33A

Discussion Worksheet 5

1. *Proof.* We'll start with the preliminary questions – namely, why do these definitions make any sense? These are infinite sums, so we should be concerned about convergence here. Luckily, the very definition of ℓ_2 will save us.

   Let's start with well definition of the norm $||\vec{x}|| = \sqrt{x_1^2 + x_2^2 + \dots}$. Well the term inside the square root is $x_1^2 + x_2^2 + \dots$, and by definition of ℓ_2 this converges. Furthermore, this sum is certainly nonnegative as each v_i^2 is nonnegative. Hence, $||\vec{x}||$ is a well defined nonnegative real number.

   Now how about this infinite dimensional dot product? Why does this sum converge? Well let's get our hands dirty with some partial sums. Actually, to get a well behaved theory we will actually want to show absolute convergence of the sum. Indeed, consider

$$s_N = |x_1 y_1| + \dots + |x_N y_N|$$

   We want to show that this limit exists. It suffices to bound it above, as s_N is an increasing sequence. Indeed, we use the Cauchy – Schwarz inequality to get

$$|x_1||y_1| + \dots + |x_N||y_N| \leq \sqrt{x_1^2 + \dots + x_N^2}\sqrt{y_1^2 + \dots + y_N^2}$$

   Now, the right side here is like a truncation of the expression for $||\vec{x}||||\vec{y}||$. Indeed, we have

$$(x_1^2 + \dots + x_N^2)(y_1^2 + \dots + y_N^2) \leq (x_1^2 + \dots + x_N^2 + \dots)(y_1^2 + \dots + y_N^2 + \dots) = ||\vec{x}||^2||\vec{y}||^2$$

   Thus, we have the chain of inequalities

$$|x_1||y_1| + \dots + |x_N||y_N| \leq \sqrt{x_1^2 + \dots + x_N^2}\sqrt{y_1^2 + \dots + y_N^2} \leq ||\vec{x}||||\vec{y}||$$

   and we showed above that this infinite dimensional norm was well defined, so we have shown that s_N is bounded above. Hence, as it's increasing in N, it converges. So the series for $\vec{x} \cdot \vec{y}$ is absolutely convergent.

   (a) As suggested, we proceed via geometric series. Indeed, $x_i = 2^{-(i-1)}$. Notice that the next term in the sequence for $\vec{x}$ is always half the preceding term. Now, we are interested in

$$||\vec{x}|| = \sqrt{x_1^2 + x_2^2 + \dots}$$

   So let's first consider the part inside the square root.

$$\begin{aligned} x_1^2 + x_2^2 + \dots &= \sum_{i=1}^{\infty} 2^{-2(i-1)} \\ &= \sum_{i=1}^{\infty} (1/4)^{-(i-1)} \end{aligned}$$

This is a geometric series with first term 1 and common ration 1/4. As $|1/4| < 1$, this converges. And in fact, we may compute its value as

$$\sum_{i=1}^{\infty}(1/4)^{-(i-1)} = \frac{1}{1-\frac{1}{4}} = \frac{4}{3}$$

so we conclude that

$$||\vec{x}|| = \sqrt{4/3} = \frac{2}{\sqrt{3}}$$

(b) What's an angle between these infinite sequences?????

This is in some sense defined by fiat. Recall that for nice normal finite dimensional vectors like

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \\ w_n \end{pmatrix}$$

we have the formula

$$\vec{v} \cdot \vec{w} = ||\vec{v}||||\vec{w}|| \cos(\theta)$$

where θ is the angle between $\vec{v}$ and $\vec{w}$ (taken so that θ is between 0 and π). So what should the "angle" between infinite dimensional vectors be? Well three out of the four terms in the above formula make sense in infinite dimensions. We know from above what the norm and dot product of these infinite dimensional vectors in ℓ_2. So we will *define* the angle between these infinite dimensional vectors to be the θ which makes the above formula work. That is, we take

$$\theta = \arccos\left(\frac{\vec{x} \cdot \vec{y}}{||\vec{x}||||\vec{y}||}\right)$$

which we say takes values in $[0, \pi]$.

Anyways, now we're onto the problem itself. Let's let $\vec{x} = (1, 1/2, 1/4, 1/8, \dots)$ and $\vec{y} = (1, 0, 0, \dots)$. We computed above that $||\vec{x}|| = 2/\sqrt{3}$. Furthermore, $||\vec{y}|| = \sqrt{1^2 + 0^2 + 0^2 + \dots} = 1$. We must also compute $\vec{x} \cdot \vec{y}$. Indeed, this is

$$\begin{aligned} \vec{x} \cdot \vec{y} &= x_1 y_1 + x_2 y_2 + \dots \\ &= 1 * 1 + (1/2) * 0 + (1/4) * 0 + \dots \\ &= 1 \end{aligned}$$

Notice the similarity between $\vec{y}$ and our usual standard basis vector $\vec{e_1}$.

Anyways, we now know all the terms in our formula for the angle between $\vec{x}$ and $\vec{y}$.

$$\begin{aligned} \frac{\vec{x} \cdot \vec{y}}{||\vec{x}||||\vec{y}||} &= \frac{1}{(2/\sqrt{3}) * 1} \\ &= \frac{\sqrt{3}}{2} \end{aligned}$$

And we know that $\arccos(\sqrt{3}/2) = \pi/6$. Hence, the angle between $\vec{x}$ and $\vec{y}$ is $\pi/6$.

(c) Note that if $x_n \to 0$ as $n \to \infty$ then we also have $x_n^2 \to 0$ as $n \to \infty$. So if we find such an example, we will end up with a sequence x_n^2 which tends to 0 as $n \to \infty$ but whose sum diverges. There are many examples of these, and one common one is the harmonic series

$$1, 1/2, 1/3, \ldots, 1/n, \ldots$$

Certainly, its terms tend to 0 as $n \to \infty$. But its sum diverges! That is,

$$1 + 1/2 + 1/3 + \cdots + 1/n + \cdots = \infty$$

I won't prove this here, as it's really just calculus, but you can certainly find arguments online if you look up "divergence of the harmonic series" or something. Anyways, with that we are led to take $x_n^2 = 1/n$. So we can set $x_n = 1/\sqrt{n}$. This will be our desired example.

(d) Once again, the geometry here is confusing. Projecting onto a subspace in an infinite dimensional space is a very abstract notion. But just like we did with the angles in (b), the geometric ideas will be defined by extending what we know from the nice and safe finite dimensional things we've done so far.

If we have a line L spanned by a vector $\vec{v}$ in $\mathbb{R}^n$, what is the orthogonal projection of $\vec{w}$ onto L? We previously derived a formula

$$proj_L(\vec{w}) = \frac{\vec{v} \cdot \vec{w}}{\vec{v} \cdot \vec{v}} \vec{v}$$

Now, the way we derived this formula was through geometric arguments. Basically, we drew some pictures and reasoned from there. But we certainly don't have access to pictures for these infinite dimensional spaces. However, we have a definition of a dot product in our infinite dimensional space ℓ_2. As such, we can just copy the same formula. This may seem hacky, but it's actually quite meaningful. The projection in the infinite dimensional case will satsify many of the properties we'd want. For instance, its image should be L and its kernel should be the orthogonal complement of L. And in fact, $proj_L(\vec{w})$ will end up being the closest point in L to $\vec{w}$. The proofs of these properties will be nearly identicaly to the case of $\mathbb{R}^n$. This is a deep notion – defining a dot product allows us to define a sensible notion of geometry on immense spaces like these!

Anyways, onto the computation. Let's let $\vec{x} = (1, 1/2, 1/4, 1/8, \dots)$ and $\vec{y} = (1, 0, 0, \dots)$ as in (b). Then the projection we seek to compute is

$$proj_L(\vec{y}) = \frac{\vec{x} \cdot \vec{y}}{\vec{x} \cdot \vec{x}} \vec{x}$$

We computed the norm of $\vec{x}$ in (a) as $||\vec{x}|| = 2/\sqrt{3}$. Hence, $\vec{x} \cdot \vec{x} = ||\vec{x}||^2 = 4/3$. And we computed the dot product of $\vec{x}$ and $\vec{y}$ in (b) as $\vec{x} \cdot \vec{y} = 1$. We can therefore conclude that

$$\begin{aligned} proj_L(\vec{y}) &= \frac{\vec{x} \cdot \vec{y}}{\vec{x} \cdot \vec{x}} \vec{x} \\ &= \frac{1}{(4/3)} \vec{x} \\ &= \frac{3}{4} \vec{x} \\ &= \left(\frac{3}{4}, \frac{3}{8}, \frac{3}{16}, \dots \right) \end{aligned}$$

So we have computed the projection. A formula for the i^{th} term of this is $3 * 2^{-(i-3)}$.

$\square$

2. *Proof.* First of all, square roots suck and the length of a vector has a square root. Let's fix that. Indeed, we know that if a and b are nonnegative then $a \leq b$ if and only if $a^2 \leq b^2$. Essentially, this is because $f(x) = x^2$ and $g(x) = \sqrt{x}$ are both increasing functions, but g is only defined on $[0, \infty)$. So to prove our "triangle inequality" it suffices to prove instead that

$$||\vec{v} + \vec{w}||^2 \leq (||\vec{v}|| + ||\vec{w}||)^2$$

Now, we get the square of a length on the left hand side, which we can understand using the dot product.

$$||\vec{v} + \vec{w}||^2 = (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w})$$

And now we're in luck, as the dot product distributes over addition in a very similar fashion to usual multiplication. Indeed, we have

$$\begin{aligned} (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) &= \vec{v} \cdot (\vec{v} + \vec{w}) + \vec{w} \cdot (\vec{v} + \vec{w}) \\ &= \vec{v} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \vec{w} \cdot (\vec{v} + \vec{w}) \\ &= \vec{v} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} + \vec{w} \cdot \vec{w} \end{aligned}$$

By the way, this is remarkably similar to the "FOIL" thing from, say, precalculus. That't not an accident – both of them arise from the "distributivity" over addition. Another fancy term is that both normal multiplication and the dot product are "bilinear".

Anyways, let's also recall that $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$. We can rewrite the above as

$$\vec{v} \cdot \vec{v} + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} + \vec{w} \cdot \vec{w} = \vec{v} \cdot \vec{v} + 2\vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{w}$$

Recall again that dotting a vector with itself yields the square of its length, so this becomes

$$\vec{v} \cdot \vec{v} + 2\vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{w} = ||v||^2 + 2\vec{v} \cdot \vec{w} + ||w||^2$$

Putting this together, we've shown

$$||\vec{v} + \vec{w}||^2 = ||v||^2 + 2\vec{v} \cdot \vec{w} + ||w||^2$$

There really isn't anything else we can do to simplify this, so now let's consider the other side of the inequality.

$$(||\vec{v}|| + ||\vec{w}||)^2 = ||\vec{v}||^2 + 2||\vec{v}||||\vec{w}|| + ||\vec{w}||^2$$

Notice how many terms we have in common from before! Remember that our goal is to show that

$$||\vec{v} + \vec{w}||^2 \leq (||\vec{v}|| + ||\vec{w}||)^2$$

We can substitute both sides of this with what we computed above. So we want to show

$$||v||^2 + 2\vec{v} \cdot \vec{w} + ||w||^2 \leq ||\vec{v}||^2 + 2||\vec{v}||||\vec{w}|| + ||\vec{w}||^2$$

Only the middle terms here are different. Hence, we will be done if we can show that

$$\vec{v} \cdot \vec{w} \leq ||\vec{v}||||\vec{w}||$$

And we're in luck, as this is exactly what the Cauchy – Schwarz inequality says!

To summarize this argument, we have

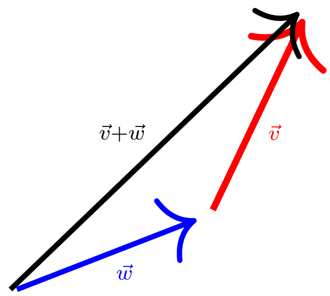
$$\begin{aligned} ||\vec{v} + \vec{w}||^2 &= (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) \\ &= ||\vec{v}||^2 + 2\vec{v} \cdot \vec{w} + ||\vec{w}||^2 \\ (*) &\leq ||\vec{v}||^2 + 2||\vec{v}||||\vec{w}|| + ||\vec{w}||^2 \\ &= (||\vec{v}|| + ||\vec{w}||)^2 \end{aligned}$$

with $(*)$ coming from Cauchy – Schwarz. Now take square roots on both sides to conclude the triangle inequality

$$||\vec{v} + \vec{w}|| \leq ||\vec{v} + \vec{w}||$$

The moral of the story is to approach problems like this by exploring all the different ways we can rewrite the expressions in question. We rewrote our lengths with dot products, and rewrote our dot products using distributivity. We finally got stuck at showing $\vec{v} \cdot \vec{w} \leq ||\vec{v}||||\vec{w}||$, so we had to appeal to a known result called Cauchy – Schwartz. We often call this exploratory approach "following your nose".

Anyways, here's the picture.



Here', I'm drawing $\vec{v}$ as starting from the head of $\vec{w}$. This allows us to clearly see where $v \vec{+} w$ lies. Now, this is a triangle, which was in the name of the inequality we're proving. Indeed, this said that

$$||v \vec{+} w|| \leq ||\vec{v}|| + ||\vec{w}||$$

In words, the length of $\vec{v} + \vec{w}$ is less than or equal to the length of $\vec{v}$ plus the length of $\vec{w}$. How about on our picture? That says that the black length is at most the sum of the red and blue lengths. In other words, if I was trying to get from the bottom left of the triangle to the top right, the black path is shorter than the blue–then–red path. This makes sense, the fastest way to get from one point to another is to walk in a straight line.

Here's some food for thought: can you think of conditions of $\vec{v}$ and $\vec{w}$ so that we we attain equality? By this, I mean so that $||\vec{v} + \vec{w}|| = ||\vec{v}|| + ||\vec{w}||$ rather than just the inequality. Think geometrically first!

As a hint, this works for $\vec{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\vec{w} = \begin{pmatrix} 2357 \\ 0 \end{pmatrix}$.

Can you prove that your conditions work? As a hint, look back at the proof we wrote. Every step was an equality, except for when we used Cauchy – Schwarz, so the issue must lie there. So when do we get equality in Cauchy – Schwarz? We'll have to look more closely at that proof! □

3. *Proof.* (a) Visually, this means that if we scale a vector $\vec{v}$ by a factor of k then we scale the length by a factor of $|k|$. The absolute value is because scaling by -1 is just reflection, which preserves length. Anyways, let's prove this. As with problem 2, we'll use the dot product here. And similarly to 2 again, square roots suck so we'll include those at the end.

   Now, we have that

$$||k\vec{v}||^2 = (k\vec{v}) \cdot (k\vec{v})$$

   Let's expand the right hand side using the definition of the dot product. Note that if

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

then

$$k\vec{v} = \begin{pmatrix} kv_1 \\ kv_2 \\ \vdots \\ kv_n \end{pmatrix}$$

So we have

$$\begin{aligned}(k\vec{v}) \cdot (k\vec{v}) &= (kv_1)(kv_1) + (kv_2)(kv_2) + \cdots + (kv_n)(kv_n) \\ &= k^2v_1^2 + k^2v_2^2 + \cdots + k^2v_n^2\end{aligned}$$

We can factor out the common k^2 here to get $k^2(v_1^2 + \cdots + v_n^2)$. The term in the parentheses is exactly $||\vec{v}||^2$. So indeed, we have shown that

$$||k\vec{v}||^2 = k^2||\vec{v}||^2$$

We take square roots on both sides to conclude

$$||k\vec{v}|| = |k|||\vec{v}||$$

(b) Here, we see that $\vec{u}$ is of the form $k\vec{v}$ with $k = \frac{1}{||\vec{v}||}$. Applying part (a), we have

$$\begin{aligned}||\vec{u}|| &= \left|\frac{1}{||\vec{v}||}\right| ||\vec{v}|| \\ &= \frac{1}{||\vec{v}||}||\vec{v}|| \\ &= 1\end{aligned}$$

as desired. We needed $\vec{v}$ to be nonzero so that we aren't dividing by 0, by the way.

□

4. *Proof.* The problem is to proceed algebraically, but let me first explain the geometry a bit. Indeed, let's recall that the geometric interpretation of the dot product led us to the formula

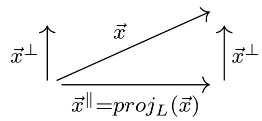
$$\vec{v} \cdot \vec{w} = ||\vec{v}||||\vec{w}|| \cos(\theta)$$

where θ is the angle between $\vec{v}$ and $\vec{w}$, taken to be between 0 and π. If we stare at this formula for a while, we might notice something nice about the right hand side. $||\vec{v}||$ is always nonnegative, as is $||\vec{w}||$. So the only term on the right hand side which could be negative is $\cos(\theta)$. This tells us that the sign of $\vec{v} \cdot \vec{w}$ is determined by the sign of $\cos(\theta)$.

So now, when is $\cos(\theta)$ positive or negative? Recall that we only consider θ to be in $[0, \pi]$. So $\cos(\theta)$ is positive for $\theta < \pi/2$ (i.e. if the angle is acute) and negative for $\theta > \pi/2$ (i.e. if the angle is obtuse). We get 0 precisely when $\theta = \pi/2$ (i.e. when the vectors are perpendicular). I'd encourage you now to try drawing some pictures.

Draw a random line on a piece of paper and then draw a random vector. Draw then the projection from that vector to the line. Is the angle obtuse? Try more lines and more vectors and convince yourself you'll never get an obtuse angle this way. Consider edge cases too – what if the vector is on the line? What if it's perpendicular?

Anyways, let's now consider the algebraic proof. Let me write $proj_L(\vec{x})$ as $\vec{x}^{\|}$. This is to indicate that the projection onto L is the component of $\vec{x}$ which lies parallel to L. There is another component of $\vec{x}$ which lies perpendicular to L, which we call $\vec{x}^{\perp}$. Indeed, we have $\vec{x}^{\perp} = \vec{x} - \vec{x}^{\|}$. Here's a diagram of what we mean here.



The point is that we can decompose $\vec{x}$ as the sum $\vec{x}^{\perp} + \vec{x}^{\|}$. Now, let's compute the dot product given this.

$$\begin{aligned}
\vec{x} \cdot proj_L(\vec{x}) &= \vec{x} \cdot \vec{x}^{\|} \\
&= (\vec{x}^{\perp} + \vec{x}^{\|}) \cdot \vec{x^{\|}} \\
&= \vec{x}^{\perp} \cdot \vec{x}^{\|} + \vec{x}^{\|} \cdot \vec{x}^{\|}
\end{aligned}$$

Now, we notice that $\vec{x}^{\perp}$ and $\vec{x}^{\|}$ are perpendicular to each other, so $\vec{x}^{\perp} \cdot \vec{x}^{\|} = 0$. This leaves us with

$$\begin{aligned}
&= \vec{x}^{\|} \cdot \vec{x}^{\|} \\
&= ||\vec{x}^{\|}||^2
\end{aligned}$$

and this value cannot be negative. Hence, $\vec{x} \cdot proj_L(\vec{x})$ cannot be negative. $\square$