Math 140C Real analysis III

- Rayan Saab : rsaab@ucsd.edu
- Book : Rudin (will follow closely)
- Webpage: math.ucsd.edu/~rsaab → teaching → 140C
  Check Regularly for HW/announcement.

- Grading : • HW, Two midterms & a final 20/20/20/40
  • HW, best midterm, & final 20/20/60
  • No makeup exams.

- Your Course Expectations / why taking it / your background

- General Outline :

(I) Functions of several variables

- Linear Transformations
- Differentiation
  Defn → Chain rule → partial derivatives, gradient, ...
- Contraction Principle
- Inverse function theorem
  Roughly: f is invertible around x if $f'(x)$ is invertible
- Implicit function theorem

- Rank / Projection / Determinant
- Higher order Derivatives

(II) Lebesgue Theory

- Set Functions / ring / σ-ring
- Lebesgue measure
- Some measure theory (measure spaces / measurable functions)
- Simple functions → integrations
- Monotone Conv. Theorem
- Dominated Conv. Theorem
- Lebesgue Vs Riemann integral
- L^2 functions / Riesz-Fischer Theorem.

Today: Linear Transformations. Let's remember a few definitions:

Vector space: $X \subset \mathbb{R}^n$ is a vector space if $\forall x, y \in X$ we have $x+y \in X$ & $cx \in X$ (c: scalar)

Linear Combination: The vector $y = \sum_{i=1}^{k} c_i x_i$ is a linear combination of $x_1, \dots, x_k \in \mathbb{R}^n$ (c_i: scalar)

Span: The span of $X \subset \mathbb{R}^n$ is the set of all linear combinations of X. We say X spans E.

Independence: $x_1, \dots, x_k \in \mathbb{R}^n$ are (linearly) independent if $\sum_{i=1}^{k} c_i x_i = 0 \Rightarrow c_i = 0$, $i = 1, \dots, k$

Dimension (of a vector space): The smallest integer r so that any collection of $r+1$ vectors is dependent.
We write $\dim(X) = r$.

Basis (for a vector space X): an independent collection of vectors that spans X.

Remarks: • The 0 vector cannot be in an independent set (why?)

• $\forall x \in X$, if $x_1, \dots, x_r$ form a basis for X then $x = \sum_{i=1}^{r} c_i x_i$ for some c_i. Moreover the representation is unique.

Proof: existance is by the fact that a basis spans

uniqueness: suppose $\sum_{i=1}^{r} c_i x_i = \sum_{i=1}^{r} d_i x_i$ where the d_i's and c_i's differ in at least one index

then $\sum (c_i - d_i) x_i = 0$. By independence of x_i's $c_i = d_i \ \forall i \Rightarrow$ contradiction.

We call the c_i's **coordinates** or **coefficients** of x with respect to $\mathcal{B}$.

Standard basis for $\mathbb{R}^n$: $\{e_1, \dots, e_n\}$ where e_j is 0 except at the jth entry, where it is 1.

Theorem: If vector space X is spanned by r vectors then $\dim(X) \leq r$.

Proof: (Long Version)

Suppose the theorem is false. Then $\exists$ a vector space X spanned by r vectors with $\dim(X) > r$.

So, there exist $r+1$ vectors in X which are independent. Let's call them $y_1 \text{---} y_{r+1}$ & let $Q = \{y_1 \text{---} y_{r+1}\}$

Each y_i, $i \in \{1, \text{---}, r+1\}$ can be written as a linear combination of r vectors, say $x_1 \text{---} x_r$

Define S_0 to be the set of these r vectors (which span X)

$$\Rightarrow y_1 + \sum_{j=1}^{r} b_j x_j = 0 \quad \text{for some } b_1, \text{---}, b_r$$

that can't all be zero

$\Rightarrow x_r$ can be written as a linear comb. of $x_1 \text{---} x_{r-1}$ & y_1 (provided $b_r \neq 0$, otherwise one of the x_j's can be written as a linear comb. of the the other x_i's and y_1 – just rename it x_r)

So, $S_1 = \{y_1, x_1, x_2, \text{---}, x_{r-1}\}$ spans the same set as $S_0 \cup \{y_1\}$ but that is X

So S_1 spans X

(e.g. see book)

So one can continue in this fashion building $S_2 = \{y_1, y_2, x_1, \text{---}, x_{r-2}\}$ all the way up to $S_r = \{y_1, \text{---}, y_r\}$ where S_r spans X

But $Q = \{y_1 \text{---} y_{r+1}\}$ is independent (we assumed this)

So $y_{r+1} \notin \text{span}(S_r)$. Contradiction.

Corollary: $\dim \mathbb{R}^n = n$

Proof: $\{e_1, \dots, e_n\}$ spans $\mathbb{R}^n \Rightarrow \dim(\mathbb{R}^n) \leq n$ (by the theorem)

$\{e_1, \dots, e_n\}$ is an indep. set $\Rightarrow \dim(\mathbb{R}^n) \geq n$ (by the def'n of dim)

Theorem: *Let* X be a vector space with $\dim(X) = n$.
Then, any set $E \subset X$ of n vectors spans X if and only if it is independent.

Proof: (I) Proof that E indep & $\dim(X) = n \Rightarrow E$ spans X

Let $E = \{x_1, \dots x_n\}$ be indep., $\dim(X) = n \Rightarrow E \cup \{y\}$ is a dependent set $\forall y \in X. \Rightarrow \forall y \in X, y = \sum_{i=1}^{n} a_i x_i$ for some a_i's

That is, E spans X.

(II) Proof that $\dim(X) = n$ & E span $X \Rightarrow E$ is indep.

This is equivalent to proving E is dep $\Rightarrow E$ does not span X or $\dim(X) \neq n$

Suppose $E = \{x_1, \dots x_n\}$ is dep. $\Rightarrow \exists j : x_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} a_i x_i$ so $E \setminus \{x_j\}$ spans X so $\dim X \leq n-1$

Corollary: If X is a vector space with $\dim(X)=n$, then X has a basis and every basis has n vectors.

Proof: $\dim(X)=n \Rightarrow \exists$ a set $E=\{x_1,\dots,x_n\} \subset X$ of indep. vectors. so by the theorem above E spans X, so it is a basis.

exercise: why does every basis have n vectors?

Exercise: Suppose X is a vector space with $\dim(X)=n$ prove that if $\{y_1 \dots y_r\}$, $1 \le r \le n$, is an indep. set then $\exists$ a basis for X containing $\{y_1 \dots y_r\}$.

Def'n: A mapping $A: X \to Y$ (Y: vector space) is a linear transformation if $\forall x_1, x_2 \in X$, c scalar,

we have $A(x_1+x_2) = Ax_1 + Ax_2$

$A(cx) = c(Ax)$

Remark: If $\{x_1 \dots x_n\}$ is a basis for X, then any $x \in X$ can be written as

$$x = \sum_{i=1}^{n} c_i x_i \Rightarrow Ax = \sum_{i=1}^{n} c_i (Ax_i)$$

Remark: $A: X \to X$, linear, we call it linear operator.

If A is a linear operator on X, one-to-one, onto, we say that A is invertible:

$$\exists A^{-1}: X \to X \quad \text{s.t.} \quad A^{-1}(Ax) = x \quad \forall x \in X.$$

- Then we also have $A(A^{-1}x) = x$ and A^{-1} is linear.

Theorem: Let A be a linear operator on a finite dimension vector space. A is one-to-one iff the range of A is all X.

Proof: Take a basis $\{x_1, \dots, x_n\}$ for X and note that $R(A) = \text{Span}\{Ax_1, \dots, Ax_n\}$ (by linearity) but we've seen that a set of n vectors in X, $\dim(X) = n$ spans if & only if it is independent. So, $R(A) = X$ if & only if $\{Ax_1, \dots, Ax_n\}$ is independent.

Want to show that $R(A) = X$ iff A is one-to-one

Suffs. to show that A is one to one iff $\{Ax_1 \dots Ax_n\}$ is indep.

(I) Let A be one-to-one & suppose $\sum c_i Ax_i = 0$

so $A(\sum c_i x_i) = 0$, so $\sum c_i x_i = 0$ (one-to-one)

so the c_i's are all 0. $\Rightarrow$ Ax_i's are indep.

(II) Suppose $\{Ax_1 \text{---} Ax_n\}$ is independant.

Want to show that $Ax = Ay \Rightarrow x = y$

Suff. to show $A(x-y) = 0 \Rightarrow (x-y) = 0$

Let $x - y = \sum c_i x_i$

suppose $A(\sum c_i x_i) = 0$, so $\sum c_i Ax_i = 0$

so the c_i's are $0 \Rightarrow A(x-y) = 0$ when $x - y = 0$

i.e. when $x = y$.

$\blacksquare$

- The set of all linear transformations is denoted $L(X,Y)$. Instead of $L(X,X)$, we write $L(X)$.

- If $A \in L(X,Y)$, $B \in L(Y,Z)$, we define BA via

$$(BA)(x) = B(A(x))$$

- If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$: $\|A\| = \sup_{|x| \leq 1} |Ax|$

- $|Ax| \leq \|A\| |x|$ (why?)
- If $|Ax| \leq \lambda |x| \Rightarrow \|A\| \leq \lambda$ (why?)

Theorem: If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ then $\|A\| < \infty$ & A is uniformly continuous

Proof: For any $x \in \mathbb{R}^n$ with $|x| \leq 1$, we can write $x = \sum_{i=1}^{n} c_i e_i$ so $|c_i| \leq 1$

$$|Ax| = \left|\sum_{i=1}^{n} c_i Ae_i\right| \leq \sum_{i=1}^{n} |c_i| |Ae_i| \leq \sum_{i=1}^{n} |Ae_i|$$

but $\|A\| = \sup_{|x|\leq 1} |Ax| \Rightarrow \|A\| \leq \sum_{i=1}^{n} |Ae_i| < \infty$ (why?)

Moreover $|A(x-y)| \leq \|A\| |x-y|$ so A is uniformly continuous.

Exercise: 1) Show that $\|A+B\| \leq \|A\| + \|B\|$ & $\|cA\| = |c| \|A\|$

2) Deduce that $L(\mathbb{R}^n, \mathbb{R}^m)$ is a metric space with $d(A,B) = \|A-B\|$.

Exercise: 3) Show that $\|BA\| \leq \|B\|\|A\|$.

Theorem: If $\|B-A\|\,\|A^{-1}\| < 1$ then B is invertible

Proof: It's enough to show that B is one to one

i.e. $Bx = 0 \Rightarrow x = 0$ or $x \neq 0 \Rightarrow Bx \neq 0$

suffices to show $(\alpha-\beta)|x| \leq |Bx|$ for some $(\alpha-\beta) > 0$

- Pick $\alpha = \frac{1}{\|A^{-1}\|}$ & $\beta = \|B-A\|$

So $\beta < \alpha$ & $(\alpha-\beta) > 0$

- Now, $\alpha|x| = \alpha|A^{-1}Ax| \leq \alpha\|A^{-1}\||Ax|$

$= |Ax| \leq |(A-B)x| + |Bx|$

$\leq \beta|x| + |Bx|$

So $\underbrace{(\alpha-\beta)}_{>0}|x| \leq |Bx| \Rightarrow B$ is one-to-one, hence invertible.

Theorem: If A is invertible, the map $A \to A^{-1}$ is continuous

Proof: We had from the previous proof $(\alpha-\beta)|x| \leq |Bx|$

so $(\alpha-\beta)|B^{-1}y| \leq |y|$

$\Rightarrow \|B^{-1}\| \leq (\alpha-\beta)^{-1}$

$$\|B^{-1}-A^{-1}\| = \|B^{-1}(A-B)A^{-1}\| \leq \|B^{-1}\|\,\|A-B\|\,\|A^{-1}\| \leq \frac{\beta}{\alpha(\alpha-\beta)}$$

Note that $\beta \to 0$ & $B \to A$

so we have continuity.

Matrices:

Let $\{x_1, \dots x_n\}$ & $\{y_1, \dots y_m\}$ be bases for X & Y.

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

Then $\exists a_{ij}$ s.t. $Ax_j = \sum_{i=1}^{m} a_{ij} y_i$

$$\begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ | & & | & & | \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

→ call this $[A]$

Suppose $x = \sum_{j=1}^{n} c_j x_j \Rightarrow Ax = \sum_{j=1}^{n} c_j Ax_j$

$$= \sum_{j=1}^{n} c_j \sum_{i=1}^{m} a_{ij} y_i$$

$$= \sum_{i=1}^{m} \Big(\underbrace{\sum_{j=1}^{n} a_{ij} c_j}_{\text{coordinates of } x} \Big) y_i$$

Remark: The matrix $[A]$ does not only depend on X & Y but also on the choice of bases.

Remark: Define $[B][A] := [BA] = [C]$

then $c_{kj} = \sum_i b_{ki} a_{ij}$

(Here $A \in L(X,Y)$, $B \in L(Y,Z) \Rightarrow BA \in L(X,Z)$)

(See book for more)

—×—

Differentiation:

Recall: If $f: (a,b) \to \mathbb{R}$ & $x \in (a,b)$

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

when the limit exists. So for a fixed h:

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

($r(h)$ → remainder)

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

We can think of $f'(x)$ as being in $L(\mathbb{R}^1)$.

It maps $h \mapsto f'(x)h$

Now: Let $\vec{F}: (a,b) \to \mathbb{R}^m$

here $\vec{F}'(x) = \lim_{h \to 0} \dfrac{\vec{F}(x+h) - \vec{F}(x)}{h}$

when the limit exists

As before $\vec{F}(x+h) - \vec{F}(x) = h\vec{F}'(x) + \vec{r}(h)$

$\lim_{h \to 0} \dfrac{\vec{r}(h)}{h} = 0$

Here again $\vec{F}'(x) \in L(\mathbb{R}^1, \mathbb{R}^m)$

via the map $h \to \vec{F}'(x)h$

A more convenient way to think of $\vec{F}'(x)$ is as the linear transformation satisfying

$$\lim_{h \to 0} \frac{|\vec{F}(x+h) - \vec{F}(x) - h\vec{F}'(x)|}{|h|} = 0$$

(provided $\vec{F}'(x)$ exists as defined above)

i.e provided $\vec{F}$ is differentiable)

We are now ready for the general case

Def'n

Let $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$ be a function from the open set E to $\mathbb{R}^m$. Fix $\vec{x} \in E$.

If $\exists A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ s.t.

($\exists h$ small enough, so $x+h \in E$ by openness of E.)

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

Then we say f is differentiable at x and we denote $\vec{f}'(\vec{x}) = A$.

If f is differentiable at every $x \in E$ then we say f is differentiable on E.

Uniqueness: Suppose $\exists A_1, A_2$ s.t.

$$0 = \lim_{h \to 0} \left| \frac{f(x+h) - f(x) - A_1 h}{|h|} \right| = \lim_{h \to 0} \left| \frac{f(x+h) - f(x) - A_1 h}{|h|} \right|$$

Let $B = A_1 - A_2 \Rightarrow |Bh| = |f(x+h) - f(x) - A_1 h - (f(x+h) - f(x) - A_2 h)|$

So $\frac{|Bh|}{|h|} \leq \frac{|f(x+h)-f(x)-A_1h|}{h} + \frac{|f(x+h)-f(x)-A_2h|}{h}$

$\frac{|Bh|}{|h|} \downarrow_{h\to 0} 0$ bec $\frac{|f(x+h)-f(x)-A_1h|}{h} \downarrow_{h\to 0} 0$, $\frac{|f(x+h)-f(x)-A_2h|}{h} \downarrow_{h\to 0} 0$

As this is true for any h, fixing h, we have

$\lim_{t\to 0} \frac{|B(th)|}{|th|} = 0$ but $\frac{|B(th)|}{|th|} = \frac{t|Bh|}{t|h|} = \frac{|Bh|}{|h|}$

So $\frac{|Bh|}{|h|} = 0$ So $Bh = 0$. As this is true for any h, $B = 0 \Rightarrow A_1 = A_2$.

- As before, $f'(x)$ is a function $\in L(\mathbb{R}^n, \mathbb{R}^m)$ and f' is a map from $E \to L(\mathbb{R}^n, \mathbb{R}^m)$
- Since $f(x+h) - f(x) = f'(x)h + r(h)$ with $\frac{|r(h)|}{|h|} \to 0$

we see that if f is differentiable at x, it is cont.'s.

Example: Let $A \in L(\mathbb{R}^n, \mathbb{R}^m)$

$A'(x) = A$ Proof: $A(x+h) - Ax = Ah$
$\Rightarrow |A(x+h) - Ax - Ah| = 0$

Chain Rule:

Theorem (Chain Rule):

Suppose E is an open set in $\mathbb{R}^n$, $f: E \to \mathbb{R}^m$, f is differentiable at $x_0 \in E$, $g: C \to \mathbb{R}^k$ where $f(E) \subset C$, and g is differentiable at $f(x_0)$. Then if $F(x) = g(f(x))$, F is differentiable at x_0 & $F'(x_0) = g'(f(x_0))\, f'(x_0)$

↑ Product of two linear transformations.

Recall the chain rule in 1D:

$$\text{If } F(x) = g(f(x)) \Rightarrow F'(x_0) = g'(f(x_0))\, f'(x_0)$$

Before we prove the general chain rule let us recall

Sketch of proof in 1D:

We seek: $\lim_{h\to 0} \dfrac{g(f(x+h)) - g(f(x))}{h}$

Let $v = \left(\dfrac{f(x+k)-f(x)}{k} - f'(x)\right) \xrightarrow[k\to 0]{} 0$

Let $u = \dfrac{g(y+h)-g(y)}{h} - g'(y) \xrightarrow[h\to 0]{} 0$

Thus: $f(x+k) = f(x) + (v + f'(x))k$ —— (1)

$g(y+h) = g(y) + (u + g'(y))h$ —— (2)

(1) $\Rightarrow g(f(x+k)) = g\Big(\underbrace{f(x)}_{y} + \underbrace{k(v+f'(x))}_{h}\Big)$

$\overset{\text{by (2)}}{=} g(f(x)) + \big(u + g'(f(x))\big)k(v+f'(x))$

But: $h \to 0$ as $k \to 0$

So $\dfrac{g(f(x+k)) - g(f(x))}{k} = \underbrace{(u + g'(f(x)))}_{\downarrow h\to 0,\ 0}\underbrace{(v+f'(x))}_{\downarrow k\to 0,\ h\to 0,\ 0}$

Taking limits: $[g(f(x))]' = g'(f(x))f'(x)$

Proof of Chain Rule

We seek: $\lim_{h\to 0} \dfrac{|g(f(x+h)) - g(f(x))|}{|h|}$

Let $v(k) = \left(f(x+k) - f(x) - f'(x)k\right) \Rightarrow \dfrac{|v(k)|}{|k|} \xrightarrow[k\to 0]{} 0$

Let $u(h) = g(y+h) - g(y) - g'(y)h$

$\Rightarrow \dfrac{|u(h)|}{|h|} \xrightarrow[h\to 0]{} 0$

Thus: $f(x+k) = f(x) + v(k) + f'(x)k$ —— (1)

$g(y+h) = g(y) + u(h) + g'(y)h$ —— (2)

$(1) \Rightarrow g(f(x+k)) = g\Big(\underbrace{f(x)}_{y} + \underbrace{v(k) + f'(x)k}_{h}\Big)$

$\Rightarrow h \to 0$ as $k \to 0$

$\overset{\text{by (2)}}{=} g(f(x)) + u + g'(f(x))\left(v + f'(x)k\right)$

$\Rightarrow$

$g(f(x+k)) - g(f(x)) = u(h) + g'(f(x))\left(v(k) + f'(x)k\right)$

$\Rightarrow g(f(x+k)) - g(f(x)) - g'(f(x))f'(x)k$

$= u(h) + g'(f(x))v(k)$

$$\left|g(F(x+k)) - g(F(x)) - g'(F(x))F'(x)k\right| / |k|$$

$$= \frac{\left|u(h) + g'(F(x))v(k)\right|}{|k|}$$

$$\leq \underbrace{\frac{|u(h)|}{|h|}\frac{|h|}{|k|}}_{\to 0 \text{ as } k \to 0} + \|g'(F(x))\| \underbrace{\frac{|v(k)|}{|k|}}_{\to 0 \text{ as } k \to 0}$$

$$|h| = |v(k) + F'(x)k|$$

$$\leq \|F'(x)\||k| + |v(k)|$$

$$\Rightarrow \frac{|h|}{|k|} \leq \|F'(x)\| + \frac{|v(k)|}{|k|}$$

Taking limits, we are done since $\frac{|u(h)|}{|h|} \to 0$ & $\frac{|v(k)|}{|k|} \to 0$.

Partial Derivative

Let $F: E \subset \mathbb{R}^n \to \mathbb{R}^m$ (E open)

where $\vec{F}(\vec{x}) = \sum_{i=1}^{m} f_i(x) u_i$ (u_i: standard basis for $\mathbb{R}^m$) $= \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$

Let $x \in E$ & define

$$\frac{\partial f_i}{\partial x_j}(\vec{x}) := (D_j f_i)(\vec{x}) := \lim_{t \to 0} \frac{f_i(\vec{x} + t\vec{e}_i) - f_i(\vec{x})}{t}$$

(partial derivative)

(provided the limit exists)

Remark: The existance of all partial derivatives does NOT imply differentiability (See HW 2).

- However:

Theorem: suppose $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$ (E open) is differentiable is differentiable at $x \in E$. Then the partial derivatives $(D_j f_i)(x)$ exist &

$$\vec{f}'(\vec{x})\vec{e}_j = \sum_{i=1}^{m} (D_j f_i)(\vec{x})\vec{u}_i \qquad 1 \le j \le n$$

In words, If f is diff. at x, then the partials exists, and determine $\vec{f}'(\vec{x})$.

Proof: We have by differentiability, for any $\vec{h} \in \mathbb{R}^n$

$$\vec{f}(\vec{x}+\vec{h}) = \vec{f}(\vec{x}) + \vec{f}'(\vec{x})\vec{h} + \vec{r}(\vec{h})$$

with $\dfrac{|\vec{r}(\vec{h})|}{|\vec{h}|} \xrightarrow[\vec{h}\to 0]{} 0$

So with $\vec{h} = t\vec{e}_j$:

$$\vec{f}(\vec{x}+t\vec{e}_j) = \vec{f}(\vec{x}) + t\vec{f}'(\vec{x})\vec{e}_j + \vec{r}(t\vec{e}_j). \text{ So:}$$

$$\lim_{t\to 0} \frac{\vec{f}(\vec{x}+t\vec{e}_j) - \vec{f}(\vec{x})}{t} = \vec{f}'(\vec{x})\vec{e}_j$$

$$\iff \lim_{t\to 0} \sum_{i=1}^{m} \frac{f_i(x+te_j) - f_i(x)}{t} \vec{u}_i = \vec{f}'(x)\vec{e}_j$$

i.e. each summand has a limit & the $D_j f_i$'s exist. ∎

Essentially, this means

$$\left[\vec{f}'(\vec{x})\right] = \begin{bmatrix} (D_1 f_1)(x) & \cdots & (D_n f_1)(x) \\ \vdots & & \\ (D_1 f_m)(x) & \cdots & (D_n f_m)(x) \end{bmatrix}.$$

Special Case: Let γ be a differentiable map from $(a,b) \subset \mathbb{R}^1$ to $E \subset \mathbb{R}^n$, where E is open. Let $f: E \to \mathbb{R}$ be differentiable & defn

$$g(t) = f(\gamma(t)), \quad t \in (a,b)$$

Then by the chain rule, we have $g'(t) = f'(\gamma(t))\,\gamma'(t)$

$g'(t) \in L(\mathbb{R}^1, \mathbb{R}^1)$ ← $f'(\gamma(t)) \in L(\mathbb{R}^n, \mathbb{R}^1)$ $\quad \gamma'(t) \in L(\mathbb{R}^1, \mathbb{R}^n)$

So $g'(t)$ is a scalar.

In matrix notation $[g'(t)] = \underbrace{[f'(\gamma(t))]}_{1\times n}\underbrace{[\gamma'(t)]}_{n\times 1}$

$$[(D_1 f)(\gamma(t)) \mid \cdots \mid (D_n f)(\gamma(t))] \begin{bmatrix} \gamma_1'(t) \\ \vdots \\ \gamma_n'(t) \end{bmatrix}$$

So $$g'(t) = \sum_{i=1}^{n} D_i f(\gamma(t))\,\gamma_i'(t)$$

$$\iff g'(t) = (\nabla f)(\gamma(t)) \cdot \gamma'(t)$$

∇f = vector of partials, $(\nabla f)(x) = \sum_{i=1}^{n} D_i f(x)\, e_i$

Directional Derivatives:

Staying in the same setting, let $\vec{\gamma}(t) = \vec{x} + t\vec{u}$,

where $t \in \mathbb{R}$, $u \in \mathbb{R}^n$ with $|u| = 1$, & $\vec{x} \in E$ is fixed

- Since $g(t) = f(\gamma(t))$, then $g(t) - g(0) = f(\vec{x}+t\vec{u}) - f(\vec{x})$

$$\Rightarrow \lim_{t\to 0} \frac{f(\vec{x}+t\vec{u}) - f(\vec{x})}{t} = g'(0)$$

but $g'(t) = f'(\gamma(t))\gamma'(t) = (\nabla f)(\gamma(t)) \cdot \gamma'(t)$

$\gamma'(t) = \vec{u}$

so $g'(0) = (\nabla f)(\vec{x}) \cdot \vec{u}$

Consequently $\lim_{t\to 0} \frac{f(\vec{x}+t\vec{u}) - f(\vec{x})}{t} = (\nabla f)(\vec{x}) \cdot \vec{u}$

$=: (D_u f)(\vec{x})$

Directional derivative of f in the direction $\vec{u}$

Theorem: Let $\vec{F}: E \subset \mathbb{R}^n \to \mathbb{R}^m$, E is convex & open, $\vec{F}$ is differentiable. Suppose $\|\vec{F}'(\vec{x})\| \leq M$ for some real number M, and for every $x \in E$.

Then: $|\vec{F}(\vec{b}) - \vec{F}(\vec{a})| \leq M|b-a|$ $\forall a, b \in E$

Proof: We will define functions $\gamma: \mathbb{R}^1 \to \mathbb{R}^n$ & $\vec{g}(t) = \vec{F}(\gamma(t))$

So that $\vec{g}'(t) = F'(\gamma(t))\gamma'(t)$

$\Rightarrow$ $|\vec{g}'(t)| = \|F'(\gamma(t))\| |\gamma'(t)|$

$|\vec{g}'(t)|$: want $|\vec{F}(\vec{b}) - \vec{F}(\vec{a})|$

$\|F'(\gamma(t))\| \leq M$

$|\gamma'(t)|$: want $|\vec{b} - \vec{a}|$

Pick $\gamma(t) = (1-t)\vec{a} + t\vec{b}$, so that $\gamma'(t) = \vec{b} - \vec{a}$ & $\vec{g}(0) = F(\vec{a})$, $\vec{g}(1) = \vec{F}(\vec{b})$

$\Rightarrow |\vec{g}'(t)| \leq M|b-a|$ $\forall t \in [0,1]$

(all this is well defined since $\gamma(t) \in E$ by convexity of E)

$\Rightarrow \sup_t |\vec{g}'(t)| \leq M|\vec{b} - \vec{a}|$ ——— (*)

So we'd be done if $|\vec{g}(1) - \vec{g}(0)| \leq (1-0)|g'(t)|$ for some $t \in (0,1)$

Let $\vec{z} = \vec{g}(1) - \vec{g}(0)$ & $\varphi(t) = \vec{z} \cdot \vec{g}(t)$, $0 \le t \le 1$

$\Rightarrow$ by MVT $\varphi(1) - \varphi(0) = (1-0)\varphi'(t)$ for some $t \in (0,1)$

$\varphi(1) = \vec{z}\cdot\vec{g}(1)$, $\varphi(0) = \vec{z}\cdot\vec{g}(0)$, $\varphi'(t) = \vec{z}\cdot\vec{g}'(t)$

so $\vec{z}\cdot\vec{g}(1) - \vec{z}\cdot\vec{g}(0) = (1-0)\vec{z}\cdot\vec{g}'(t)$

so $|\vec{z}|^2 = \vec{z}\cdot\vec{g}'(t) \le |\vec{z}|\,|\vec{g}'(t)|$

$|\vec{g}(1) - \vec{g}(0)| = |\vec{z}| \le |\vec{g}'(t)|$ for some $t \in (0,1)$

so $|\vec{F}(\vec{b}) - \vec{F}(\vec{a})| \le \sup_t |g'(t)| \le M|b-a|$

($\sup_t |g'(t)| \le M$ by (*))

Remark: If in the theorem above $\vec{F}'(\vec{x}) = \vec{0}$ $\forall x \in E$ then $M = 0 \Rightarrow |F(b) - F(a)| \le 0$ $\forall a, b \in E$

so F is constant!

Definition: A differentiable map $\vec{F}$ of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is said to be in $C'(E)$ (or continuously differentiable) if F' is a continuous mapping of E into $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. That is $F \in C'(E)$ if $\forall x \in E$, $\varepsilon > 0$, $\exists \delta > 0$ s.t.

$$\|F'(x) - F'(y)\| \le \varepsilon \quad \forall y \in E \text{ with } |x-y| < \delta$$

Theorem: $f: E \subset \mathbb{R}^n \to \mathbb{R}^m$, E open. $f \in C'(E)$ iff all partials $D_j f_i$ exist and are continuous on E.

Proof: (I) Want to prove $f \in C'(E) \Rightarrow D_j f_i$ are cont.

But $(D_j f_i)(\vec{x}) = (f'(x)\vec{e}_j) \cdot \vec{u}_i \quad \forall \vec{x}$,

(In matrix language, $(D_j f_i)(x)$ is the $(j,i)^{th}$ entry of $[f'(x)]$)

So $$\left|(D_j f_i)(\vec{x}) - (D_j f_i)(\vec{y})\right| = \left|\left((f'(x) - f'(y))\vec{e}_j\right) \cdot \vec{u}_i\right|$$

$$\leq \|f'(x) - f'(y)\| \underbrace{|e_j|}_{1} \underbrace{|u_i|}_{1}$$

$\Rightarrow D_j f_i$ is continuous.

(II) Want to show $D_j f_i$ are cont $\Rightarrow f \in C'(E)$.

We only need to consider the case $m=1$, since $m \geq 1$ follows trivially. (why?)

Let $\vec{x} \in E$ & fix $\varepsilon > 0$.

$\exists$ an open Ball $rB(x) \subset E$ such that

$\forall y \in rB(x)$

$|D_j f(y) - D_j f(x)| < \varepsilon/n$ by the cont. of $D_j f$.

Let $\vec{h} = h_1\vec{e}_1 + \dots + h_n\vec{e}_n$ with $|h| < r$

& define $\vec{v}_0 = 0$

$\vec{v}_k = h_1\vec{e}_1 + \dots h_k\vec{e}_k = \vec{v}_{k-1} + h_k\vec{e}_k$

So that $f(\vec{x}+\vec{h}) - f(\vec{x}) \overset{\text{telescoping}}{=} \sum_{j=1}^{n} \left[f(\underbrace{\vec{x}+\vec{v}_j}_{\in rB(x)}) - f(\underbrace{\vec{x}+\vec{v}_{j-1}}_{\in rB(x)}) \right]$

$= \sum_{j=1}^{n} \underbrace{\left[f(\vec{x}+\vec{v}_{j-1}+\vec{e}_j h_j) - f(\vec{x}+\vec{v}_{j-1}) \right]}_{\text{by MVT: } h_j (D_j f)(x+v_{j-1}+\theta_j h_j \vec{e}_j)}$

for some $\theta_j \in (0,1)$

Subtracting $\sum h_j (D_j f)(x)$ on both sides and using Δ-ineq.

$\Rightarrow |f(\vec{x}+\vec{h}) - f(\vec{x}) - \sum_{j=1}^{n} h_j (D_j f)(\vec{x})|$

$\sum_{j=1}^{n} |h_j| \left| (D_j f)(x+v_j+\theta_j h_j \vec{e}_j) - (D_j f)(\vec{x}) \right|$

$\leq \sum_{j=1}^{n} |h_j| \frac{1}{n}\varepsilon = \left(\sum_{j=1}^{n} |h_j|/n \right) \varepsilon$

$\leq |\vec{h}|\varepsilon$

$\forall \vec{h}$, $|\vec{h}| < r$ $\therefore$ f is diff. & $f'(x) : h \to \sum h_j (D_j f)(x)$

Exercise: Finish the proof. ▨

(i.e. show continuity).