Foundations of Analysis II
Week 6

Domingo Toledo

University of Utah

Spring 2019

$|f(x_0) - f(x_0 - t)|$ $\forall \varepsilon > 0 \ \exists \delta$

$\underbrace{}_{|t| < \delta}$ s.t. $< \varepsilon$ on $|t| < \delta$

$\varepsilon \int_{-\pi}^{\pi} K_N(t)\, dt$

$\forall \varepsilon > 0 \ \exists N_\varepsilon$ s.t. $\underbrace{|\sigma_N(f, x_0) - f(x_0)|}_{} < \varepsilon$

if $N > N_\varepsilon$

abs cont.

$\dfrac{M'}{(N+1)(1-\cos\delta)}$ for $\int_{-\pi}^{-\delta}$ & $\int_{\delta}^{\pi}$

$\dfrac{1}{2\pi}\int_{-\delta}^{\delta} \underbrace{|f(x_0) - f(x_0 - t)|}_{\varepsilon} K_N(t)\, dt$

$< \varepsilon \underbrace{\int_{-\delta}^{\delta} K_N(t)\, dt}_{\leq \int_{-\pi}^{\pi} = 1}$

if δ chosen so that $|f(x_0) - f(x_0 - t)| < \varepsilon$

for $|t| < \delta$) $\left(\int_{-\delta}^{\delta} \right) < \varepsilon$

$\dfrac{M'}{N+1} \dfrac{1}{1-\cos\delta}$ $\exists N_\varepsilon$

s.t. $\delta = \delta_\varepsilon$

$< \varepsilon$

if $N > N_\varepsilon$

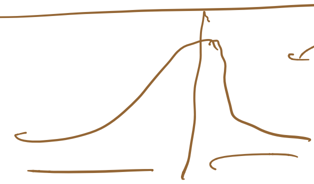
$< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon$.

bdd & cont at x_0

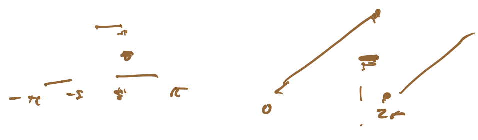
$f(x^+), f(x^-)$

$f(x^+)$ $\qquad \int_{-\pi}^{\pi} f(x-t)\,K$

$f(x^-)$ $\qquad \int_0^{\pi}$



$1/2 \qquad 1/2$



$\delta = \pi/2$

HW due Friday

$$\int_0^{2\pi} x e^{inx} = x \frac{e^{inx}}{in}\Big|_0^{2\pi} - \int$$

$$\sum_1^{\infty} 1/n^2 = \pi^2/6$$

$$\sum 1/n^3 = ??$$

Differentiable Functions of Several Variables

- Simplest Example:
  Linear transformations $A : \mathbb{R}^m \to \mathbb{R}^n$

- Recall Linear algebra vocabulary:

  - Vector space
  - Linear combinations
  - Linear independence
  - Span
  - Basis, dimension

For finite dimensional vector space

Basis: linearly indep / span.

$e_1, \dots, e_n$ a basis for a v.e. X

mean $\forall x \in X \ \exists ! \ a_1, \dots, a_n \in \mathbb{R}$

$$x = a_1 e_1 + \cdots a_n e_n$$

$X = \mathbb{R}^n \quad e_i = (0 \dots \underset{i}{1} \dots 0)$

$$x = (x_1, \dots, x_n) = x_1 e_1 + x_2 e_2 \cdots x_n e_n$$

Span

$v_1, \ldots, v_k$ span X

means every $x \in X$

is a linear comb of $v_1, \ldots, v_k$

$\exists\, b_1, \ldots, b_k \in \mathbb{R}$

s.t. $x = b_1 v_1 + \cdots + b_k v_k$

$\cancel{A}!$ $\exists!$ basis

Basis → Span

→ Lin indep

Lemma a

X finite dimensional

means $\exists$ finite spanning set.

$\exists \{v_1, \ldots, v_k\}$ s.t.

$\forall x \in X$, $\exists\, b_1, \ldots, b_k$

$x = b_1 v_1 + \cdots + b_k v_k$

$\Downarrow$ clean

basis

$\Downarrow$

Define $\dim(X)$

$= \#$ of elements in a basis.

Need: all bases have
same #

Steinitz replacement Thm
(exchange lemma)
in a finite dim X

Suppose $v_1, \ldots, v_m$ independent
$w_1, \ldots, w_n$ spans

$\Rightarrow$ after possibly reordering
$w_1, \ldots, w_n$,

$\{v_1, \ldots, v_m, w_{m+1}, \ldots, w_n\}$

spans X.

One conc'n $m \leq n$

I = indep set

S = spanning set

$\# I \leq \# S$

$\Rightarrow$ i.e. B_1, B_2 are bases

$$\underset{I}{\# B_1} \leq \underset{S}{\# B_2}$$

$$\underset{S}{\# B_1} \geq \underset{I}{\# B_2}$$

$\Rightarrow \# B_1 = \# B_2$

Examples of ($\mathbb{R}$-)VectorSpaces

- Main example: $\mathbb{R}^n$:
  - $\dim(\mathbb{R}^n) = n$
  - Every n-dimensional vector space is isomorphic to $\mathbb{R}^n$
- Another example: $\mathcal{P}^n \subset \mathcal{C}(\mathbb{R}, \mathbb{R})$, the space of polynomials of degree $\leq n$.

$$\mathcal{P}^n = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \ : \ a_0, \ldots a_n \in \mathbb{R}\}$$

dim $= n+1$

$\mathbb{R}$

- Similarly $\mathcal{T}^N \subset \mathcal{C}(\mathbb{R}/\mathbb{Z}, \mathbb{R})$, the space of trigonometric polynomials of degree $\leq N$:

$$\mathcal{T}^N = \{a_0 + \sum_{n=1}^{N} (a_n \cos(nt) + b_n \sin(nt)) \ : \ \underline{a_0}, \underline{a_n}, \underline{b_n} \in \mathbb{R}\}$$

dim: $2N+1$

- What are the dimensions of $\mathcal{P}^n, \mathcal{T}^N$?

- $\mathbb{C}$-versions (complex vector spaces) $\mathbb{C}$
  - Take $a_i \in \mathbb{C}$ in definition of $\mathcal{P}^n$.
  - Take $\sum_{-N}^{N} c_n e^{int}$ to define $\mathcal{T}^N$. $\mathbb{F}$

Some Infinite Dimensional Vector Spaces

$C[0,1]$

P = polys of any degree

~~Every fin~~

Every n-dim vc sp.

$\cong \mathbb{R}^n$

Linear Transformations

- X, Y vector spaces.
- $A : X \to Y$ is a Linear Transformation

$A(x+y) = Ax + Ay \quad \forall\, x, y \in X$

$A(ax) = a(Ax) \quad \forall \left\{ \begin{array}{l} a \in \mathbb{R} \\ x \in X \end{array} \right.$

Examples

$A = id$

$a \quad x \to ax$ mult by a scalar

i

$P^n \to P^{n-1}$

$p \to p'$

$P^n \to P^{n+1}$

$p \to \int_0^x p(t)\,dt$

Linear Transformations and Matrices

- X, Y finite dimensional with bases $\{e_1, \dots, e_m\}$ for X, $\{f_1, \dots, f_n\}$ for Y.

$x \in X$, $\quad x = x_1 e_1 + \cdots + x_m e_m$

$Ax = \sum x_1 \underbrace{\boxed{Ae_1}}_{\in Y} + \cdots + x_m \underbrace{\boxed{Ae_m}}_{\in Y}$

- $A : \mathbb{R}^m \to \mathbb{R}^n$

$Ae_j = \boxed{\sum_{i=1}^{n} a_{ij}} f_i$

matrix of A with

Standard basis

respect to bases

$\{e_1, \dots, e_n\}, \{f_1, \dots, f_n\}$

$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}}_{\in \mathbb{R}^m} = \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}_{\in \mathbb{R}^n}$$

Columns of matrix are Ae_j

Invertible Linear Transformations

- X, Y finite dimensional,
- $\dim(X) = \dim(Y)$,
- $A : X \to Y$ linear.
- Then A is one-to-one $\Leftrightarrow$ A is onto.

$$X \xrightarrow{A} Y \qquad \dim X = \dim Y$$

$e_1, \dots, e_n$ basis for X

$\Rightarrow$ $Ae_1, \dots, Ae_n$ is linearly independent set of # n in Y

$\Rightarrow$ $Ae_1, \dots, Ae_n$ basis for Y

Suppose $\exists$ $a_1, \dots, a_n$

$a_1 Ae_1 + \dots + a_n Ae_n = 0$

$\parallel$

$A(a_1 e_1 + \dots + a_n e_n) = 0$

A 1-1

$\Rightarrow a_1 e_1 + \dots + a_n e_n = 0$

$\Rightarrow a_1 = \dots = a_n = 0$

So every $y \in Y$ $\exists\, a_1, \dots, a_m \in \mathbb{R}$

s.t. $y = a_1 A e_1 + \dots + a_m A e_m$

$= A(\underbrace{a_1 e_1 + \dots + a_m e_m}_{x})$

$\forall y \in Y$
$\exists x \in X$ $Ax = y$ onto

define $A^{-1}G$ $A^{-1} y = x \iff Ax = y$

$A A^{-1} = id_Y$

$A^{-1} A = id_X$

The Space $L(X, Y)$

$$= \{ A : X \to Y : \text{linear} \}$$

is a vector space

$A, B \in L(X,Y)$

$\Rightarrow A+B \in L(X,Y)$

$(A+B)(x) = Ax + Bx$

$(a A)(x) = a(Ax)$

Norm of $A \in L(\mathbb{R}^m, \mathbb{R}^n)$

Def
$$\|A\| = \sup\{|Ax| : |x| = 1\}$$

$x \in \mathbb{R}^m$, $\neq 0$: $|Ax|$

$$= \left|A\left(|x| \frac{x}{|x|}\right)\right| = |x| \left|A\left(\frac{x}{|x|}\right)\right|$$

$$\leq \|A\| \, |x|$$

$L(X, Y)$ $L(Y, Z)$
A B

$BA \in L(X, Z)$

$L(X, X)$ is an algebra.

$$\boxed{|Ax| \le \|A\|\,|x|}$$

another def of norm

$$\|A\| = \inf \{ C : |Ax| \le C|x| \quad \forall x \in \mathbb{R}^m \}$$

$$|Ax| \le C|x|$$

metric space

Lipschitz condition

C Lipschitz constant.

- $A \in L(\mathbb{R}^m, \mathbb{R}^n) \Rightarrow A$ is uniformly continuous.
(In fact A is Lipschitz,with Lipschitz constant$||A||$.)

$$\varepsilon \to \delta = \varepsilon / ||A||$$

- $A, B \in L(\mathbb{R}^m, \mathbb{R}^n) \Rightarrow ||A + B|| \leq ||A|| + ||B||$.

- $A \in L(\mathbb{R}^M, \mathbb{R}^n), B \in L(\mathbb{R}^n, \mathbb{R}^k) \Rightarrow ||BA|| \leq ||B|| \; ||A||$

$$|BAx| \leq ||B|| \, |Ax| \leq ||B|| \, ||A|| \, |x|$$

$|BAx| \le C \|x\|$

$$|BAx| \le \underbrace{\|B\| \, \|A\|}_{C} |x|$$

$$\|BA\| \le \|B\| \, \|A\|$$

- $L(\mathbb{R}^m, \mathbb{R}^n)$ is a *normed* vectorspace.
- $L(\mathbb{R}^n, \mathbb{R}^n)$ is a *normed algebra*

Inversion

- Let $\Omega = \{A \in L(\mathbb{R}^n, \mathbb{R}^n) : A \text{ is invertible }\}$.
- Then Ω is open.

$A \neq I$

$\exists$ Nbd of I consisting of

invertible transf.

$A = I + B$ where $\|B\| < 1$

$$(I+B)^{-1} = \underbrace{I - B + B^2 - B^3 + \cdots}_{\text{converges if } \|B\|<1}$$

$$\|I - B + B^2 - \cdots\| \le 1 + \|B\| + \|B\|^2 + \cdots < \infty \quad \text{if } \|B\|<1$$

A invertible $\qquad A + B = A(I + A^{-1}B)$

$A \in \Omega$, $\{A+B : \|A^{-1}B\| < 1\} \subset \Omega$

- The map $\Omega \to \Omega$ defined by $A \to A^{-1}$ is continuous.

$\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$

$\|A\|$ $\quad$ $\|x\|$

Recall Norm of $A \in L(\mathbb{R}^m, \mathbb{R}^n)$

- $\underbrace{||A|| = \sup\{\frac{|Ax|}{|x|} \;:\; x \in \mathbb{R}^m, x \neq 0\}}$
- Equivalent: $||A|| = \sup\{|Ax| \;:\; x \in \mathbb{R}^m, |x| = 1\}$
- Characterization:

$$\underbrace{||A||} = \inf\{C > 0 \;:\; \boxed{|Ax| \leq C|x|} \text{ for all } x \in \mathbb{R}^m\}$$

Map of metric spaces satisfy this it's called Lipschitz

C : Lipschitz

$\Downarrow$

uniform cont

$d(x,y) = |y-x|$

$|Ax - Ay| \leq C\,|x-y|$

$$\boxed{d(Ax, Ay) \leq C\, d(x,y)}$$

Comparison with other norms

- Suppose $A \in L(\mathbb{R}^m, \mathbb{R}^n)$ has matrix $(a_{i,j})$
- A takes the column vector x with entries $(x_1, \ldots, x_m)$ to the column vector y with entries $(y_1, \ldots, y_n)$ given by the matrix product:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \to \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

$$\boxed{y_i = \sum_{j=1}^{m} a_{i,j}\, x_j}$$

$(a_{ij}) \in \mathbb{R}^{mn}$

another possible norm on $L(X,Y)$

$\left(\sum a_{ij}^2\right)^{1/2}$

- Schwarz inequality gives, for each $i \in \{1, \ldots, n\}$

for i $$y_i^2 \le \Big(\sum_{j=1}^{m} a_{i,j}^2\Big)\Big(\sum_{j=1}^{m} x_j^2\Big)$$

$$\sum y_i^2 = \sum \Big(\quad \Big)$$

- summing over i get

$$\sum_{i=1}^{n} y_i^2 \le \Big(\sum_{i=1,j=1}^{n,m} a_{i,j}^2 \Big) \Big(\sum_{j=1}^{m} x_j^2 \Big)$$

- Thus

$$\boxed{||A|| \le \Big(\sum_{i=1,j=1}^{n,m} a_{i,j}^2 \Big)^{\frac{1}{2}}}$$

- This inequality almost never an equality.

Natural Questions

- When is the last inequality an equality?
- Is there a formula for $||A||$?

if A is diagonal $\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_m \end{pmatrix}$

$$\|A\| = \max\{|d_1|, \ldots, |d_n|\}$$

$$\neq \left(\sum_{i=1}^{n} |d_i|^2\right)^{1/2}$$

$$\boxed{\|A\| = \sqrt{\max |d| \text{ is } d \text{ eigenvalue of } A^t A}}$$

- $A \in L(\mathbb{R}^m, \mathbb{R}^n) \Rightarrow A$ is uniformly continuous.
  (In fact A is Lipschitz,with Lipschitz constant$||A||$.)
- $A, B \in L(\mathbb{R}^m, \mathbb{R}^n) \;\Rightarrow ||A + B|| \leq ||A|| + ||B||$. Δ' ineq.
- $A \in L(\mathbb{R}^m, \mathbb{R}^n)$, $B \in L(\mathbb{R}^n, \mathbb{R}^k) \;\Rightarrow\; ||BA|| \leq ||B||\; ||A||$
- $L(\mathbb{R}^m, \mathbb{R}^n)$ is a *normed* vector space.
- $L(\mathbb{R}^n, \mathbb{R}^n)$ is a *normed algebra*

Invertible Transformations

- Write $L(\mathbb{R}^n)$ for $L(\mathbb{R}^n, \mathbb{R}^n)$.
- Suppose $A \in L(\mathbb{R}^n)$ is invertible, so A^{-1} exists.
- $AA^{-1} = \mathbf{I}$ (the unit matrix)
- Then $\boxed{|| \mathbf{I} || \leq ||A|| \, ||A^{-1}||}$ gives
$$\boxed{||A^{-1}|| \geq \frac{1}{||A||}}$$
- Warning: almost never equality!

$A A^{-1} = I$

$\underbrace{||A A^{-1}||}_{\leq ||A|| \, ||A^{-1}||} = |I| = 1$

- Since $x = A^{-1}Ax$, get $\boxed{|x| \leq ||A^{-1}||\,|Ax|}$
- Equivalently

$$\boxed{|Ax| \geq \frac{1}{||A^{-1}||}|x|} \text{ for all } x \in \mathbb{R}^n.$$

$\Downarrow$

quantitative refinement of

A invertible

$\Downarrow$

$|Ax| \neq 0$

if $x \neq 0$

- We see:

Theorem

- A *is invertible* $\iff$ *there exists a constant* $C > 0$ *so that*

$$\boxed{|Ax| \geq C\,|x|} \quad \text{for all } x \in \mathbb{R}^n$$

$Ax = 0 \Rightarrow x = 0$

- *If* γ *is the supremum of all such* C*, then*

$$\boxed{||A^{-1}|| = \frac{1}{\gamma}} \longrightarrow \inf\{|Ax| : |x| = 1\}$$

Inversion

Rudin: Ω

- Let $\Omega = \{A \in L(\mathbb{R}^n, \mathbb{R}^n) : A \text{ is invertible } \}$.
- Then Ω is open.
- More precisely:

$$A \in \Omega \text{ and } ||A - B|| < \frac{1}{||A^{-1}||} \Rightarrow B \in \Omega$$

- In other words, $A \in \Omega \Rightarrow \mathcal{B}(A, \frac{1}{||A^{-1}||}) \subset \Omega$

$$\sum_{n=1}^{\infty} A_n$$

if $\sum \|A_n\| < \infty$

$\Rightarrow \sum A_n$ converges in Cauchy seq

$\Downarrow \sum_{i}^{n} A_i$

Geometric series

$$(I + A^{-1}(B-A))^{-1}$$

= geom sum

provided

$$\|A^{-1}(B-A)\| < 1$$

find $r>0$ s.t. $\|A-B\|<r$ $\Rightarrow$ B invertible

$A \in S \Rightarrow \exists r>0$ $B(A,r) \subset S$

- Write

$$B = A + (B - A) = A(\mathbf{I} + A^{-1}(B - A))$$

- Thus

$$B^{-1} = (\mathbf{I} + A^{-1}(B - A))^{-1})A^{-1}$$

- The geometric series, if convergent, gives

$$B^{-1} = (\mathbf{I} - (A^{-1}(B - A) + (A^{-1}(B - A))^2 - \dots)A^{-1}$$

if $\|C\|<1 \Rightarrow (I + C)^{-1} = I - C + C^2 - C^3 \dots$

if conv

So if cond for conv $\|I - C + C^2 + \dots (-1)^n C^n\|$ conv (Cauchy seq)

$C =$

$\|C\|<1 \Rightarrow \exists A$

$\le \sum_{n=0}^{\infty} \|C\|^n < \infty$ iff $\|C\|<1$

- The norms of the partial sums are majorized by

$$\sum_{0}^{\infty}(-1)^n(||A^{-1}||\,||B-A||)^n = \frac{1}{1-||A^{-1}||\,||B-A||}$$

- Converges by the assumption $||B-A|| < \frac{1}{||A^{-1}||}$.

- We also get the estimate (Rudin, proof of Thm 9.8)

$$||B^{-1}|| \leq \frac{||A^{-1}||}{1-||A^{-1}||\,||B-A||}$$

$||A^{-1}(B-A)|| < 1$

$\leq ||A^{-1}||\,||B-A||$

- The map $\Omega \to \Omega$ defined by $A \to A^{-1}$ is continuous.
- Fix $A \in \Omega$ and for $B \in \mathcal{B}(A, \frac{1}{||A^{-1}||})$ write

$$B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1}$$

- Get

$$||B^{-1} - A^{-1}|| < \frac{||A^{-1}||^2 \, ||A - B||}{1 - ||A^{-1}|| \, ||B - A||}$$

which converges to 0 as $||A - B|| \to 0$

$\to 0$ as $||A - B|| \to 0$

$B \to A$

$\Rightarrow I$ cont at A.

Inversion is Rational

- From Linear Algebra we know more facts about A^{-1}.
- For example, if A has matrix $(a_{i,j})$, there is a polynomial $\det(A)$ of degree n in the $a_{i,j}$ called the *determinant*.
- $A \in \Omega \iff \det(A) \neq 0$
- Shows Ω is (Zariski) open.

poly $\neq 0$

$\det(a_{ij})$

poly in the n entries

$\det(A) \neq 0$

$\Leftarrow$ A^{-1} exists.

$\Rightarrow \Omega$ is open

$$\Omega = \{A \in L(\mathbb{R}^3) : \det(A) \neq 0\}$$

Formula for A^{-1}

- Given $a \in L(\mathbb{R}^n)$, let $C(A)$ denote the *matrix of cofactors of* A.
- The entries of $C(A)$ are polynomials (of degree $n-1$) in the entries of A.
- Classical formula

$$A^{-1} = \frac{1}{\det(A)} C^t$$

- If $n = 2$ and

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \xrightarrow{\text{tran}} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Cofactor

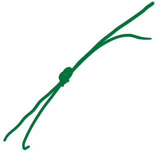
- Then

$$\boxed{A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

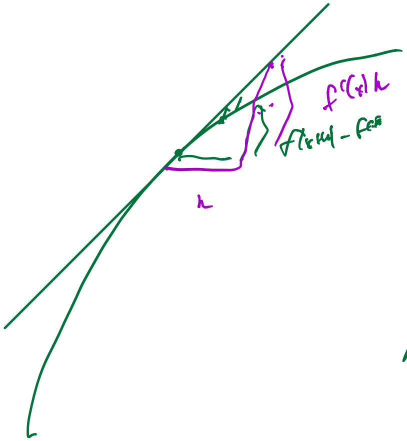
Differentiation in Several Variables

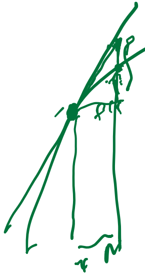
- Setting: $U \subset \mathbb{R}^m$ open, $f : U \to \mathbb{R}^n$, $x \in U$.
- What does it mean for f to be differentiable at x?
- What is the multi-variable analogue of $\quad \mathbb{R} \xrightarrow{f} \mathbb{R}$

$$\boxed{f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}}$$



- Problem: can't divide by $h \in \mathbb{R}^m$ for $m > 1$.





$$f(x+h) - f(x) \approx f'(x)h$$

- Let $r(h) = f(x+h) - f(x) - f'(x)h$

"small"

- Then
$$f(x+h) - f(x) = f'(x)h + r(h)$$
where $r(h)$ is "small".

- How small?

$r(h) \to 0$ as $h \to 0$

$f(x+h) - f(x) - ah = o(h)$

$f(x+h) - f(x) - bh = o(h)$

$(a-b)h = o(h)$

$\frac{o(h)}{h} \to 0$ as $h \to 0$

$\exists$ at most one a

if a exists, unique.

$h/h \to 0$

$\frac{h^2}{h} \to 0$

$a-b = h$

$a-b = 0$

- Need

$$\boxed{\lim_{h\to 0} \frac{r(h)}{h} = 0}$$

- means $r(h) \to 0$ faster than any linear function of h.
- Another notation:

$$r(h) = o(|h|)$$

$f(x) = o(x)$ as $x \to 0$

$\frac{f(x)}{|x|} \to 0$

Definition of differentiability, derivative

- Let $U \subset \mathbb{R}^m$ be $\boxed{\text{open}}$, let $f : U \to \mathbb{R}^n$, let $x \in U$.
- f is *differentiable at x* if there exists a linear transformation
$$A : \mathbb{R}^m \to \mathbb{R}^n$$
so that
$$f(x+h) - f(x) = Ah + o(|h|)$$
Bh

- Equivalent formulation:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Ah}{|h|} = 0$$

$(A-B)h = o(|h|)$

$(A-B)\frac{h}{|h|} \to 0$

A exists ⇒ unique

$|(A-B)h|$

$\le \|(A-B)\| |h|$

$\exists \varepsilon > 0 : \{h : |h| < \varepsilon\}$ for a

$$\frac{\|(A-B)h\|}{|h|} = \|A-B\| \to 0$$

$$\Rightarrow \|A-B\| = 0$$

- If A exists, it is unique
- If A exists, it is called the *derivative of f at x*
- Notation: $\boxed{d_x f}$ or (Rudin) $\boxed{f'(x)}$.

Note: f diff at $x \Rightarrow f$ cont at x

$$f(x+h) - f(x) = Ah + o(|h|)$$

$h \to 0$ $\qquad$ $f(x+h) \to f(x)$ as $h \to 0$ $\qquad$ Continuity

$o(|h|)$ = a vector fun of h

$\varphi(h)$

Hardy Lefflec. $O(1)$

$o(|h|)$ $\quad \dfrac{\varphi(h)}{|h|} \to 0 \iff \dfrac{|\varphi(h)|}{|h|} \to 0$

as $h \to 0$

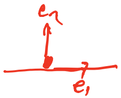
$O(|h|) = \left(\dfrac{\varphi(h)}{|h|}\right) \leq C$

$o(h) = \dfrac{\varphi(h)}{|h|} \to 0$

e geom Rader $\quad \dfrac{|f(x+h) - f(x) - Ah|}{|h|} \to 0$

Partial Derivatives, Jacobian Matrix

$e_1, \dots, e_n$ standard bases for $\mathbb{R}^n$



$$\lim_{t \to 0} \frac{f(x+te_i) - f(x)}{t} = \frac{\partial f}{\partial x_i}(x)$$

if f is diff at x

$$\frac{f(x+te_i) - f(x) - (d_x f)(te_i)}{t} \to 0 \quad \text{as } t \to 0$$

$$\frac{f(x+te_i) - f(x)}{t} - (d_x f)(e_i) \to 0$$

$\frac{\partial f}{\partial x_i}(x)$

$(d_x f)(e_i) = \frac{\partial f}{\partial x_i}$

Jacobian matrix

Standard bases $e_1, \dots, e_m$ for $\mathbb{R}^m$

$\bar{e}_1, \dots, \bar{e}_n$ for $\mathbb{R}^n$

$$(d_x f)(e_i) = \begin{array}{c} i\text{th} \\ \left(\quad \boxed{} \quad \right) \end{array}$$

$$\boxed{d_x f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & & \frac{\partial f_n}{\partial x_m} \end{pmatrix}}$$

$f : \mathbb{R}^m \to \mathbb{R}^n$

$(x_1, \dots, x_m) \to (f_1(x_1, \dots, x_m), f_2(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m))$

f diff at $x \Rightarrow$ all partial derivs exist at x & $d_x f =$ Jacobian matrix

Chain Rule

diff $\Rightarrow$ ∂'s exist.

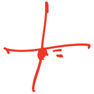
$f: \mathbb{R}^2 \to \mathbb{R}$ $\quad f(x,y) = \cancel{\frac{x^2 y}{x^2+y^2}}$

$$f(x,y) = \begin{cases} \cancel{\dfrac{x^2}{x^2+y^2}} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

$f(x,0) = \dfrac{x^2}{x^2+0} = 1 \qquad \dfrac{\partial f}{\partial x}(0) = 0$

$f(0,y) = \dfrac{0}{y^2} = 0$

$$\frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r} = \cos\theta$$



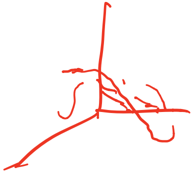
$f(x,0) = 1$

$f(0,y) = 0.$

$\frac{\partial f}{\partial x}(0,0) = 0$

$\frac{\partial f}{\partial y}(0,0) = 0$

$\lim_{(x,y)\to(0,0)} f(x,y)$ doesn't exist



$$(\sqrt{\pi})^n \longrightarrow \boxed{\begin{aligned}&\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{-(x_1^2+\cdots+x_n^2)}\,dx_1\cdots dx_n\\ &= \int_0^{\infty} e^{-r^2}\,\mathrm{vol}_{n-1}(S^{n-1}(1))\, r^{n-1}\,dr\end{aligned}}$$

$$\mathrm{vol}(S^{n-1}(1))\, r^{n-1}$$

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$$

$$f(r) \qquad \int_{\mathbb{R}^n} f(r)\,dV$$

$$= \int_0^{\infty} f(r)\,\mathrm{vol}_{n-1}(S^{n-1}(r))\,dr$$

Recall:Definition of differentiability, derivative

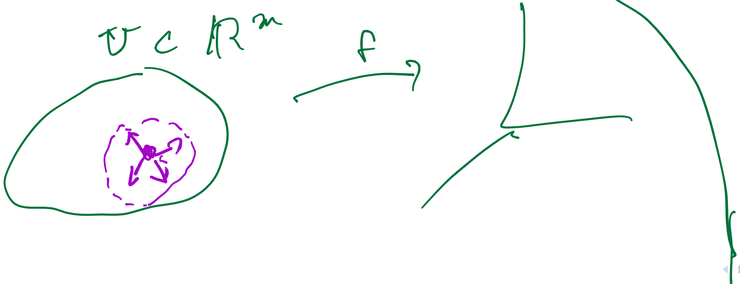
- Let $U \subset \mathbb{R}^m$ be open, let $f : U \to \mathbb{R}^n$, let $x \in U$.
- f is *differentiable at* x if there exists a linear transformation

$$A : \mathbb{R}^m \to \mathbb{R}^n$$

so that

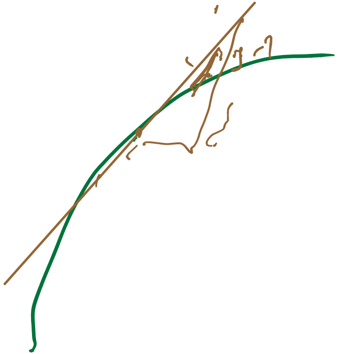
$$\vec{f}(\vec{x} + \vec{h}) - \vec{f}(\vec{x}) = A\vec{h} + o(|\vec{h}|)$$

- Picture



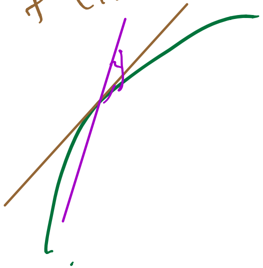
$\mathbb{R} \to \mathbb{R}$

$f(x+h) - f(x)$ is approximately linear



$f(x+h) - f(x) = (\text{linear in } h) + \to 0$ faster than linear

$f'(x)h$



- If A exists, it is unique
- If A exists, it is called the *derivative of f at x*
- Notation: $d_x f$ or (Rudin) $f'(x)$.
- $d_x f$ is the best linear approximation to f at x.
- f differentiable at $x \Rightarrow f$ continuous at x.

$d_x f(h) \in \mathbb{R}^n$

fun of $x \in \mathbb{R}$

$h \in \mathbb{R}^m$

$f(x+h) - f(x)$.

$= \underbrace{Ah}_{\to 0} + \underbrace{o(|h|)}_{\to 0}$

$\Rightarrow f(x+h) - f(x) \to 0$

as $h \to 0$

$\Longrightarrow$ Cont at x

Partial Derivatives, Jacobian Matrix

- f differentiable at $x \Rightarrow$ all partial derivatives of f at x exist.
- In fact.

$$\boxed{\frac{\partial f}{\partial x_i}(x) = d_x f(e_i)}$$

$x = (x_1, \dots, x_m)$

$f = (f_1(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m))$

$$\frac{\partial f}{\partial x_i}(x) = \lim_{t \to 0} \frac{f(x + t e_i) - f(x)}{t}$$

$$= \lim \frac{d_x f(t e_i) + o(|t|)}{t} \leftarrow 0$$

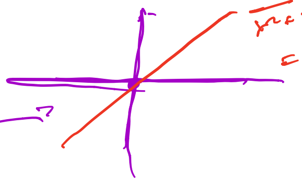
e_i

$= (d_x f)(e_j)$

- Warning: Existence of partials $\not\Rightarrow$ differentiable.
- Example: $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

ex last time not correct

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 \text{ if } (x,y) = (0,0). \end{cases}$$



$\frac{x^2}{x^2+x^2} = 1/2$

- $f(x,0) = 0$ for all $x \Rightarrow \frac{\partial f}{\partial x}(0,0) = 0$.
- Similarly $\frac{\partial f}{\partial y}(0,0) = 0$.
- But $f(x,x) = \frac{1}{2}$ for $x \neq 0 \Rightarrow f$ not continuous at $(0,0)$.

$f(x,y)$ constant on each line through 0

f diff at x

$\to$ linear transf

$d_x f : \mathbb{R}^m \to \mathbb{R}^n$

what's its matrix w.r.t. standard basis

$y = ax$

$$\frac{(ax)\,x}{x^2 + a^2x^2} = \frac{a}{1+a^2}$$

$x = r\cos\theta$

$y = r\sin\theta$

$$\frac{r^2 \cos\theta \sin\theta}{r^2} = \cos\theta \sin\theta$$

$$= \frac{\sin(2\theta)}{2}.$$

$$\lim_{r\to 0} f(r\cos\theta, r\sin\theta) = \frac{\sin(2\theta)}{2}$$

depends on θ

$\theta = 0, \pi/2$

$2\theta = 0, \pi$ $\sin(2\theta) = 0$

$$(d_x f)(e_i) = \frac{\partial f}{\partial x_i}$$

i'th column of matrix

- Suppose f is differentiable at x.

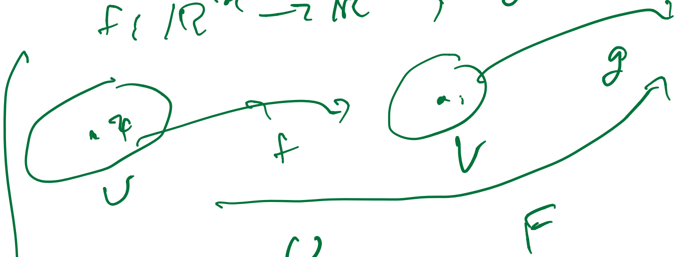
- $d_x f(e_i) = \frac{\partial f}{\partial x_i} \Rightarrow$the matrix of $d_x f$ is the *Jacobian matrix*

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_m} \end{pmatrix}$$

- Note that the i^{th} column is the vector $d_x f(e_i)$, where $\{e_1, \dots, e_m\}$ is the standard basis of $\mathbb{R}^m$.
- More precisely, the entries of the columns are the components of $d_x f(e_i)$ in the standard basis of $\mathbb{R}^n$.

Chain Rule

$f: \mathbb{R}^n \to \mathbb{R}^n$, $g: V \to \mathbb{R}^k$



$x \in U$

Suppose $g \circ f(x) \in V$

$F: U \to \mathbb{R}^k$

$F(x) = g(f(x))$

Suppose f diff at x, g diff at $f(x)$ $\Rightarrow$ F diff at x and

$$d_x F = d_{f(x)} g \left(d_x f \right)$$

$$F(x+h) - F(x) = g\left[f(x+h)\right] - g\left(f(x)\right)$$

$$g\Big(f(x) + \underbrace{\overbrace{\left(f(x+h) - f(x) \right)}^{k}}_{k} \Big) - g\left(f(x)\right)$$

$$= d_{f(x)} g \left(f(x+h) - f(x) \right) + o\left(\underbrace{\left| f(x+h) - f(x) \right|}_{|k|} \right)$$

$$= d_{f(x)} g \left(d_x f(h) + o(|h|) \right) + o(|h|)$$

$$= \underline{(d_{f(x)} g)(d_x f)(h)} + \underbrace{d_{f(x)} g (o(|h|)) + o(|k|)}_{o(|h|)}$$

$$\underline{d_{f(x)} g} \left(\frac{o(|h|)}{|h|} \right)$$

$\downarrow 0$ $\quad \downarrow 0$

$$F(x+h) - F(x) = g(f(x) + k) - g(f(x))$$

$k = f(x+h) - f(x)$

diff of g at $f(x)$:

$$\underline{(d_{f(x)} g)(k) + o(|k|)} \qquad k = k(x, h)$$

$$\underline{k = f(x+h) - f(x) = (d_x f)(h) + o(|h|)}$$

$$= (d_{f(x)} g) \left(d_x f(h) + o(|h|) \right)$$

~~The Gradient~~

$$\underbrace{(d_{f(x)}g)(d_x f(h))}_{\text{Want}} + \underbrace{(d_{f(x)}g)(o(|h|)) + o(|h|)}_{\text{want } o(|h|)}$$

$|h|$

$$\frac{o(|h|)}{|h|} \to 0 \text{ as } |h| \to 0$$

$$(d_{f(x)}g)\left(\frac{o(|h|)}{|h|}\right) + \frac{o(|h|)}{|h|}$$

$\downarrow 0$

$$\frac{o(|k|)}{|h|} = \frac{o(|k|)}{|k|} \frac{|k|}{|h|}$$

$\frac{o(|k|)}{|k|} \to 0$, $\frac{|k|}{|h|} \to$ bdd

$$\frac{f(x+h) - f(x)}{|h|} \to (d_x f)(h)$$

$h \to 0$

$\to 0$

$$\frac{|k|}{|h|} \le \|d_x f\| \frac{|h|}{h} \le \|d_x f\|$$

in terms of Jacobian matrix

Jacob matrix of F at x

$=$ (Jacob of g at $f(x)$) (Jacob f at x)

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \\ \vdots & \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \\ \vdots & \\ \partial \cdots & \end{pmatrix} \begin{pmatrix} & \end{pmatrix}$$

$$f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$$

$$f(x+h) - f(x) = (d_x f)(h) + o(|h|)$$

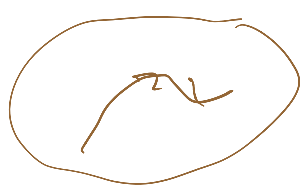
$$\left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_m} \right) \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix}$$

$$= \underbrace{(\vec{\nabla} f)}_{} \cdot \vec{h}$$

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_m} \right)$$

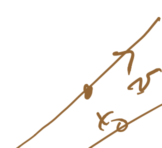
chain rule: $\gamma(t) = (x_1(t), \dots, x_m(t))$

$$\frac{d}{dt} f(\gamma(t)) = \nabla_{\gamma(t)} f \cdot \gamma'(t)$$



in particular $x_0 \in U$

$\underline{x_0 + tv}$ $\quad \|v\| = 1$

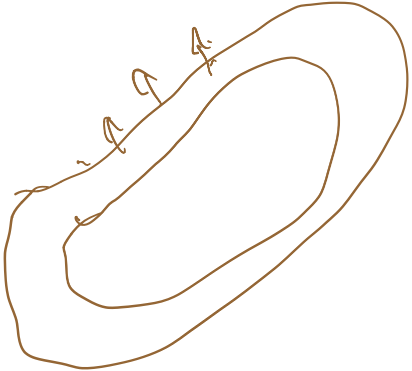


$$\frac{d}{dt} f(x_0 + tv)\Big|_{t=0}$$

chain rule

$$= (\nabla_{x_0} f) \cdot v$$

directional deriv of f at x_0 in direction v



~~level~~

$\gamma(t)$ c-level set

$\Rightarrow f(\gamma(t)) \equiv$ const

$\frac{d}{dt} f(\gamma(t)) \equiv 0$

$\parallel$

$\nabla f_{\gamma(t)} \cdot \gamma'(t) \equiv 0$

$\Rightarrow \nabla_{\gamma(t)} f \perp \gamma'(t)$

$||df||$ bounded $\Rightarrow$ Lipschitz

Theorem

$U \subset \mathbb{R}^m$ open and convex, $f : U \to \mathbb{R}^n$ differentiable. Suppose there is constant M such that

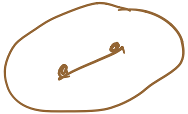
$$\boxed{||d_x f|| \leq M} \text{ for all } x \in U$$

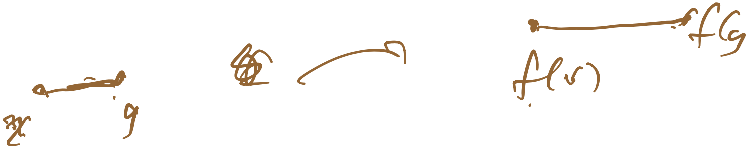
then

$$|f(y) - f(x)| \leq M|y - x| \text{ for all } x, y \in U.$$

U $\quad \forall x, y \in U, \quad \overrightarrow{xy} \subset U$

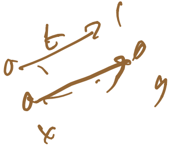
$$\overrightarrow{xy} = \{(1-t)x + ty \; ; \; 0 \leq t \leq 1\}$$





$$|f(x)-f(y)| \leq \underline{\underline{\underline{M}}}\,|x-y|$$

$$f(y) - f(x)$$



$$f((1-t)x + ty)$$

$$\frac{d}{dt} f\left((1-t)x + ty\right)$$

$$= \left(\frac{d}{d\ell(t)} f\right)(y-x)$$

$\ell(t) = (1-t)x + ty$

$\ell'(t) = \cancel{-tx} - x + y$

f scalar function

$\cancel{f(x)}$

$$f(y) - f(x) = f'_{\ell(\xi)}(y-x)$$

inte

Rudin claims

$$\left\{ \begin{aligned} &|f(y) - f(x)| \\ &\leq M |y - x| \end{aligned} \right.$$

$|f(y) - f(x)| \leq M |y-x|$

$|f'(x)| \leq M$

also for $f : \mathbb{R} \to \mathbb{R}$

Theorem

$U \subset \mathbb{R}^m$ open and convex, $f : U \to \mathbb{R}^n$ differentiable. Suppose that

$$d_x f = 0 \text{ for all } x \in U$$

then f is constant.

$M = 0$ works

$|f(y) - f(x)| \leq 0\,|y - x| = 0$

connected enough

C^1 Rudin C'

Functions of class C^1

$f : \overset{\subset \mathbb{R}^m}{U} \to \mathbb{R}^n$ is of class C^1

$\iff$ f diff on U ($d_x f$ exists $\forall x \in U$)

and the map

$U \longrightarrow L(\mathbb{R}^m, \mathbb{R}^n)$

$$x \longrightarrow d_x f$$

is continuous

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{s.t.} \quad |x-y| < \delta \Rightarrow \|d_x f - d_y f\| < \varepsilon$$

Thm (clear) $f \in C^1$

$\iff$ all partial derivatives of f are continuous

$$d_x f \iff \left(\frac{\partial f_i}{\partial x_j}\right) \text{ exists cont}$$

~~Inverse Function Theorem~~

Later

$f \in C^1$

$\ell(t) = (1-t)x + ty$

$$f(y) - f(x) = \int_0^1 \frac{d}{dt} f(\ell(t))\, dt$$

$$= \int_0^1 (d_{\ell(t)} f)(y-x)\, dt$$

$$|f(y) - f(x)| = \left| \int_0^1 (d_{\ell(t)} f)(y-x)\, dt \right|$$

$$\le \int_0^1 |d_{\ell(t)} f (y-x)|\, dt$$

$$\leq \int_0^1 \underbrace{\| d_{c(s)} \|}_{\leq M} \, |(y - x)|$$

$$\leq M \, |(y - x)|$$