L18Orthogonality

April 12, 2015

1 Orthogonality

Today we'll start to bring **geometry** to center stage in our discussion.

We'll concern ourselves with simple notions: length, distance, perpendicularity, and angle. However we will take these notions that are familiar from our 3D world and see how to define them for spaces of arbitrary dimension, ie, $\mathbb{R}^n$.

To start, we'll return and review the **inner product.**

1.1 Inner Product (Revisited)

Recall that we consider vectors such as $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ to be $n \times 1$ matrices.

Then $\mathbf{u}^T\mathbf{v}$ is a scalar, called the **inner product** of $\mathbf{u}$ and $\mathbf{v}$.

You will also see this called the **dot product**. It sometimes written as $\mathbf{u}\cdot\mathbf{v}$ but we will always write $\mathbf{u}^T\mathbf{v}$.

The inner product is the sum of the componentwise product of $\mathbf{u}$ and $\mathbf{v}$:

If $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, then the inner product of $\mathbf{u}$ and $\mathbf{v}$ is:

$$\begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n = \sum_{i=1}^n u_iv_i.$$

Let's remind ourselves of the properties of the inner product:

Theorem. Let $\mathbf{u}$,$\mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then:

1. $\mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$
2. $(\mathbf{u}+\mathbf{v})^T\mathbf{w} = \mathbf{u}^T\mathbf{w} + \mathbf{v}^T\mathbf{w}$
3. $(c\mathbf{u})^T\mathbf{v} = c(\mathbf{u}^T\mathbf{v}) = \mathbf{u}^T(c\mathbf{v})$
4. $\mathbf{u}^T\mathbf{u} \geq 0, \quad$ and $\mathbf{u}^T\mathbf{u} = 0$ if and only if $\mathbf{u} = 0$

1.2 Vector Norm

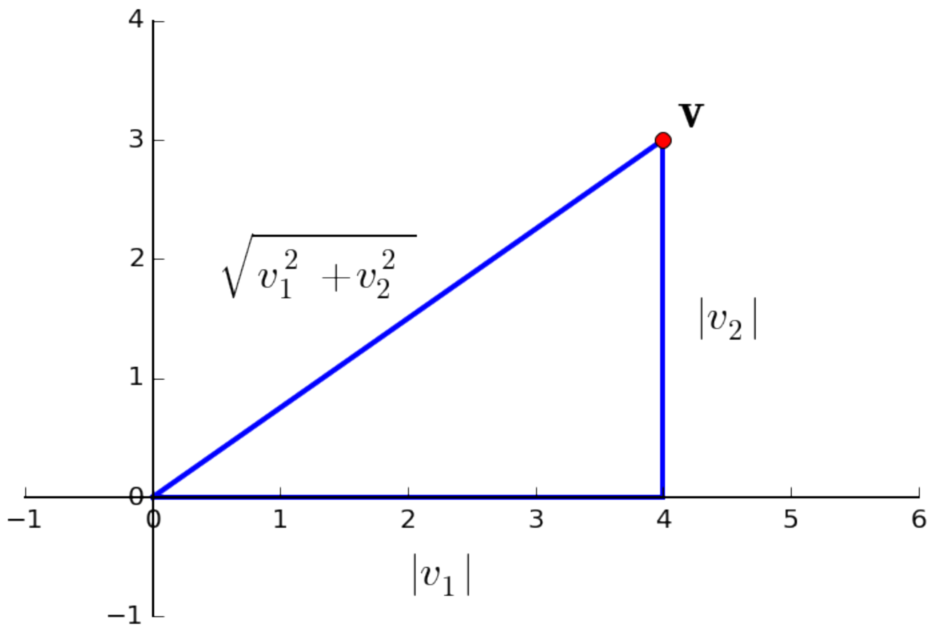
If $\mathbf{v}$ is in $\mathbb{R}^n$, with entries $v_1, \dots, v_n$, then the square root of $\mathbf{v}^\mathbf{v}$ is defined because $\mathbf{v}^T\mathbf{v}$ is nonnegative.

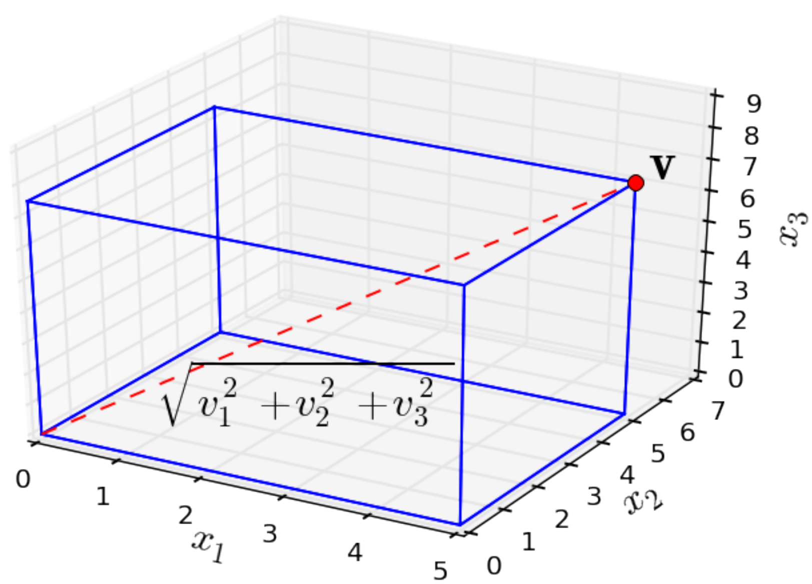
Definition. The **norm** of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}^T\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = \sqrt{\sum_{i=1}^n {v_i}^2}.$$

The norm of $\mathbf{v}$ is its **length** in the usual sense.

This follows directly from the Pythagorean theorem:





For any scalar c, the length of $c\mathbf{v}$ is $|c|$ times the length of $\mathbf{v}$. That is,

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|.$$

So, for example, $\|(-2)\mathbf{v}\| = \|2\mathbf{v}\|$.
A vector of length 1 is called a **unit vector**.
If we divide a nonzero vector $\mathbf{v}$ by its length – that is, multiply by $1/\|\mathbf{v}\|$ – we obatin a unit vector $\mathbf{u}$.
We say that we have *normalized* $\mathbf{v}$, and that $\mathbf{u}$ is *in the same direction* as $\mathbf{v}$.

Example. Let $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix}$. Find the unit vector $\mathbf{u}$ in the same direction as $\mathbf{v}$.

Solution.
First, compute the length of $\mathbf{v}$:

$$\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9$$

$$\|\mathbf{v}\| = \sqrt{9} = 3$$

Then multiply $\mathbf{v}$ by $1/\|\mathbf{v}\|$ to obtain

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{3}\mathbf{v} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

It's interesting that we can't actually visualize $\mathbf{u}$ but we can still reason geometrically about it as a unit vector.

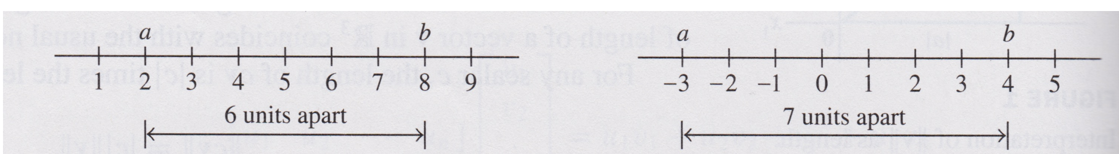
Example. Let W be the subspace of $\mathbb{R}^2$ spanned by $\mathbf{x} = \begin{bmatrix} 2/3 \\ 1 \end{bmatrix}$. Find a unit vector $\mathbf{z}$ that is a basis for W.

Solution.

1.3 Distance in $\mathbb{R}^n$

It's very useful to be able to talk about the **distance** between two points (or vectors) in $\mathbb{R}^n$.

We can start from basics:



On the number line, the distance between two points a and b is $|a - b|$.
The same is true in $\mathbb{R}^n$.

Definition. For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the **distance between u and v,** written as $\text{dist}(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

This definition agrees with the usual formulas for the Euclidean distance between two points. The usual formula is

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \sqrt{(v_1 - u_1)^2 + (v_2 - u_2)^2 + \cdots + (v_n - u_n)^2}.$$

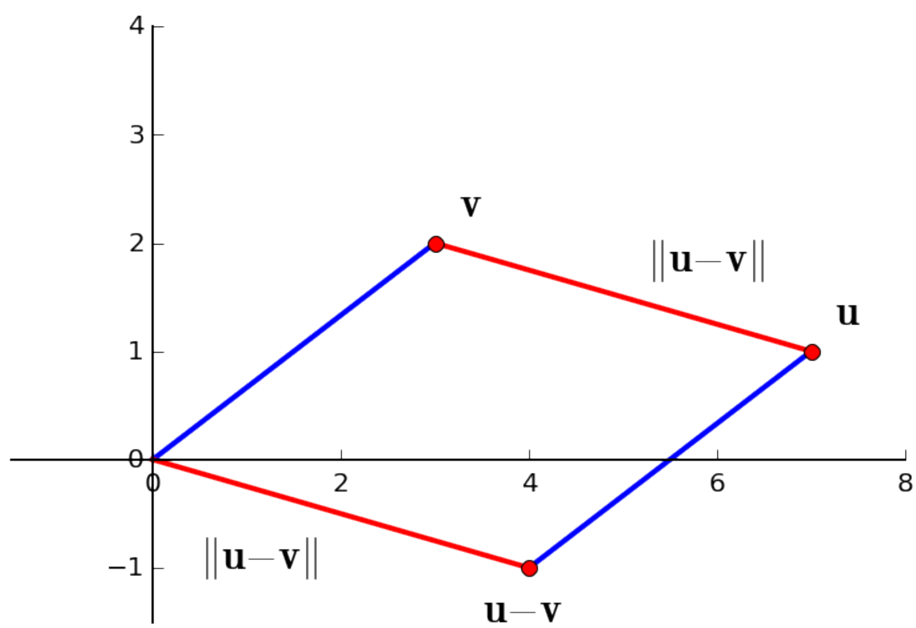
Which you can see is equal to

$$\|\mathbf{u}-\mathbf{v}\| = \sqrt{(\mathbf{u}-\mathbf{v})^T(\mathbf{u}-\mathbf{v})} = \sqrt{\begin{bmatrix} u_1-v_1 & u_2-v_2 & \ldots & u_n-v_n \end{bmatrix} \begin{bmatrix} u_1-v_1 \\ u_2-v_2 \\ \vdots \\ u_n-v_n \end{bmatrix}}$$

There is a geometric view as well.

For example, consider the vectors $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ in $\mathbb{R}^2$.

Then one can see that the distance from $\mathbf{u}$ to $\mathbf{v}$ is the same as the length of the vector $\mathbf{u}-\mathbf{v}$.

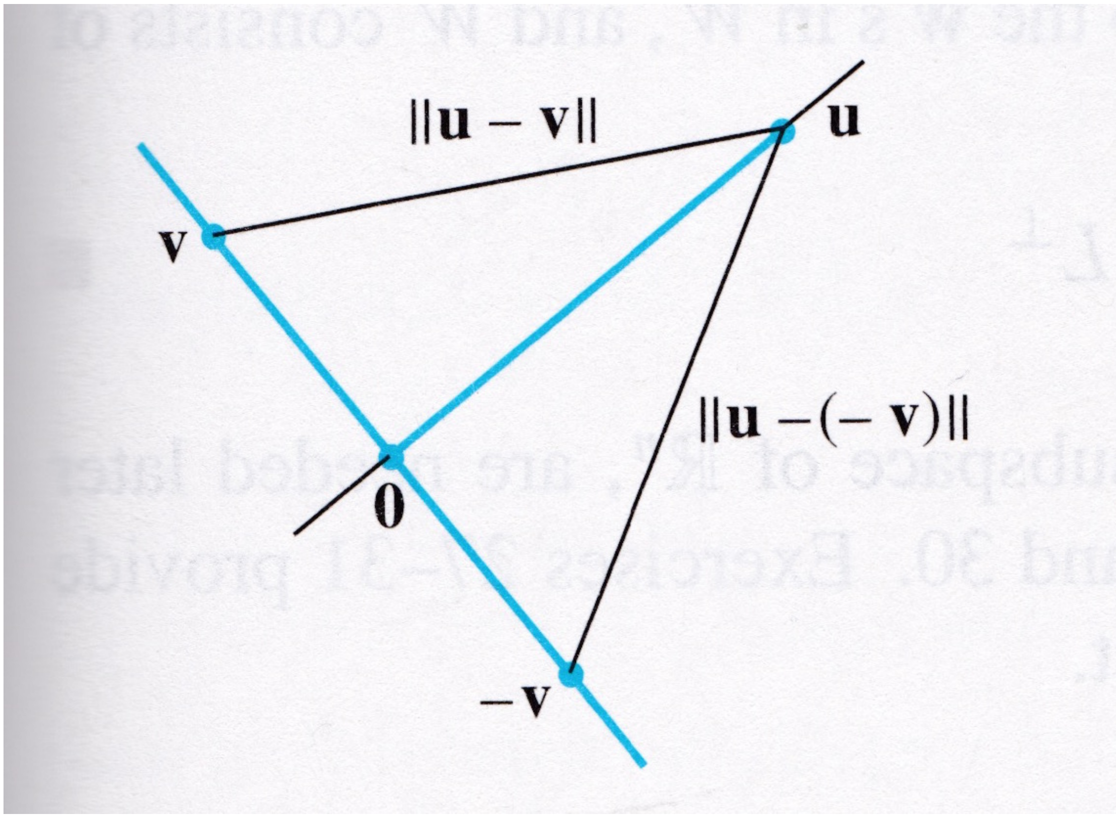


This shows that the distance between two vectors is the length of their difference.

1.4 Orthogonal Vectors

Now we turn to another familiar notion from 2D geometry, which we'll generalize to $\mathbb{R}^n$: the notion of being perpendicular.

You may recall the classic Euclidean way to construct a line perpedicular to another line at a point:



One constructs an isoceles triangle centered at the point. Because the sides are equal, the two inner triangles are right triangles.

So the two blue lines are perpendicular if and only if the distance from $\mathbf{u}$ to $\mathbf{v}$ is equal to the distance from $\mathbf{u}$ to $-\mathbf{v}$.

This is the same as requiring the squares of their distances to be equal.

Let's see what this implies from an algebraic standpoint:

$$[\text{dist}(\mathbf{u}, -\mathbf{v})]^2 = \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2$$

$$= (\mathbf{u} + \mathbf{v})^T(\mathbf{u} + \mathbf{v})$$

$$= \mathbf{u}^T(\mathbf{u} + \mathbf{v}) + \mathbf{v}^T(\mathbf{u} + \mathbf{v})$$

$$= \mathbf{u}^T\mathbf{u} + \mathbf{u}^T\mathbf{v} + \mathbf{v}^T\mathbf{u} + \mathbf{v}^T\mathbf{v}$$

$$= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u}^T\mathbf{v}$$

Now let's find the distance from $\mathbf{u}$ to $\mathbf{v}$:

$$[\text{dist}(\mathbf{u}, \mathbf{v})]^2 = \|\mathbf{u} - \mathbf{v}\|^2$$

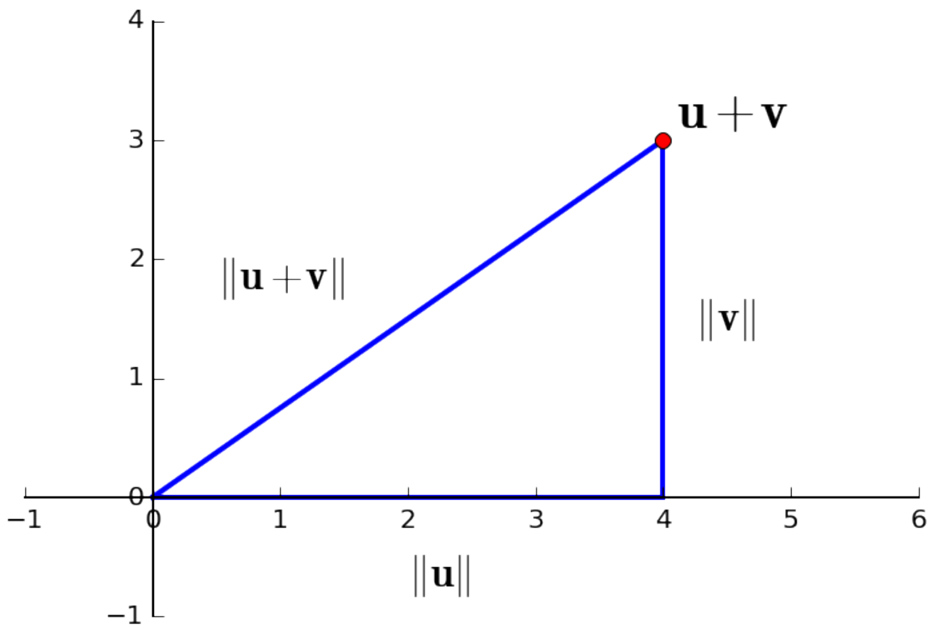
$$= (\mathbf{u} - \mathbf{v})^T(\mathbf{u} - \mathbf{v})$$

$$= \mathbf{u}^T(\mathbf{u} - \mathbf{v}) - \mathbf{v}^T(\mathbf{u} - \mathbf{v})$$

$$= \mathbf{u}^T\mathbf{u} - \mathbf{u}^T\mathbf{v} - \mathbf{v}^T\mathbf{u} + \mathbf{v}^T\mathbf{v}$$

$$= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}^T\mathbf{v}$$

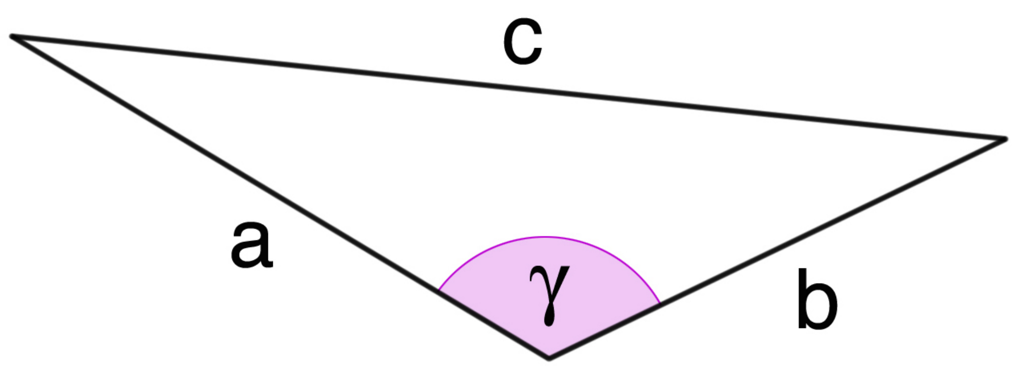
So $\text{dist}(\mathbf{u}, \mathbf{v}) = \text{dist}(\mathbf{u}, -\mathbf{v})$ if and only if $\mathbf{u}^T\mathbf{v} = 0$.
So now we can define perpendicularity in $\mathbb{R}^n$:
Definition. Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are **orthogonal** to each other if $\mathbf{u}^T\mathbf{v} = 0$.
This also allows us to restate the Pythagorean Theorem:
Theorem. Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.



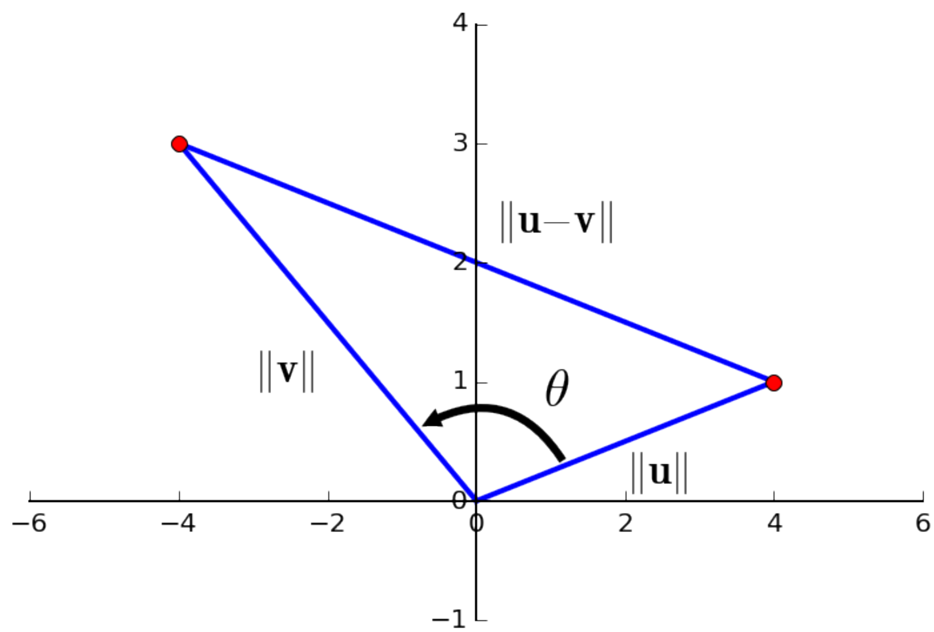
1.5 The Angle Between Two Vectors

There is an important connection between the inner product of two vectors and the **angle** between them.
This connection is very useful, eg, in visualizing data mining operations.
We start from the **law of cosines:**

$$c^2 = a^2 + b^2 - 2ab\cos\gamma$$



Now let's interpret this law in terms of vectors $\mathbf{u}$ and $\mathbf{v}$:



Applying the law of cosines we get:

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$$

Now, previously we calculated that:

$$\|\mathbf{u} - \mathbf{v}\|^2 = (\mathbf{u} - \mathbf{v})^T(\mathbf{u} - \mathbf{v})$$

$$= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}^T\mathbf{v}$$

Which means that

$$2\mathbf{u}^T\mathbf{v} = 2\|\mathbf{u}\|\|\mathbf{v}\| \cos\theta$$

So

$$\mathbf{u}^T\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta$$

This is a very important connection between inner product and trigonometry.

One implication in particular concerns **unit vectors.**

$$\mathbf{u}^T\mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\|\cos\theta$$

So

$$\frac{\mathbf{u}^T\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \cos\theta$$

$$\frac{\mathbf{u}^T}{\|\mathbf{u}\|}\frac{\mathbf{v}}{\|\mathbf{v}\|} = \cos\theta$$

Note that $\frac{\mathbf{u}}{\|\mathbf{u}\|}$ and $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ are unit vectors.

So we have the very simple rule, that for two unit vectors, their inner product is the cosine of the angle between them!

1.6 Orthogonal Complements

In 3D geometry we also encounter the notion of a *line perpendicular to a plane.*

We generalize this to $\mathbb{R}^n$ by realizing that in $\mathbb{R}^n$, the plane becomes a *subspace.*

Definition. If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$ then $\mathbf{z}$ is said to be **orthogonal to** W**.**

```
In [ ]:
```