Multidimensional Derivatives

Friday, February 26, 2021 8:16 AM

Last Lecture Review

Recall a linear function $f: \mathbb{R}^n \to \mathbb{R}^p$ is a fcn satisfying

$$f(\vec{x}+\vec{y}) = f(\vec{x}) + f(\vec{y}) \quad \forall\ \vec{x},\vec{y} \in \mathbb{R}^n$$
$$f(a\vec{x}) = af(\vec{x}) \quad a \in \mathbb{R}$$

There's a natural 1-1 correspondence b/tw linear fcs $\mathbb{R}^n$ to $\mathbb{R}^p$ and $p \times n$ matrices w/ real coefficients.

What is correspondence?

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & & \\ \vdots & & \ddots & \\ a_{p1} & & & a_{pn} \end{pmatrix}$$ corresponds to f

if any of the following equivalent conditions are true:

- $f(e_j) = (a_{1j}, a_{2j}, \dots, a_{pj})$ for $j = 1, \dots, n$
- jth column of A is $f(e_j)$ (viewed as a column vector)
- $f(x_1, x_2, \dots, x_n)$ has ith component $\sum_{j=1}^{n} a_{ij} x_j \quad i = 1, \dots, p$

If $f_1, f_2, \dots, f_p$ denote component vectors in $\mathbb{R}^p$, then

$$f\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{p} a_{ij} x_j f_i\right)$$
$$= \sum_{i=1}^{p}\left[\sum_{j=1}^{n} a_{ij} x_j\right] f_i$$

$$f(x_1, \dots, x_n) = A\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Now what if we compose fcns?

Say
$f: \mathbb{R}^n \to \mathbb{R}^p$
$g: \mathbb{R}^p \to \mathbb{R}^q$
$g \circ f: \mathbb{R}^n \to \mathbb{R}^q$
sends $x \in \mathbb{R}^n$ to $g(f(\vec{x})) \in \mathbb{R}^q$

Fact: If f and g are linear, then so is $g \circ f$

e.g. $g(f(\vec{x}+\vec{y})) = g(f(\vec{x}) + f(\vec{y}))$
$= g(f(\vec{x})) + g(f(\vec{y}))$

Say f corresponds to matrix A ($p \times n$ matrix) and g corresponds to B ($q \times p$ matrix)

Q/ Then which $q \times n$ matrix corresponds to $g \circ f$?
A/ Matrix product BA

Suppose C corresponds to $g \circ f$. Then, its jth column is $g(f(e_j))$

Q/ What is $g(f(e_j))$ in terms of A & B?

$$g(f(e_j)) = g(j\text{th column of } A)$$
$$= g(a_{1j}, a_{2j}, \dots, a_{pj})$$ a vector of length p

its ith component is $\sum_{k=1}^{p} b_{ik} a_{kj}$

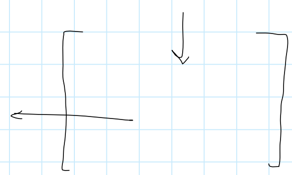
its ith component is $\sum_{k=1}^{p} b_{ik} a_{kj}$

so this is the ij component/coeff of C (def'd as the matrix representing $g \circ f$)

$$\boxed{C = BA}$$

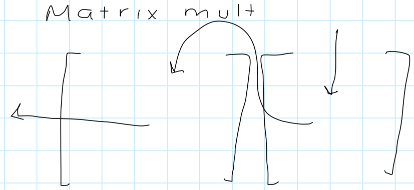
This is matrix multiplication
→ C is the matrix w/columns $BF(e_j)$
⇒ col(C) correspond to col(A)
⇒ row(C) correspond to row(B)
⇒ col(B) & row(A) just get jumbled around

Thinking in terms of input/output:
col(A) correspond to components of input of F and rows of A correspond to components of the output of F
(similar for B and g)



in through the columns, out through the rows

Matrix mult



Limits & Interior

Open ball

$$B(\vec{a}; r) = \left\{ \vec{x} \in \mathbb{R}^n \,\middle|\, \|\vec{x} - \vec{a}\| < r \right\}$$

Closed ball

$$\overline{B}(\vec{a}; r) = \left\{ \vec{x} \in \mathbb{R}^n \,\middle|\, \|\vec{x} - \vec{a}\| \leq r \right\}$$

Let D be a subset of $\mathbb{R}^n$. We say that $\vec{x} \in D$ is an interior of D if

Intuitively: if $\vec{y}$ is near $\vec{x}$, then $\vec{y} \in D$

formally: $\exists\, \delta > 0$ s.t $B(\vec{x}; \delta) \subseteq D$

We say O is open if every point if every point of O is an interior point
ie, an "open subset of $\mathbb{R}^n$"

Examples
- open interval for $n=1$
- union of open intervals ($n=1$)

- union of open intervals ($n=1$)
- all of $\mathbb{R}^n$
- $\emptyset$ the empty set
- open ball $B(a;r)$ (any n) ⟵ inside of a circle
- $n=2$: interior of a square/any polygon
- $n=2$: set of (x,y) satisfying a $\underbrace{\text{strict}}$ linear inequality like:

  less than, rather than less than or equal to

  $x>0$
  $x>-7$
  $y<3$
  $x+y<4$
  $ax+by<c$
- same for linear inequalities for any n

Usually we prefer to consider a fcn def'd on an open domain.
<u>Why?</u>
- If f def'd at $\vec{x}$, then f is def'd <u>near</u> $\vec{x}$ and therefore, we can talk about $\lim_{\vec{y}\to\vec{x}}$ and f will be def'd at $\vec{y}$ near $\vec{x}$

- Another way to state defn of open set:
  D is open if whenever $\vec{x}\in D$, then all points sufficiently close to $\vec{x}$ are also in D.

<u>Nonexample</u>
$D=$ a point not open
and indeed if f is def'd at only a single point, we can't talk about derivatives or limits at that point.

<u>Other non-open sets</u>
- a line in $\mathbb{R}^2$ (or $\mathbb{R}^3$, etc)
- a plane in $\mathbb{R}^3$
- closed ball $\overline{B}(a;r)$ $r\geq 0$
- closed interval
- half-open interval
- square (incl. the boundary) in $\mathbb{R}^2$
- an open square along with a single point on the boundary

<u>Another intuitive def of open</u>
- a set is open if it has no boundary points

Derivatives in Multiple Dimensions

<u>Idea</u> derivative of a fcn $f: D\longrightarrow\mathbb{R}$ with $D\subseteq\mathbb{R}$ (open) at a point $x_0\in D$ is a number $f'(x_0)$
We should think of it as a 1×1 matrix, ie, as a linear fcn from $\mathbb{R}^1$ to $\mathbb{R}^1$ that approximates f near (x_0,y_0) $y_0=f(x_0)$

<u>Technical Note.</u> If L is linear, then $L(0)=0$,
So if we want to translate L to the point (x_0,y_0), we really consider the affine fcn
$$y=L(x-x_0)+y_0$$
$$=L(x)+y_0-L(x_0)$$

<u>So</u> when we say L approximates f near (x_0,y_0) we really mean $L(x-x_0)+y_0$ approximates f.
$(\Rightarrow)$ L itself approximates $f(x+x_0)-y_0$

This applies to $x_0\in\mathbb{R}^n$, $y_0\in\mathbb{R}^p$ and $L:\mathbb{R}^n\to\mathbb{R}^p$
ie, if we translate L to (x_0,y_0), we take $L(\vec{x}-\vec{x_0})+\vec{y_0}$.

<u>Suppose</u> $f: D\longrightarrow\mathbb{R}^p$ where D is an open subset of $\mathbb{R}^n$

For $\vec{x}_0 \in D$, we want to define what we mean by $f'(x_0)$.
It will be a linear fcn from $\mathbb{R}^n$ to $\mathbb{R}^p$.
s.t.

$$\underbrace{f'(x_0)}_{\text{linear fcn}} \underbrace{(\vec{x} - \vec{x}_0)}_{\text{input to linear fcn}} + f(\vec{x}_0)$$

is the best affine approximation to f near $\vec{x}_0$

More precisely we define "best" and what it means for f to be diff'able

f is diff'able at $\vec{x}_0$ iff $f'(\vec{x}_0)$ exists

Even more precisely
Recall precise def for $n = p = 1$

Attempt to generalize to arbitrary p, n

$$f'(\vec{x}_0) = \lim_{\vec{x} \to \vec{x}_0} \frac{f(\vec{x}) - f(\vec{x}_0)}{\vec{x} - \vec{x}_0}$$

PROBLEM
can't divide vectors by other vectors

Instead, the gen def will be:

$$f(\vec{x}) - f(\vec{x}_0) \approx f'(\vec{x}_0)(\vec{x} - \vec{x}_0)$$
for $\vec{x}$ near $\vec{x}_0$

Q/ what do we mean precisely by "$\approx$"?

A/ ~~Look at~~ $f'(x_0) = \lim_{x \to x_0} \frac{1\ f(x) - f(x_0)}{}$

Notice that the eq of $f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$

is equivalent to

$$0 = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)$$

$$= \lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0}$$

Key Fact
a limit approaches 0 iff its magnitude approaches 0.

So, def of derivative is equivalent to:

$$0 = \lim_{x \to x_0} \left| \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} \right|$$

$$= \lim_{x \to x_0} \frac{|f(x) - f(x_0) - f'(x_0)(x - x_0)|}{|x - x_0|}$$

So if we replace x, x_0 with $\vec{x}, \vec{x}_0$ then we get

$$|x - x_0|$$

so if we replace x, x_0 with $\vec{x}, \vec{x}_0$, then we get.

$$0 = \lim_{\vec{x} \to \vec{x}_0} \frac{\| F(\vec{x}) - F(\vec{x}_0) - F'(\vec{x}_0)(\vec{x} - \vec{x}_0) \|}{\| \vec{x} - \vec{x}_0 \|}$$

NO more division by vectors !

use this as def of $F'(\vec{x}_0)$