Orthogonality

Transcript of oral arguments before the US Supreme Court, Briscoe v. Virginia, January 11, 2010:

MR. FRIEDMAN: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging –

CHIEF JUSTICE ROBERTS: I'm sorry. Entirely what?

MR. FRIEDMAN: Orthogonal. Right angle. Unrelated. Irrelevant.

CHIEF JUSTICE ROBERTS: Oh.

JUSTICE SCALIA: What was that adjective? I liked that.

MR. FRIEDMAN: Orthogonal.

CHIEF JUSTICE ROBERTS: Orthogonal.

MR. FRIEDMAN: Right, right.

JUSTICE SCALIA: Orthogonal, ooh.

(Laughter.)

JUSTICE KENNEDY: I knew this case presented us a problem.

(Laughter.)

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

We'll concern ourselves with simple notions:
* length, * distance, * orthogonality (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, $\mathbb{R}^n$.

Interestingly, it turns out that these notions (length, distance, purpendicularity, angle) **all** depend on one key notion: the **inner product**.

In fact, the notion is so important that we refer to a vector space for which there is an inner product as an **inner product space.**

So let's briefly return to and review the **inner product.**

Inner Product (Review)

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:

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

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 $\mathbf{u} = 0$

The first three are restatements of facts about matrix-vector products.

The last one is straightforward, but important.

Now, given that review, let's start talking about geometry.

Vector Norm

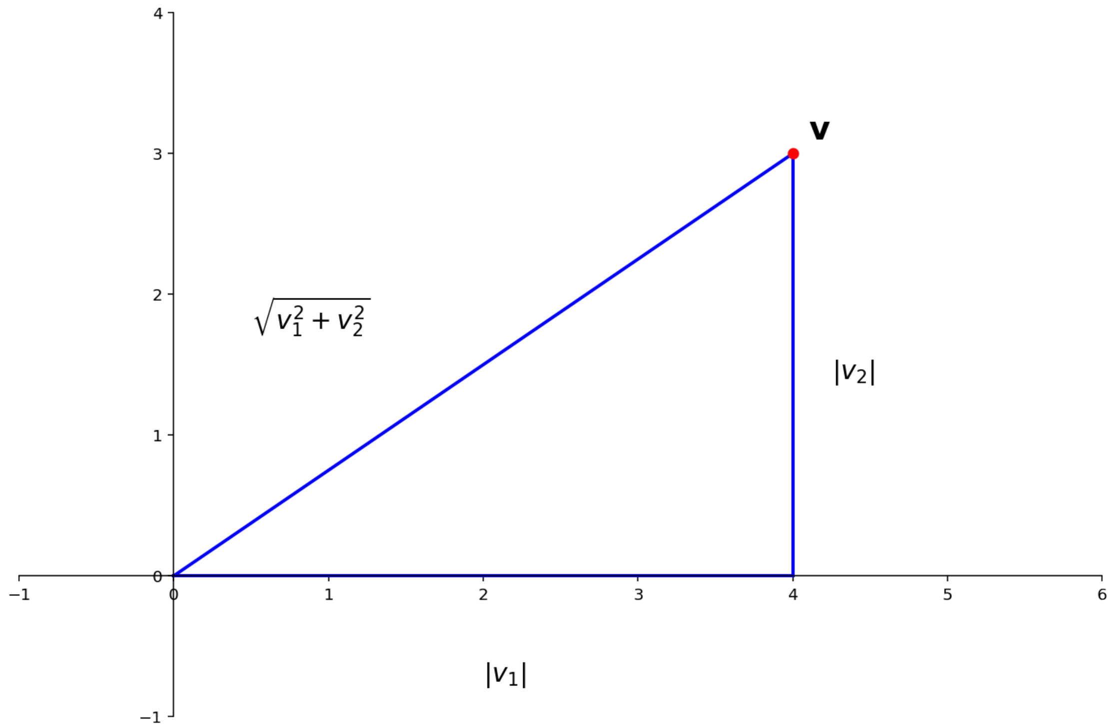
If $\mathbf{v}$ is in $\mathbb{R}^n$, with entries $v_1, \ldots, v_n$, then the square root of $\mathbf{v}^T\mathbf{v}$ is defined because $\mathbf{v}^T\mathbf{v}$ is nonnegative.

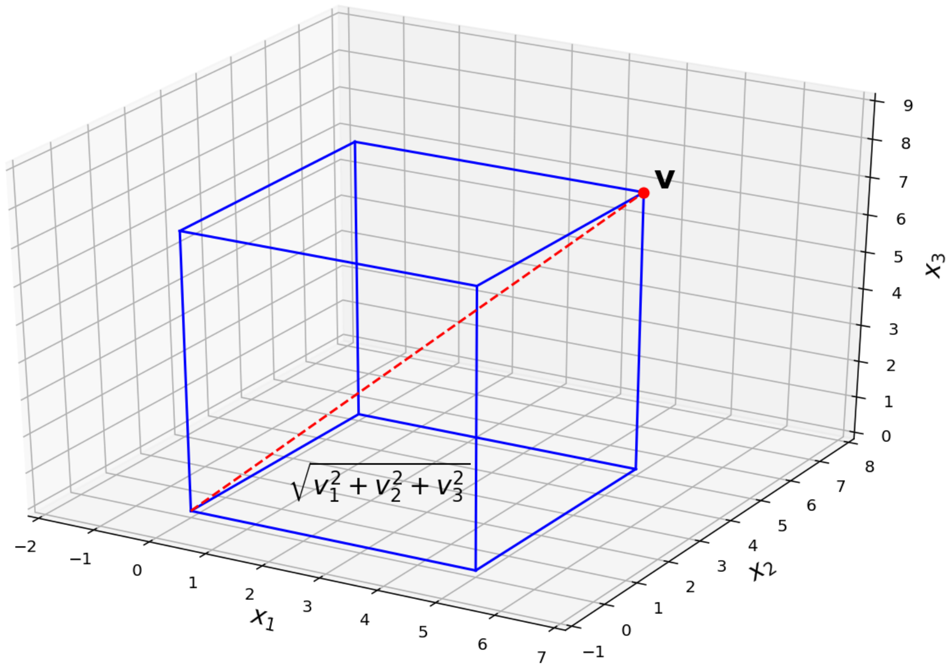
Definition. The **norm** of $\mathbf{v}$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

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

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 obtain 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

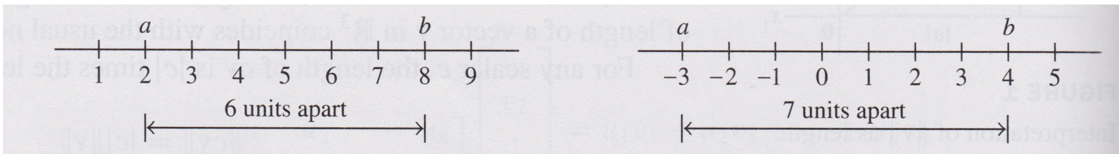
$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{3}\mathbf{v} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$$

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

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

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

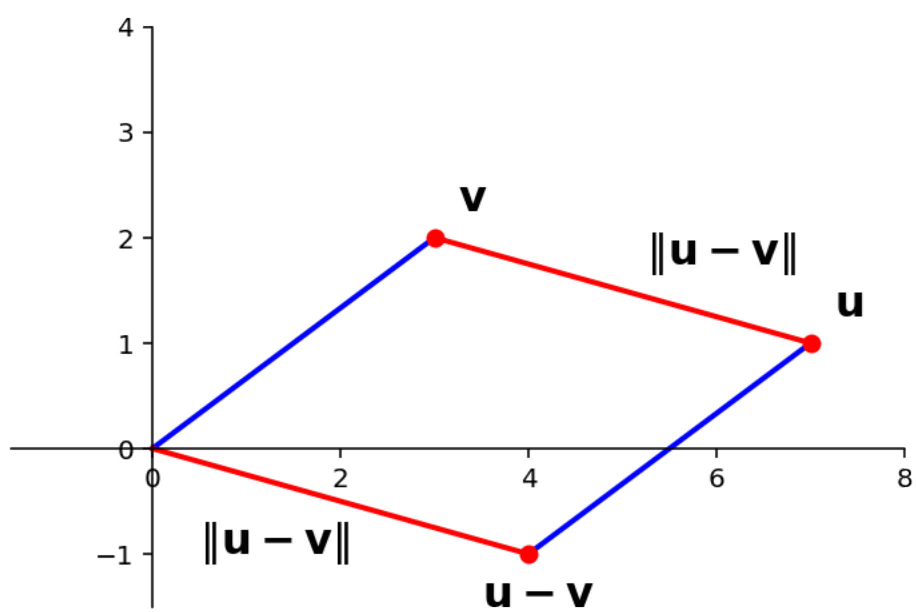
Which you can see is equal to

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

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 $\mathbf{u} - \mathbf{v}$.



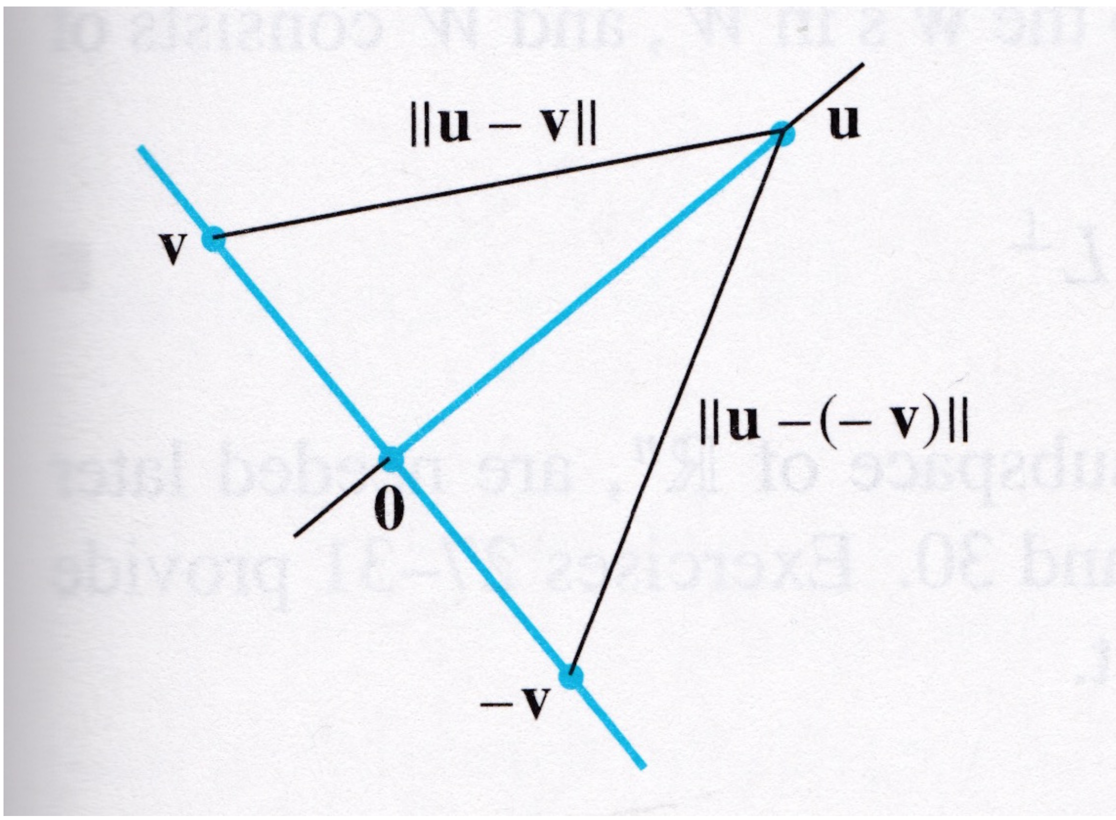
This shows that the distance between two vectors is the length of their difference.

Question Time! Q 20.1

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 method from Euclid for how to construct a line perpendicular 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. First we'll find the distance from $\mathbf{u}$ to $-\mathbf{v}$:

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

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

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

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

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

Now let's find the distance from **u** to **v**:

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

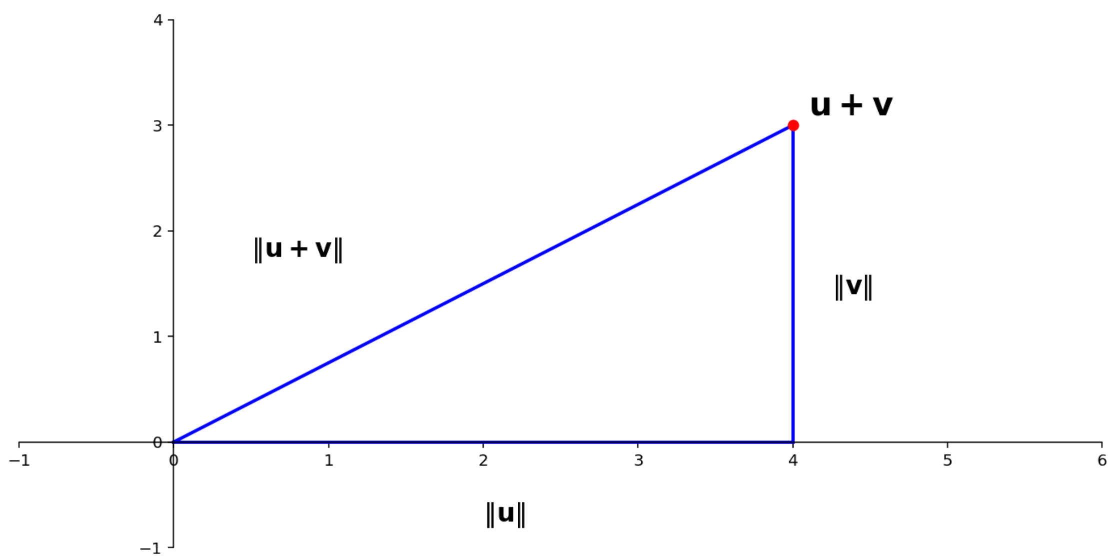
$$= (\mathbf{u} - \mathbf{v})^T(\mathbf{u} - \mathbf{v})$$

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

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

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

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 **u** and **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 **u** and **v** are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

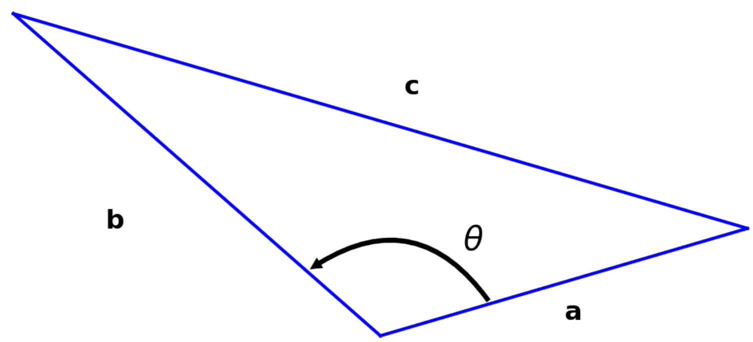


Question Time! Q20.2

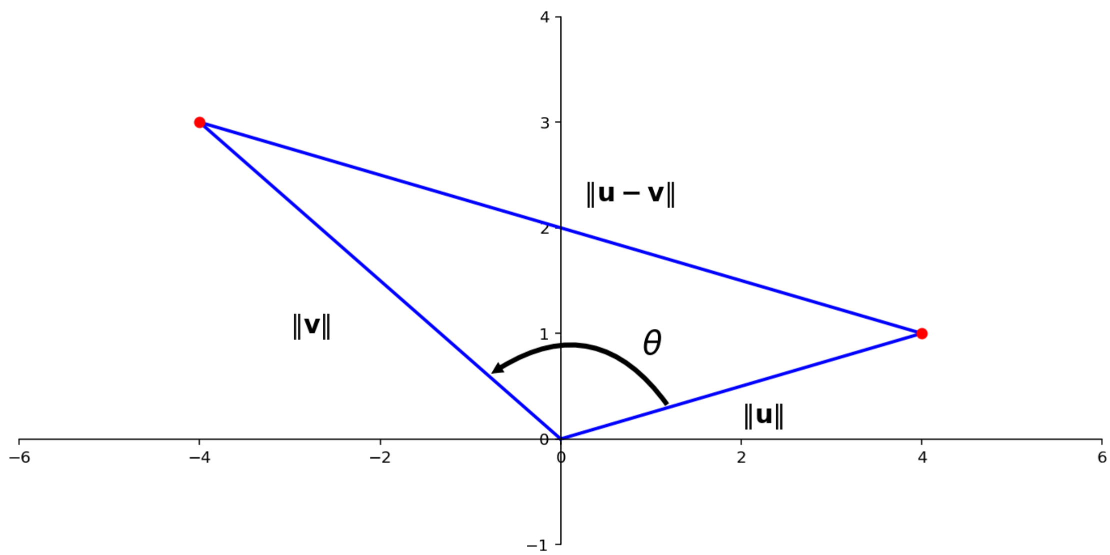
The Angle Between Two Vectors

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:**

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



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 the notion of inner product and trigonometry.

As a quick check, note that if $\mathbf{u}$ and $\mathbf{v}$ are nonzero, and $\mathbf{u}^T\mathbf{v} = 0$, then $\cos\theta = 0$.

In other words, the angle between $\mathbf{u}$ and $\mathbf{v}$ is 90 degrees (or 270 degrees). So this agrees with our definition of orthogonality.

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

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

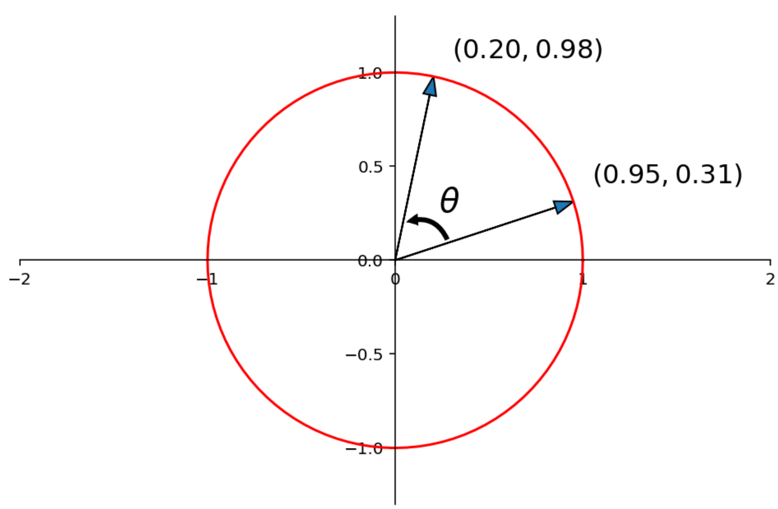
So

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

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

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!



Here $\mathbf{u} = \begin{bmatrix}0.95\\0.31\end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix}0.20\\0.98\end{bmatrix}$.

So $\mathbf{u}^T\mathbf{v} = (0.95 \cdot 0.20) + (0.31 \cdot 0.98) = 0.5$

So $\cos\theta = 0.5$.

So $\theta = 60$ degrees.

Example. Find the angle formed by the vectors:

$$\mathbf{u} = \begin{bmatrix}1\\3\\-7\\-2\end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix}8\\-2\\4\\6\end{bmatrix}$$

Solution.
First normalize the vectors:

$$\|\mathbf{u}\| = \sqrt{1^2 + 3^2 + (-7)^2 + (-2)^2} = 7.93$$

$$\|\mathbf{v}\| = \sqrt{8^2 + (-2)^2 + 4^2 + 6^2} = 10.95$$

So

$$\frac{\mathbf{u}}{\|\mathbf{u}\|} = \begin{bmatrix} 0.13 \\ 0.38 \\ -0.88 \\ -0.25 \end{bmatrix} \quad \text{and} \quad \frac{\mathbf{v}}{\|\mathbf{v}\|} = \begin{bmatrix} 0.73 \\ -0.18 \\ 0.36 \\ 0.54 \end{bmatrix}$$

Then calculate the cosine of the angle between them:

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

$$= (0.13 \cdot 0.73) + (0.38 \cdot -0.18) + (-0.88 \cdot 0.36) + (-0.25 \cdot 0.54)$$

$$= -0.44$$

Then:

$$\theta = \cos^{-1}(-0.44)$$

$$= 116 \text{ degrees.}$$

```
[28]: u = np.array([1.,3,-7,-2])
      print (u/np.sqrt(u.T.dot(u)))
      v = np.array([8.,-2,4,6])
      print(v/np.sqrt(v.T.dot(v)))
      print((v/np.sqrt(v.T.dot(v)).T).dot(u/np.sqrt(u.T.dot(u))))
      print(180*np.arccos((v/np.sqrt(v.T.dot(v)).T).dot(u/np.sqrt(u.T.dot(u))))/np.pi)
```

```
[ 0.12598816  0.37796447 -0.8819171  -0.25197632]
[ 0.73029674 -0.18257419  0.36514837  0.54772256]
-0.4370415209168243
115.91527033906208
```