\ y = 3e^t \cos s \\ z = 4e^t \end{cases}$$

Find $\frac{\partial w}{\partial t}$ at $s=t=0$

Sol

$$\begin{aligned} \frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial s} + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial s} \\ &= \frac{x}{\sqrt{x^2 + y^2 + z^2}} \cdot 3e^t \cos s \\ &\quad - \frac{y}{\sqrt{x^2 + y^2 + z^2}} \cdot 3e^t \sin s + \frac{z}{\sqrt{x^2 + y^2 + z^2}} (0) \end{aligned}$$

$$s=t=0 \Rightarrow (x, y, z) = (0, 3, 4)$$

$$\therefore \left. \frac{\partial w}{\partial s} \right|_{(s,t)=(0,0)} = 0 - \frac{3}{5}(0) + 0 = 0$$

eg John is walking with position at time t

$$\text{given by } \begin{cases} x(t) = t^3 + 1 \\ y(t) = 2t^2 \end{cases}$$

Altitude is $H(x,y) = x^2 - y^2 + 100$

① Is John going up/down at $t=1$?

② Which direction should he go instead at $t=1$ to go down most quickly?

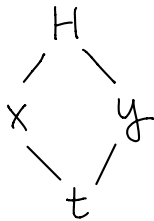
Sol ① Find $\frac{dH}{dt} \Big|_{t=1}$

$$\frac{dH}{dt} = \frac{\partial H}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial H}{\partial y} \cdot \frac{dy}{dt}$$

$$= (2x)(3t^2) + (-2y)(4t)$$

$$= 2(t^3+1)(3t^2) - 2(2t^2)(4t)$$

$$= 6t^5 - 16t^3 + 6t^2$$



$$\therefore \frac{dH}{dt} \Big|_{t=1} = 6 - 16 + 6 = -4 < 0$$

\therefore John is going down at $t=1$

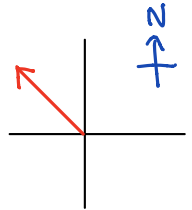
② At $t=1$, $(x,y) = (2,2)$

$$\nabla H = (2x, -2y)$$

$$\nabla H(2,2) = (4, -4)$$

\therefore H decreases most rapidly in the direction of $-\nabla H(2,2) = (-4, 4)$

\therefore John should go NW



Rmk

slope in x - and y -direction

gradient

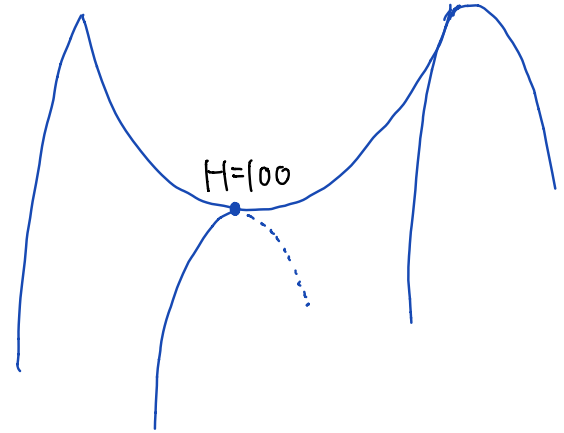
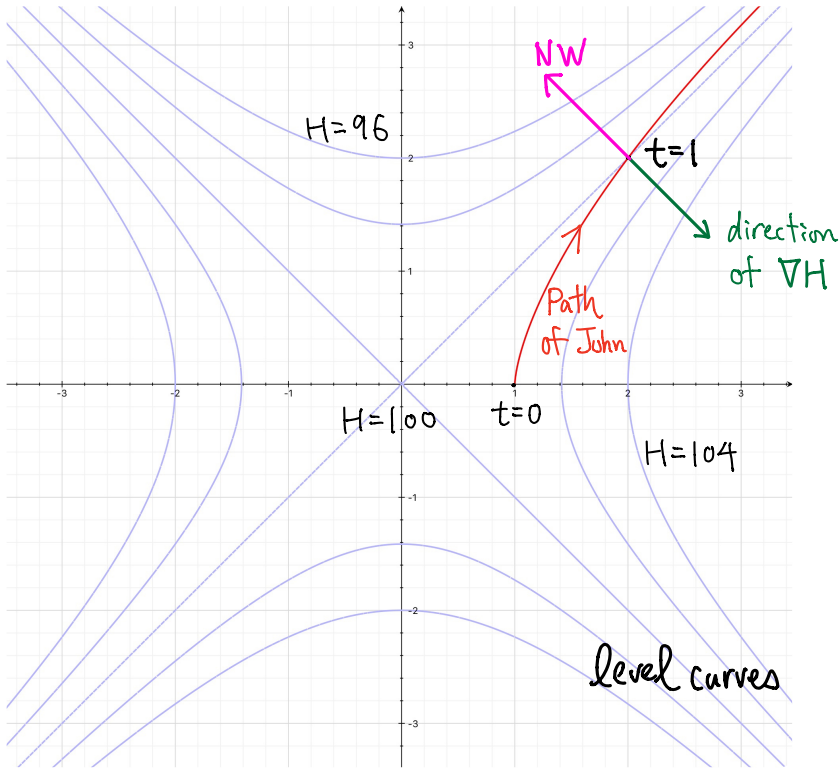
$$\frac{dH}{dt} = \frac{\partial H}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial H}{\partial y} \cdot \frac{dy}{dt} = \nabla H \cdot \begin{bmatrix} \frac{dx}{dt} & \frac{dy}{dt} \end{bmatrix}$$

velocity in x - and y -direction

velocity

Picture Altitude: $H(x,y) = x^2 - y^2 + 100$

$$\text{Position: } \begin{cases} x(t) = t^3 + 1 \\ y(t) = 2t^2 \end{cases}$$



Graph of $z = H(x,y)$

Idea of PF of Chain Rule

Suppose $f: \mathcal{S}_1 \subseteq \mathbb{R}^k \rightarrow \mathbb{R}^n$, differentiable at a

$g: \mathcal{S}_2 \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$, differentiable at $b = f(a) \in \mathcal{S}_2$

For $x \in \mathcal{S}_1$, $f(x) - f(a) = Df(a)(x-a) + \mathcal{E}_f(x) \dots$ (i)

$y \in \mathcal{S}_2$, $g(y) - g(b) = Dg(b)(y-b) + \mathcal{E}_g(y) \dots$ (ii)

Put $y = f(x)$, $b = f(a)$ and (i) into (ii):

$$\begin{aligned} g(f(x)) - g(f(a)) &= Dg(f(a)) [Df(a)(x-a) + \mathcal{E}_f(x)] + \mathcal{E}_g(f(x)) \\ &= \underbrace{Dg(f(a)) Df(a)(x-a)}_{\text{linear in } x-a} + \underbrace{Dg(f(a)) \mathcal{E}_f(x) + \mathcal{E}_g(f(x))}_{\text{let it be } \mathcal{E}_{g \circ f}(x)} \end{aligned}$$

Math 2050 for details
when $m=n=k=1$

Show that $\lim_{x \rightarrow a} \frac{\|\mathcal{E}_{g \circ f}(x)\|}{\|x-a\|} = 0$ $\left(\begin{array}{l} f \text{ continuous at } a \Rightarrow \|f(x) - f(a)\| \text{ small} \Rightarrow \mathcal{E}_g(f(x)) \text{ small} \\ \mathcal{E}_f(x) \text{ small} \Rightarrow Dg(f(a)) \mathcal{E}_f(x) \text{ small} \end{array} \right)$

$\Rightarrow g \circ f$ is differentiable at a with $D(g \circ f)(a) = Dg(f(a)) Df(a)$

Summary: Jacobian Matrix

① $f: \mathcal{U} \subseteq \mathbb{R} \rightarrow \mathbb{R}$ (1-variable, real-valued)

$$x \longmapsto f(x)$$

$$Df(x) = \frac{df}{dx} \quad (\text{scalar, } 1 \times 1 \text{ matrix})$$

② $f: \mathcal{U} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (multivariable, real-valued)

$$x = (x_1, \dots, x_n) \longmapsto f(x) = f(x_1, \dots, x_n)$$

$$Df(x) = \nabla f(x) \quad (\text{vectors in } \mathbb{R}^n)$$
$$= \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \quad (1 \times n \text{ matrix})$$

③ $f: \mathcal{U} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ (multivariable, vector-valued)

$$x = (x_1, \dots, x_n) \longmapsto \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} \quad f_i(x) = f_i(x_1, \dots, x_n)$$

$$Df(x) = \begin{bmatrix} -\nabla f_1 - \\ \vdots \\ -\nabla f_m - \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \quad (m \times n \text{ matrix})$$

Chain Rule (in classical notations)

$$(x_1, \dots, x_k) \xrightarrow{f} (y_1, \dots, y_n) \xrightarrow{g} (g_1, \dots, g_m)$$

$g_i = g_i(y_1, \dots, y_n)$ are functions of y_1, \dots, y_n

$y_j = f_j = f_j(x_1, \dots, x_k)$ are functions of x_1, \dots, x_k

\therefore We can regard $g_i = g_i(x_1, \dots, x_k)$ as functions of x_1, \dots, x_k

Chain rule in matrix form:

$$\begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_k} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_k} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial y_1} & \dots & \frac{\partial g_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial y_1} & \dots & \frac{\partial g_m}{\partial y_n} \end{bmatrix} \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_k} \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1} & \dots & \frac{\partial y_n}{\partial x_k} \end{bmatrix}$$

$m \times k \qquad m \times n \qquad n \times k$

i -th row, j -th column \Rightarrow

$$\frac{\partial g_i}{\partial x_j} = \frac{\partial g_i}{\partial y_1} \frac{\partial y_1}{\partial x_j} + \dots + \frac{\partial g_i}{\partial y_n} \frac{\partial y_n}{\partial x_j}$$