LECTURE 11: MULTIVARIABLE ANALYSIS

Welcome to the heart of the course, where we generalize single-variable concepts like derivatives to multivariable calculus.

1. Preliminaries

Notation: $x = (x_1, x_2, \dots, x_n)$ is a point in $\mathbb{R}^n$

The **length** of x is denoted by

$$|x| = \sqrt{(x_1)^2 + \cdots + (x_n)^2}$$

We will frequently deal with linear mappings between spaces:

Definition: $A : \mathbb{R}^n \to \mathbb{R}^m$ is a **linear transformation** if:

(1) For all x and y, $A(x+y) = Ax + Ay$

(2) For al x and all scalars c, $A(cx) = cAx$

The set of linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ is denoted as $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$

Mnemonic: $n =$ **IN**put, $m =$ **M**outhput, it goes from $\mathbb{R}^n$ to $\mathbb{R}^m$

Date: Wednesday, July 20, 2022.

2. Norm of A

Given a linear transformation A, we can define $\|A\|$, the norm of A, which like a "maximum spread" of A:

Definition: If $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ then

$$\|A\| = \sup_{x \in \mathbb{R}^n} \frac{|Ax|}{|x|}$$

Note: By definition of sup we have $\|A\| \geq \frac{|Ax|}{|x|}$ for all x and hence

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

This is the form we'll most frequently use.

Intuitively: Think of $\|A\|$ as the largest possible spread of A. For example, if $\|A\| = 2$ then we have $|Ax| \leq 2\,|x|$ for all x, so $|Ax|$ is never more than twice as big as $|x|$.

$\|A\| = \infty$ concretely means there is a sequence x_k of points such that

$$\lim_{k \to \infty} \frac{|Ax_k|}{|x_k|} = \infty$$

So the stretch $\frac{|Ax|}{|x|}$ gets uncontrollably big.

Note: In the definition of $\|A\|$, it's actually enough to assume $|x| \leq 1$ or even just $|x| = 1$ because

$$\frac{|A(cx)|}{|cx|} = \frac{|cAx|}{|c|\,|x|} = \frac{|c|\,|Ax|}{|c|\,|x|} = \frac{|Ax|}{|x|}$$

So if the property holds for x it holds for any multiple cx.

3. Properties of $\|A\|$

Let's now prove some properties of $\|A\|$.

First of all, could $\|A\| = \infty$? Not in finite dimensions!

Fact: If $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ then $\|A\| < \infty$ (that is A is **bounded**)

Proof: Finite-dimensionality is crucial here!

Let $\{e_1, \dots, e_n\}$ be the standard basis of $\mathbb{R}^n$ and let $x \in \mathbb{R}^n$ be given. Then there are $c_1, \dots, c_n$ such that

$$x = c_1 e_1 + \cdots + c_n e_n$$

$$\text{Then } |Ax| = \left| A\left(\sum_{i=1}^{n} c_i e_i\right)\right| \overset{\text{LIN}}{=} \left|\sum_{i=1}^{n} c_i \left(Ae_i\right)\right| \leq \sum_{i=1}^{n} |c_i| \, |Ae_i|$$

However, for each i, we have

$$|c_i| = \sqrt{(c_i)^2} \leq \sqrt{(c_1)^2 + \cdots + (c_n)^2} = |x|$$

$$\text{Hence } |Ax| \leq \sum_{i=1}^{n} \underbrace{|c_i|}_{\leq |x|} |Ae_i| \leq \left(\sum_{i=1}^{n} |Ae_i|\right) |x| = C\,|x|$$

Where $C =: \sum_{i=1}^{n} |Ae_i|$ (doesn't depend on x).

From this it follows that for all x, $\frac{|Ax|}{|x|} \leq C < \infty$ $\square$

One of the most surprising facts about linear transformations is that A bounded implies A is continuous!

Fact: If $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ then A is uniformly continuous

Proof: Let $\epsilon > 0$ be given, let $\delta = \frac{\epsilon}{\|A\|}$ then if $|x - y| < \delta$ then

$$|Ax - Ay| = |A(x - y)| \leq \|A\| \, |x - y| < \|A\| \left(\frac{\epsilon}{\|A\|} \right) = \epsilon \checkmark \quad \Box$$

Next we want to show that $\|A\|$ behaves nicely with respect to sums and compositions

Facts:

(1) $\|A + B\| \leq \|A\| + \|B\|$

(2) $\|AB\| \leq \|A\| \, \|B\|$

Here AB is the composition of A and B whenever it's defined

Why? (1) follows because for all x

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

And (2) follows similarly because for all x

$$|(AB)x| = |A(Bx)| \leq \|A\| \, |Bx| \leq \|A\| \, \|B\| \, |x|$$

Fun fact: If you define $d(A, B) = \|A - B\|$, then $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ actually becomes a metric space! So all the concepts of open sets and continuity makes sense for linear transformations as well. We'll discuss this below.

Relation to matrices: If A is a linear transformation with matrix

$$[A] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

Then by calculating $|Ax|$ and using Cauchy-Schwarz, you can actually show that (see section 9.9 of Rudin for details)

$$\|A\| \leq \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{ij})^2}$$

There are examples where we get a strict inequality. For example, if

$$A(x_1, x_2) = (x_1, 2x_2)$$

Then can show $\|A\| = 2$ but

$$[A] = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

And so the square root of the sum of squares of components is $\sqrt{5}$

4. A^{-1}

Later in this chapter, we will talk a lot about inverse transformations. In order to do this, let's study the inverse transformation A^{-1} a bit more carefully.

Notation: $\mathcal{L}(\mathbb{R}^n) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ (space of linear operators on $\mathbb{R}^n$)

Recall: $A \in \mathcal{L}(\mathbb{R}^n)$ is **invertible** if there is $A^{-1} \in \mathcal{L}(\mathbb{R}^n)$ such that $AA^{-1} = A^{-1}A = I$ (the identity transformation)

Notation: Ω = set of invertible linear transformations on $\mathbb{R}^n$

Goal: Show Ω is open and moreover the mapping $A \to A^{-1}$ is continuous (with respect to the metric $d(A, B) = \|B - A\|$)

Lemma: If $A \in \Omega$ and $B \in \mathcal{L}(\mathbb{R}^n)$ is such that

$$\|B - A\| \left\|A^{-1}\right\| < 1$$

$$\text{Then } B \in \Omega$$

So if B is close enough to A, then B is invertible as well. This makes sense *if* Ω were open, and in fact helps us to show open-ness (see below)

Proof:

STEP 1: Let $\alpha = \frac{1}{\|A^{-1}\|}$ and $\beta = \|B - A\|$

$$\text{Then } \beta = \|B - A\| < \frac{1}{\|A^{-1}\|} = \alpha \Rightarrow \beta < \alpha$$

Then for every $x \in \mathbb{R}^n$, consider:

$$\begin{aligned}
\alpha |x| =& \alpha \left|A^{-1}Ax\right| \leq \alpha \underbrace{\left\|A^{-1}\right\|}_{\frac{1}{\alpha}} |Ax| = |Ax| \\
=& |(A - B)x + Bx| \\
\leq& |(A - B)x| + |Bx| \\
\leq& \|A - B\| \, |x| + |Bx| \\
=& \beta |x| + |Bx|
\end{aligned}$$

Therefore we obtain

$$(\alpha - \beta) |x| \leq |Bx|$$

STEP 2: Claim: B is one-to-one

If $Bx = 0$ then since $\alpha > \beta$ we get

$$\underbrace{(\alpha - \beta)}_{>0} |x| \leq |Bx| = |0| = 0$$

Which in turn implies $|x| = 0$ and so $x = 0$ ✓

Since $B : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one, it follows that B is also onto and hence B is invertible, that is $B \in \Omega$ □

Corollary: Ω is open

Suppose $A \in \Omega$ and let $r = \frac{1}{\|A^{-1}\|} > 0$ then the previous result shows that if $d(B, A) < r$ then $B \in \Omega$, so Ω is open. □

Fact: The mapping $A \to A^{-1}$ is continuous

STEP 1: We need to study $B^{-1} - A^{-1}$ for B "close" to A

Claim # 1:

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

Why?

$$\begin{aligned} B^{-1}(A - B)A^{-1} =& B^{-1}\left(AA^{-1} - BA^{-1}\right) = B^{-1}\left(I - BA^{-1}\right) = B^{-1} - B^{-1}BA^{-1} \\ =& B^{-1} - A^{-1} \checkmark \end{aligned}$$

From the claim, it follows that

$$\left\|B^{-1} - A^{-1}\right\| \leq \left\|B^{-1}\right\| \underbrace{\|A - B\|}_{\beta} \underbrace{\left\|A^{-1}\right\|}_{\frac{1}{\alpha}} = \left\|B^{-1}\right\| \left(\frac{\beta}{\alpha}\right)$$

So all is left is to study $\left\|B^{-1}\right\|$

STEP 2:

Claim # 2:

$$\left\|B^{-1}\right\| \leq \frac{1}{\alpha - \beta}$$

Why? In the Lemma, we showed that for all x, we have

$$(\alpha - \beta)\,|x| \leq |Bx|$$

Replacing x with $B^{-1}x$ in the above, we get

$$(\alpha - \beta)\left|B^{-1}x\right| \leq \left|BB^{-1}x\right| \Rightarrow (\alpha - \beta)\left|B^{-1}x\right| < |x|$$

And therefore $\frac{\left|B^{-1}x\right|}{|x|} < \frac{1}{\alpha-\beta}$ and hence $\left\|B^{-1}\right\| \leq \frac{1}{\alpha-\beta}$

STEP 3: Combining steps 1 and 2 we get

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

Let $\epsilon > 0$ be given, then since

$$\lim_{\beta \to 0} \frac{\beta}{(\alpha - \beta)\alpha} = 0$$

There is $\delta > 0$ such that if $|\beta| < \delta$ then $\left|\frac{\beta}{(\alpha-\beta)\alpha}\right| < \epsilon$.

With that δ, if $\underbrace{\|A - B\|}_{\beta} < \delta$ then by the above, $\left\|B^{-1} - A^{-1}\right\| < \epsilon$. $\square$

5. The Derivative in $\mathbb{R}^n$

With those preliminaries out of our way, we can finally embark on our exploration of derivatives in $\mathbb{R}^n$.

Goal: If $f : \mathbb{R}^n \to \mathbb{R}^m$, how to define the derivative $f'(x)$?

First guess: By analogy with the scalar case, if $x \in \mathbb{R}^n$

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

But here h is a vector, so it makes no sense to divide by h

So we need a definition of derivatives that doesn't have division in it.

What saves us is the concept of linear approximation from calculus:

Analogy: $(n = 1)$ Note that if h is small, then

$$f(x+h) = f(x) + f'(x)h + \text{ Smaller terms}$$

(This was used to approximate quantities like $\sqrt{4.02}$ or $\ln(0.97)$)

This is the point of view that we'll take:

Important definition: Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ and $x \in \mathbb{R}^n$.

If there is a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$f(x+h) = f(x) + Ah + r(h)$$

$$\text{Where } \lim_{h\to 0} \frac{|r(h)|}{|h|} = 0$$

Then we say f is **differentiable at** x and $f'(x) = A$

Definition: f is **differentiable** if f is differentiable at all x

In other words, if you can expand $f(x+h)$ out with a small remainder, then the linear part is the derivative of f.

Before, $f'(x)$ was just a number, but now it's something more dynamic, it's a linear transformation. Intuitively, if f distorts space, then $f'(x)$ describes the linear part of the distortion.

Note: More commonly, people write $o(h)$ instead of $r(h)$, it's a term smaller than h