CLARK UNIVERSITY

Cauchy's inequality
Math 130 Linear Algebra
D Joyce, Fall 2015

The triangle inequality and angles in n-space. We worked from principles of geometry to develop the triangle inequality in dimension 2, and it works in dimension 3 as well, but we can't rely on a geometry for higher dimensions since we don't have it. Likewise we used the law of cosines to connect inner products to angles, but that only worked in low dimensions, too. The question is, how do we do these things in higher dimensions?

Also, we were working in real vector spaces. Can we do the same things in complex vector spaces? The answer is yes, but it's not based on geometry, but on algebra, in particular, something called Cauchy's inequality

$$|\langle \mathbf{v}|\mathbf{w}\rangle| \leq \|\mathbf{v}\|\,\|\mathbf{w}\|.$$

It holds in any dimension and it works for complex vector spaces, too. It's been generalized all over the place, to infinite dimensional space, to integrals, and to probabilities. It sometimes goes by the name Cauchy-Bunyakovsky-Schwarz inequality, but it started with Cauchy in 1821.

An elementary proof of the Cauchy inequality. The early proofs of the Cauchy inequality used coordinates, and that's what we'll do here. This proof is only valid for the standard spaces $\mathbf{R}^n$ and $\mathbf{C}^n$. There are more recent proofs that work for general abstract inner product spaces.

Since we're working with n-vectors, summation notation will facilitate the proof. There's nothing tricky or unexpected in this proof, just a lot of details. The early proof only applied to $\mathbf{R}^n$, but it works just as well for $\mathbf{C}^n$ at the expense of adding complex conjugates to half the terms. They can be ignored for the real case since the complex conjugate of a real number is itself.

Proof. In order to prove the Cauchy inequality, we'll prove its square instead

$$|\langle \mathbf{v}|\mathbf{w}\rangle|^2 \leq \|\mathbf{v}\|^2\,\|\mathbf{w}\|^2.$$

We're considering the case of complex vector spaces as well as real vector spaces, so we'll have to write bars for complex conjugation. Starting with the left hand side of the inequality, we have

$$|\langle \mathbf{v}|\mathbf{w}\rangle|^2 = \langle \mathbf{v}|\mathbf{w}\rangle\overline{\langle \mathbf{v}|\mathbf{w}\rangle} = \langle \mathbf{v}|\mathbf{w}\rangle\langle \mathbf{w}|\mathbf{v}\rangle = \left(\sum_i v_i\overline{w_i}\right)\left(\sum_j w_j\overline{v_j}\right) = \sum_{i,j} v_i\overline{w_i}w_j\overline{v_j}$$

The right hand side of the inequality is

$$\|\mathbf{v}\|^2 \, \|\mathbf{w}\|^2 = \left(\sum_i v_i \overline{v_i} \right) \left(\sum_j w_j \overline{w_j} \right) = \sum_{i,j} v_i \overline{v_i} w_j \overline{w_j}$$

Thus, our goal is to show

$$\sum_{i,j} v_i \overline{w_i} w_j \overline{v_j} \leq \sum_{i,j} v_i \overline{v_i} w_j \overline{w_j},$$

which is equivalent to

$$0 \leq \sum_{i,j} (v_i \overline{v_i} w_j \overline{w_j} - v_i \overline{w_i} w_j \overline{v_j}).$$

Now, the terms in which $i = j$ are all 0, so we can drop them. We'll pair the ij^{th} term with the ji^{th} term to write the inequality as

$$0 \leq \sum_{i<j} (v_i \overline{v_i} w_j \overline{w_j} - v_i \overline{w_i} w_j \overline{v_j} + v_j \overline{v_j} w_i \overline{w_i} - v_j \overline{w_j} w_i \overline{v_i}).$$

Now the ij^{th} term can be factored

$$(v_i \overline{v_i} w_j \overline{w_j} - v_i \overline{w_i} w_j \overline{v_j} + v_j \overline{v_j} w_i \overline{w_i} - v_j \overline{w_j} w_i \overline{v_i}) = (v_i w_j - v_j w_i)(\overline{v_i} \overline{w_j} - \overline{v_j} \overline{w_i}) = |v_i w_j - v_j w_i|^2$$

which is nonnegative. Since all the terms are greater than or zero, so is their sum, and we've proved the Cauchy inequality. Q.E.D.

The triangle inequality. We can use the Cauchy inequality to prove the triangle inequality

$$\|\mathbf{v} - \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

Proof. It will simplify our proof slightly if we prove the equivalent inequality $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$. That way we won't have to deal with negative signs. As usual, we'll prove that by proving its square

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

The left hand side is

$$\|\mathbf{v} + \mathbf{w}\|^2 = \langle \mathbf{v} + \mathbf{w} | \mathbf{v} + \mathbf{w} \rangle = \|\mathbf{v}\|^2 + \langle \mathbf{v} | \mathbf{w} \rangle + \langle \mathbf{w} | \mathbf{v} \rangle + \|\mathbf{w}\|^2,$$

while the right hand side is $(\|\mathbf{v}\| + \|\mathbf{w}\|)^2 = \|\mathbf{v}\|^2 + 2\|\mathbf{v}\| \, \|\mathbf{w}\| + \|\mathbf{w}\|^2$. We can subtract the common terms $\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$ from both sides to reduce our goal to showing that

$$\langle \mathbf{v} | \mathbf{w} \rangle + \langle \mathbf{w} | \mathbf{v} \rangle \leq 2\|\mathbf{v}\| \, \|\mathbf{w}\|.$$

Now the hand left side is $\overline{\langle \mathbf{w} | \mathbf{v} \rangle} + \langle \mathbf{w} | \mathbf{v} \rangle$. In general, for any complex number z, $z + \overline{z} \leq 2|z|$, so we know the left hand side is less than or equal to $2|\langle \mathbf{w} | \mathbf{v} \rangle|$, and the Cauchy inequality says that's less than or equal to the right hand side. Thus we've proved our goal. Q.E.D.

Angles. In either the real or complex case, we'll say that $\mathbf{v}$ and $\mathbf{w}$ are *orthogonal*, denoted $\mathbf{v} \perp \mathbf{w}$, if $\langle \mathbf{v}|\mathbf{w}\rangle = 0$.

The Pythagorean theorem follows directly from this definition. In this context, it says

$$\mathbf{v} \perp \mathbf{w} \quad \text{implies} \quad \|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

The converse of this Pythagorean theorem holds in the real case, but not in the complex case. (In the complex case, $\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$ only implies that $\langle \mathbf{v}|\mathbf{w}\rangle$ is purely imaginary, that is, has no real part.)

In the real case, the Cauchy inequality says that

$$-1 \leq \frac{\langle \mathbf{v}|\mathbf{w}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|} \leq 1.$$

That allows us to define the angle θ between the vectors $\mathbf{v}$ and $\mathbf{w}$ to be the arccosine of $\dfrac{\langle \mathbf{v}|\mathbf{w}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$. When $\mathbf{v}$ and $\mathbf{w}$ are unit vectors, their inner product $\langle \mathbf{v}|\mathbf{w}\rangle$ is $\cos\theta$. The law of cosines follows from this definition:

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\,\|\mathbf{w}\|\cos\theta.$$

In the complex case, we know $\dfrac{\langle \mathbf{v}|\mathbf{w}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$ is a complex number that lies in the closed unit disk, that is to say, it's a complex number less than or equal to 1 unit away from the origin. There is no good definition of the angle between two complex vectors, but this quantity, $\dfrac{\langle \mathbf{v}|\mathbf{w}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$, does a good job as a stand-in for its cosine.

(Since $\dfrac{\langle \mathbf{w}|\mathbf{v}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$ is its conjugate, therefore their sum is real, and their average lies between -1 and 1. We could take θ to be the arccosine of that average so that

$$\cos\theta = \frac{1}{2}\,\frac{\langle \mathbf{v}|\mathbf{w}\rangle + \langle \mathbf{w}|\mathbf{v}\rangle}{\|\mathbf{v}\|\,\|\mathbf{w}\|}$$

Again, the law of cosines, written above, would hold. Unfortunately, $\cos\theta$ being 0 would only say that $\langle \mathbf{v}|\mathbf{w}\rangle$ is purely imaginary, not that it's 0, and we couldn't conclude that $\mathbf{v} \perp \mathbf{w}$.)

Math 130 Home Page at `http://math.clarku.edu/~ma130/`