Solution to Proof Questions from September 1st

Olena Bormashenko

September 4, 2011

What is a proof? A proof is an airtight logical argument that proves a certain statement in general. In a sense, it's a shortcut: for example, if you know that $\vec{x} \cdot \vec{y} = 0$ if and only if $\vec{x}$ and $\vec{y}$ are perpendicular, then for each pair of vectors $\vec{x}$ and $\vec{y}$ we now have a method of checking whether they are perpendicular: we don't need to do anything sophisticated, we just need to calculate $\vec{x} \cdot \vec{y}$.

With that in mind, here are some hints:

1. A proof usually shows something *in general.* That means a number of things: for one thing, it means that it's not enough to check a statement for a number of examples. (On the other hand, to show that a statement is *false*, it suffices to present just one counterexample. The distinction often trips people up!)

2. A proof often shows that if a certain statement A is true, then a different statement B is true. A common failure mode is to forget to use statement A in the proof – since the statement B is not supposed to be *always* true, this will by definition lead to a wrong proof!

3. And finally, the perfectly formed demonstrations that are presented in textbooks (and in class) are not in fact the way anyone comes up with a proof! In the proofs below, I will have two parts: the first will include various musings and ways to think about the question, and the second will be the actual demonstration. On homework and exams, what you want to write down is the second part: the actual proof. But it should be helpful to see how people think about this.

Proofs for the Questions in Class:

1. Show that $\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$ for any $\vec{x}$ in $\mathbb{R}^n$

   Musings: This statement seems very simple: there are no assumptions, we are just showing something in general for a vector in $\mathbb{R}^n$. Therefore, it's very reasonable to start with an arbitrary vector in $\mathbb{R}^n$, and check that the left hand side is always equal to the right-hand side.

Proof: Let $\vec{x} = [x_1, x_2, \ldots, x_n]$ be an arbitrary vector in $\mathbb{R}^n$. By definition of length,

$$\begin{aligned}\|\vec{x}\|^2 &= x_1^2 + x_2^2 + \cdots + x_n^2 = x_1 \cdot x_1 + x_2 \cdot x_2 + \cdots x_n \cdot x_n \\ &= [x_1, x_2, \ldots, x_n] \cdot [x_1, x_2, \ldots, x_n] = \vec{x} \cdot \vec{x}\end{aligned}$$

using the definition of dot product. So we're done! $\square$

2. Show that $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$ if and only if $\vec{x} \cdot \vec{y} = 0$.

Musings: This is an *if and only if* statement: that is, we want to show that $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$ happens precisely when $\vec{x} \cdot \vec{y} = 0$. The best way to think of proofs like this is that they have two parts: in part 1, we show that $\vec{x} \cdot \vec{y} = 0$ implies that $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$, and in part 2 we show that $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$ implies that $\vec{x} \cdot \vec{y} = 0$. With proofs like these, it's very important to remember what the assumptions and the desired conclusions are for each part: otherwise, it's very easy to get circular arguments!

This proof also demonstrates something else: that it's often good to use previous results to make the proof nicer. Below, I will present two proofs: one that expresses everything in terms of coordinates, just like in Question 1 above, and one that uses Question 1 and rephrases the question in terms of dot products. Note that both the proofs below are perfectly valid! One is just considerably shorter and prettier.

First Proof: In this proof, just like in Question 1, we will use the coordinates of the vectors $\vec{x}$ and $\vec{y}$ in order to simplify the above expressions. Therefore, let

$$\begin{aligned}\vec{x} &= [x_1, x_2, \ldots, x_n] \\ \vec{y} &= [y_1, y_2, \ldots, y_n]\end{aligned}$$

The above expressions will be used later on in the proof. We now break down the proof into two parts, as described above.

Part 1:
Assumptions: $\vec{x} \cdot \vec{y} = 0$
Need to show: $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$

Since we're going to use the coordinates of the vectors in this proof, it would be good to rephrase our assumptions in terms of the coordinates. Therefore, we se that we're given that

$$x_1y_1 + x_2y_2 + \cdots + x_ny_n = \vec{x} \cdot \vec{y} = 0$$

Now, let's show that $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$. We will rephrase everything in terms of the coordinates, and then hope that we can work the left-hand side into the right-hand side using the assumption. We see that by definition,

$$\begin{aligned}
\|\vec{x}+\vec{y}\|^2 &= \|[x_1+y_1, x_2+y_2, \ldots, x_n+y_n]\|^2 \\
&= (x_1+y_1)^2 + (x_2+y_2)^2 + \cdots (x_n+y_n)^2 \\
&= x_1^2 + 2x_1y_1 + y_1^2 + x_2^2 + 2x_2y_2 + y_2^2 + \cdots + x_n^2 + 2x_ny_n + y_n^2 \\
&= (x_1^2 + \cdots + x_n^2) + 2(x_1y_1 + \cdots + x_ny_n) + (y_1^2 + \cdots + y_n^2)
\end{aligned}$$

Now, using the definitions, we see that $x_1^2+\cdots+x_n^2$ is $\|\vec{x}\|^2$, $y_1^2+\cdots+y_n^2$ is $\|\vec{y}\|^2$, and $x_1y_1+\cdots+x_ny_n$ is 0. Therefore, the above equation simplifies to

$$\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

which is what we wanted to show.

Part 2:
Assumptions: $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x}+\vec{y}\|^2$
Need to show: $\vec{x}\cdot\vec{y} = 0$

Writing things out similarly to above, we see that we're given

$$(x_1+y_1)^2 + \cdots (x_n+y_n)^2 = (x_1^2+\cdots+x_n^2) + (x_1^2+\cdots+x_n^2)$$

Expanding and gathering terms, this says precisely that

$$(x_1^2+\cdots+x_n^2)+2(x_1y_1+\cdots+x_ny_n)+(y_1^2+\cdots+y_n^2) = (x_1^2+\cdots+x_n^2)+(x_1^2+\cdots+x_n^2)$$

and now subtracting the right-hand side from both sides, we see that

$$2(x_1y_1+\cdots+x_ny_n) = 0$$

Dividing by 2, we now see that

$$0 = x_1y_1+\cdots+x_ny_n = \vec{x}\cdot\vec{y}$$

as desired. □

Second Proof: This proof is simpler because instead of using coordinates (which got a little messy), we use the fact that $\|\vec{z}\|^2 = \vec{z}\cdot\vec{z}$ for any vector $\vec{z}$. Otherwise, the proof is remarkably similar.

Part 1:
Assumptions: $\vec{x}\cdot\vec{y} = 0$
Need to show: $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x}+\vec{y}\|^2$

Similarly to before, we will want to manipulate $\|\vec{x}+\vec{y}\|^2$ to be $\|\vec{x}\|^2+\|\vec{y}\|^2$, in the process using the assumption that $\vec{x}\cdot\vec{y} = 0$. Using the fact that $\|\vec{z}\|^2 = \vec{z}\cdot\vec{z}$ for any vector $\vec{z}$,

$$\|\vec{x}+\vec{y}\|^2 = (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) = \vec{x}\cdot\vec{x} + 2\vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{y}$$

using the properties of dot products. Since $\vec{x} \cdot \vec{y} = 0$, this simplifies to

$$\|\vec{x} + \vec{y}\|^2 = \vec{x} \cdot \vec{x} + \vec{y} \cdot \vec{y} = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

which is what we wanted.

Part 2:
Assumptions: $\|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{x} + \vec{y}\|^2$
Need to show: $\vec{x} \cdot \vec{y} = 0$

Again, expanding everything out, what we're given is that:

$$\vec{x} \cdot \vec{x} + \vec{y} \cdot \vec{y} = (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) = \vec{x} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y}$$

and subtracting $\vec{x} \cdot \vec{x} + \vec{y} \cdot \vec{y}$ from both sides, then dividing by 2, we get

$$\vec{x} \cdot \vec{y} = 0$$

which is what we wanted! $\square$

3. Show that if $\vec{x} \in \mathbb{R}^n$ and $\vec{x} \cdot \vec{y} = 0$ for every $\vec{y} \in \mathbb{R}^n$, then $\vec{x} = \vec{0}$.

   Musings: This one seems a little tricky. Recall that the dot product of $\vec{x}$ and $\vec{y}$ is 0 whenever $\vec{x}$ and $\vec{y}$ are perpendicular (or if one of them is $\vec{0}$.) Here, we're essentially given that $\vec{x}$ is perpendicular to every single vector in $\mathbb{R}^n$. It seems like the only way that's possible is if it's the $\vec{0}$ vector... but how do we show that?

   Let's see... It really doesn't make sense that something would be perpendicular to *everything*. What's a good example of a vector $\vec{x}$ shouldn't be perpendicular to? Well... vectors don't tend to be perpendicular to themselves, do they? Maybe that's a good thing to plug in for our $\vec{y}$!

   Proof: Since $\vec{x} \cdot \vec{y} = 0$ for every $\vec{y}$ in $\mathbb{R}^n$, we have that

   $$0 = \vec{x} \cdot \vec{x} = \|\vec{x}\|^2$$

   This means that the length of $\vec{x}$ is 0, and therefore it must be $\vec{0}$. $\square$

4. Use Cauchy-Schwarz to prove the triangle inequality: that is, that $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.

   Musings: Cauchy-Schwarz states that

   $$|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$$

   We're asked to use this fact to show the triangle inequality. Dot products appear when we consider the length squared, so maybe it'd be a good idea to work with that instead. Is it enough to consider the squares? Well, let's say we managed to show that

$$\|\vec{x}+\vec{y}\|^2 \leq (\|\vec{x}\|+\|\vec{y}\|)^2$$

Are we now allowed to take the square root of both sides? (We aren't always! For example, $(1)^2 < (-4)^2$, but it's certainly not true that $1 < -4$.) Yes, we are: since lengths are non-negative, we see that $\|\vec{x}+\vec{y}\|$ and $\|\vec{x}\|+\|\vec{y}\|$ are non-negative, so taking the square root of both sides is allowed. Therefore, showing that

$$\|\vec{x}+\vec{y}\|^2 \leq (\|\vec{x}\|+\|\vec{y}\|)^2$$

is a good goal.

Here, since we're asked to use Cauchy-Schwarz explicitly, it'd be good to put it as one of our assumptions: that way, we won't forget to use it!

Proof:
Assumptions: $|\vec{x}\cdot\vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$
Need to show: $\|\vec{x}+\vec{y}\| \leq \|\vec{x}\|+\|\vec{y}\|$

Since both $\|\vec{x}+\vec{y}\|$ and $\|\vec{x}\|+\|\vec{y}\|$ are non-negative, it suffices to show that

$$\|\vec{x}+\vec{y}\|^2 \leq (\|\vec{x}\|+\|\vec{y}\|)^2$$

Now, rewriting the left-hand side in terms of dot products,

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 &= (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) = \vec{x}\cdot\vec{x}+2\vec{x}\cdot\vec{y}+\vec{y}\cdot\vec{y}\\ &= \|\vec{x}\|^2+\|\vec{y}\|^2+2\vec{x}\cdot\vec{y}\end{aligned}$$

Cauchy-Schwarz says that $|\vec{x}\cdot\vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$. Since a number is smaller than its absolute value, this means that $\vec{x}\cdot\vec{y} \leq \|\vec{x}\|\|\vec{y}\|$. Plugging that in above,

$$\begin{aligned}\|\vec{x}+\vec{y}\|^2 &\leq \|\vec{x}\|^2+\|\vec{y}\|^2+2\|\vec{x}\|\|\vec{y}\|\\ &= (\|\vec{x}\|+\|\vec{y}\|)^2\end{aligned}$$

which is precisely what we wanted. Hence, we're done. □

5. (Extra) Prove Cauchy-Schwarz: that is, prove that $|\vec{x}\cdot\vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$.

Musings: This proof, unfortunately, is just kind of tricky and hard. All the proofs I've ever seen look a bit like magic – it's not clear how someone managed to come up with them. The proof in the book on page 19 is pretty good, but it still has that property.

Ultimately, doing mathematics is precisely about doing proofs where it's not clear how to get started. But that's hard to teach, and not what these problems were about! (For now, we're just trying to get the mechanics of proofs down.) Anyway, if you're interested – take a look in the book.