MAT257: Real Analysis II

QiLin Xue

Fall 2021

Contents

1 Review: Inner Product, Cauchy–Schwarz, Triangle Inequality

Lectures 1 and 2 (Sept 10/13)

- Everything in the course will lead to Stokes' Theorem:

$$\int_C \mathrm{d}W = \int_{\partial C} W \tag{1}$$

This generalizes a well-known theorem in one-dimensional calculus, known as the Fundamental Theorem of Calculus:

$$\int_{[a,b]} F'(t) = F(b) - F(a) = \int_{\partial[a,b]} F \tag{2}$$

where $\partial[a,b] = \{\underbrace{b}_{+}, \underbrace{a}_{-}\}$.

- **Continuity in $\mathbb{R}^n$**: Recall that continuity in $\mathbb{R}$ is formally defined via $\delta - \epsilon$. However intuitively it means that if you wiggle the input by a tiny bit, you wiggle the output by a tiny bit.

A similar way can be used to view continuity in $\mathbb{R}^n$.

Definition: For $x, y \in \mathbb{R}^n$, **the standard (Euclidean) inner product** of x and y denoted

$$\langle x, y\rangle = \sum_{i=1}^n x_i y_i \tag{3}$$

The **norm squared** is defined as

$$|x|^2 = \langle x, x\rangle \tag{4}$$

and the **norm** of x is defined:

$$|x| = \sqrt{|x^2|} = \sqrt{\sum_{i=1}^n x_i^2} \tag{5}$$

Idea: There are multiple ways of defining $\mathbb{R}^n$. Some people will define it as the set of all column vectors while others define it as the set of all row vectors. In linear algebra, the distinction is important but in real analysis, this distinction is not too important.

- A **bilinear** function $f(x, y)$ means that the function is linear in each of the two variables. This means that

$$f(ax + by, z) = af(x, z) + bf(y, z) \tag{6}$$

and similarly the same thing for the other parameter.

- A **semi-linear** function $f(x)$ is one such that

$$f(ax) = |a|f(x) \tag{7}$$

Proposition 1: If $x, y, z \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$, then:

0. $\langle \cdot, \cdot\rangle$ is bilinear and $|\cdot|$ is semi-linear. Also note that $\langle x, y\rangle = \langle y, x\rangle$.
1. $|x| \geq 0$ and $|x| = 0 \iff x = 0$.
2. Cauchy–Schwarz Inequality: $|\langle x, y\rangle| \leq |x||y|$ and equality holds if and only if x, y are dependent.
3. Triangle Inequality: $|x + y| \leq |x| + |y|$
4. Polarization Identity: $\langle x, y\rangle = \dfrac{|x + y|^2 - |x - y|^2}{4}$

Proof. We prove each part separately, and skip 0:

1. We have $|x| = \sqrt{\sum x_i^2} \geq 0$ since every x_i^2 is non-negative. Then $|x| = 0$ if and only if $x_i = 0$, i.e. $x = 0$.

2. Consider and note that $\left||y|^2 x - \langle x, y\rangle y\right|^2 \geq 0$. This is equal to 0 if and only if the first term (a multiple of x) equals the second term (a multiple of y), which is equivalent to x, y being dependent.

   Next, note that

$$
\begin{aligned}
|s+t|^2 &= \langle s+t, s+t\rangle &(8)\\
&= \langle s,s\rangle + \langle s,t\rangle + \langle t,s\rangle + \langle t,t\rangle &(9)\\
&= |s|^2 + 2\langle s,t\rangle + |t|^2 &(10)
\end{aligned}
$$

   Using this result, we can simplify the earlier expression to get that

$$
\begin{aligned}
\left||y|^2 x - \langle x,y\rangle y\right|^2 &= |y|^4|x|^2 + \langle x,y\rangle^2|y|^2 - 2|y|^2\langle x,y\rangle^2 &(11)\\
&= |y|^2(|y|^2|x|^2 - \langle x,y\rangle^2) &(12)
\end{aligned}
$$

   Since this quantity is non-negative, it follows that $|y|^2|x|^2 \geq \langle x,y\rangle^2$, which is what we wanted to show.

   Note that there is another part regarding equality, which will not be proven in these notes.

3. As both $|x+y|$ and $|x|+|y|$ are nonnegative, we can square both sides. It suffices to prove that

$$
\begin{aligned}
|x+y|^2 &\overset{?}{\leq} (|x|+|y|)^2 &(13)\\
\langle x+y, x+y\rangle &\overset{?}{\leq} |x|^2 + |y|^2 + 2|x||y| &(14)\\
\langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle &\overset{?}{\leq} |x|^2 + |y|^2 + 2|x||y| &(15)\\
|x|^2 + 2\langle x,y\rangle + |y|^2 &\overset{?}{\leq} |x|^2 + |y|^2 + 2|x||y| &(16)\\
\langle x,y\rangle &\overset{?}{\leq} |x|\cdot|y| &(17)
\end{aligned}
$$

   which is true via Cauchy–Schwarz.

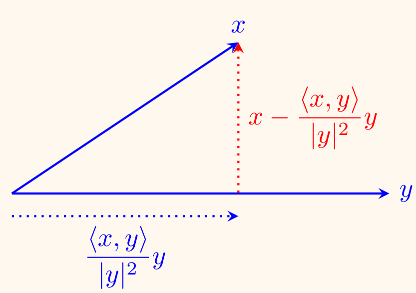
4. The proof is trivial. Expanding everything on the right hand side leads to the left hand side.

   Note: This property is important as it tells us that if we know how to compute norms, we can compute inner products.

□

Idea: In the proof for 2, a weird quantity $\left||y|^2 x - \langle x,y\rangle y\right|^2$ was introduced. There is actually an intuition behind it. Recall that a geometric interpretation of Cauchy-Shwarz in $\mathbb{R}^2$ can be given as follows: $\langle x,y\rangle = |x||y|\cos\theta$ where θ is the "angle" between the two vectors x, y. This is smaller than $|x||y|$ since $\cos\theta \leq 1$.

Similarly, we want to find a generalized way to express something equivalent to $\cos\theta \leq 1$ in $\mathbb{R}^n$. One idea to show whether two vectors are dependent or not is to look at the separation between them, i.e. the length of the red line below.



Therefore, we have

$$\left|x - \frac{\langle x,y\rangle}{|y|^2}y\right| \geq 0 \tag{18}$$

Removing the fraction gives

$$\left||y|^2 x - \langle x,y\rangle y\right| \geq 0. \tag{19}$$

2 Review: Continuity, Distance, and Linear Algebra Review

Lecture 2 and 3 (Sept 13 and 15) + Tutorial 1

- The properties in the previous section will be important when giving a formal definition of continuity.

Definition: If $x, y \in \mathbb{R}^n$, then:

$$d(x, y) = \text{“Distance between } x \text{ and } y\text{”} = |x - y| \tag{20}$$

Theorem:
1. d is **symmetric**, i.e. $d(x, y) = d(y, x)$.
2. d is **positive definite**, i.e. $d(x, y) \geq 0$ and $d(x, y) = 0 \iff x = y$.
3. Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

Proof. We prove each separately:

1. $d(x, y) = |x - y| = |-(y - x)| = |-1| \cdot |y - x| = |y - x| = d(y, x)$
2. $d(x, y) = 0 \iff |x - y| = 0 \iff x - y = 0 \iff x = y$
3. We need to check that

$$d(x, z) \overset{?}{\leq} d(x, y) + d(y, z) \tag{21}$$
$$|x - z| \overset{?}{\leq} |x - y| + |y - z| \tag{22}$$
$$|x - z| \overset{?}{\leq} |x - z| \tag{23}$$

where the third line comes from the previous triangle inequality. The last statement is true, and the steps are reversible, so we are done.

□

- This theorem is significant as these are the only properties that we need to know about distances to formally define continuity.

  Note: In a future section, we will use these properties to *define* a distance function (formally a metric), which is anything that satisfies these properties. This will allow us to generalize continuity to more abstract spaces.

- A note on notation. Our definition of a norm is known as the L^2 or *Euclidean norm,* i.e.

$$|x|_{L^2} = \sqrt{\sum x_i^2} \tag{24}$$

The L^1 norm can be defined as

$$|x|_{L^1} = \sum |x_i| \tag{25}$$

and the infinity norm:

$$|x|_\infty = \text{sup}|x_i|. \tag{26}$$

which is the maximum coordinate. Note that for a finite space, we can use max, but the use of supreme allows us to generalize to infinite spaces.

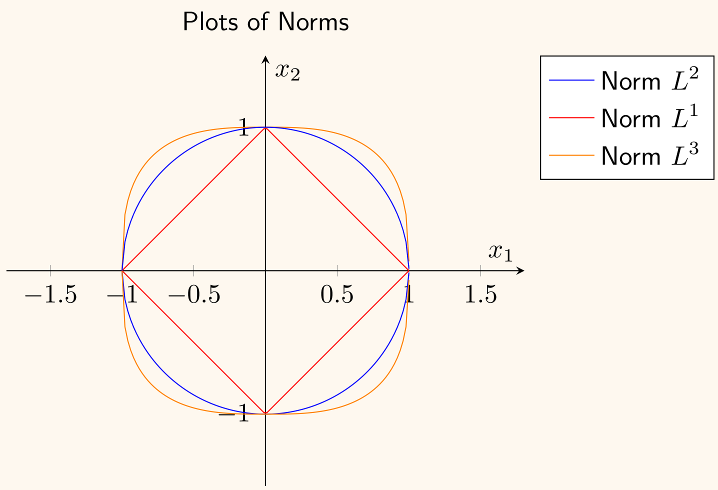
Idea: The p-norm can be defined as

$$|x|_{L^p} = \left(\sum x_i^p\right)^{1/p} \tag{27}$$

And it turns out that

$$|x|_\infty = \lim_{p \to \infty} |x|_{L^p}. \tag{28}$$

There is a nice geometric idea behind why this is the case. Consider $\mathbb{R}^2$. Then we can plot vectors with a norm of 1 in L^1, L^2, L^3 where the coordinate axes are x_1 and x_2.



We can see that as the dimension of the norm increases, the plot becomes closer and closer to a square, which can be represented by $\max\{x_1, x_2\}$.

Example 1: Suppose we have $\|\cdot\|_a$ on $\mathbb{R}^n$ and a norm $\|\cdot\|_b$ on $\mathbb{R}^n$. Suppose $\exists c > 0$ such that $\|x\|_b \le c\|x\|_a$. Then if $f : \mathbb{R}^n \to \mathbb{R}$ is continuous with respect to the b-norm, then it is also continuous with respect to the a-norm.

Proof. We know that for any $\epsilon > 0$, we have $\|x - y\|_b < \delta_b \implies |f(x) - f(y)| < \epsilon$. Let us pick $\delta = \dfrac{\delta_b}{c}$. Then we have:

$$\begin{aligned} \|x - y\|_a < \frac{\delta_b}{c} &\implies c\|x - y\|_a < \delta_b && (29)\\ &\implies \|x - y\|_b \le c\|x - y\|_a < \delta_b && (30)\\ &\implies |f(x) - f(y)| < \epsilon && (31) \end{aligned}$$

□

- Similarly, distances for L^1 and the infinity norms can be defined as $d_1(x, y) = |x - y|_1$ and $d_\infty(x, y) = |x - y|_\infty$.

  Exercise: Show that d_1 and d_∞ also satisfy the properties of a distance.

- There is a bijection from the linear map $\mathbb{R}^n \to \mathbb{R}^m$ and a $m \times n$ matrix:

$$\{T : \mathbb{R}^n \to \mathbb{R}^m\} \longleftrightarrow M_{m\times n}(\mathbb{R}) \tag{32}$$

  which is also a homomorphism. We can associate a matrix with any linear transformation, and any linear transformation is associated with a matrix. Here, the standard basis is used.

- Specifically, we have the map:

$$A \in M_{m\times n} \mapsto L_A(x) = Ax \tag{33}$$

  where $x \in \mathbb{R}^n$, and

$$T \mapsto M_T = \begin{pmatrix} Te_1 & Te_2 & \cdots & Te_n \end{pmatrix} \tag{34}$$

- We also need to show that this map is bijective, i.e both

$$L_{M_T} = T, \quad M_{L_A} = A \tag{35}$$

  are both satisfied.

- Furthermore, we can also show that this map is a homomorphism. Note that the set of linear transformations is itself a vector space. Both $A \mapsto L_A$ and $T \mapsto M_T$ is linear, so

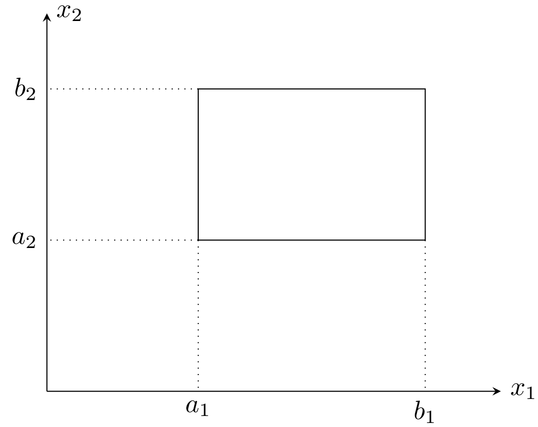
$$\begin{aligned} L_{aA+bB} &= aL_A + bL_b && (36)\\ M_{aT+bS} &= aM_T + bM_s. && (37) \end{aligned}$$

- Furthermore, suppose we have two maps $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$. To go from $\mathbb{R}^n$ to $\mathbb{R}^p$, we can take the composition
$$S \circ T \tag{38}$$
or
$$S \mathbin{/\!\!/} T \tag{39}$$
The bijection is a homomorphism, so we have
$$M_S M_T = M_{S \circ T} \tag{40}$$

3 Intervals and Regions in Higher Dimensions

Lecture 3 (Sept 15)

- In single-variable calculus, we typically focused on an interval $[a, b]$ on the real number line. Similarly, we can talk about intervals in $\mathbb{R}^2$ which can be represented as a rectangle:



- We can generalize to $\mathbb{R}^n$. Given $a_i \leq b_i$ for $i = 1, \ldots, n$, we can define the **closed rectangle** corresponding to a_i, b_i:

$$R = \prod_{i=1}^{n} [a_i, b_i] = \{x \in \mathbb{R}^n : \forall i \, a_i \leq x_i \leq b_i\} \tag{41}$$

- Recall from set theory, if X and Y are sets, then $X \times Y = \{(x, y) : x \in X, y \in Y\}$ (also referred to as direct product in group theory). This is associative (up to isomorphism), so

$$(X \times Y) \times Z \neq X \times (Y \times Z) \tag{42}$$

  but

$$(X \times Y) \times Z \cong X \times (Y \times Z). \tag{43}$$

  As a result, while they are not equal strictly speaking, we can view them as the same.

- We can then view $\mathbb{R}^n$ as

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1 \, \ldots \, x_n) : x_i \in \mathbb{R}\} \tag{44}$$

- Products of intervals can be written as

$$\prod_{i=1}^{n} [a_i, b_i] = \{(x_1, \ldots, x_n) : \forall i \, x_i \in [a_i, b_i]\} \tag{45}$$

- Likewise, there are also **open rectangles.** Specifically, the open rectangle defined by a_i, b_i is

$$\prod_{i=1}^{n} (a_i, b_i) = \{(x_1, \ldots, x_n) \forall i \, x_i \in (a_i, b_i)\} \tag{46}$$

- There is a way to define continuity using open sets.

> **Definition**: The subset $A \subset \mathbb{R}^n$ is called "open" if:
>
> For every $a \in A$, there exists an open rectangle R, such that $x \in R \subset A$.