Linear Algebra – Week 2

Lengths, Angles, and the Dot Product

This week we will learn about:

- The dot product,
- The length of vectors and the angle between them, and
- The Cauchy–Schwarz and triangle inequalities.

Extra reading and watching:

- Section 1.2 in the textbook
- Lecture videos 4, 5, 6, and 7 on YouTube
- Dot product at Wikipedia
- Cauchy–Schwarz inequality at Wikipedia

Extra textbook problems:

⋆ 1.2.1–1.2.3, 1.2.7, 1.2.8

⋆⋆ 1.2.4–1.2.6, 1.2.9–1.2.11

⋆⋆⋆ 1.2.12, 1.2.13, 1.2.17–1.2.21

☠ 1.2.23

The Dot Product

In 2D (and sometimes in 3D), it is fairly intuitive to talk about geometric quantities like lengths or angles. You have used things like similar triangles and the law of cosines for tackling problems like this in the past.

Using vectors, we can now generalize these concepts to arbitrary dimensions (even though we can't picture it)! Our main tool will be...

Definition 2.1 — Dot Product

If $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$ then the **dot product** of $\mathbf{v}$ and $\mathbf{w}$, denoted by $\mathbf{v} \cdot \mathbf{w}$, is the quantity

$$\mathbf{v} \cdot \mathbf{w} \stackrel{\text{def}}{=} v_1 w_1 + v_2 w_2 + \cdots + v_n w_n.$$

Please be wary of what types of objects go into and come out of the dot product:

Input: 2 vectors Output: a scalar

$\vec{v}/(\vec{w}\cdot\vec{x})$ makes sense, $(\vec{v}\cdot\vec{w})/\vec{x}$ does not

$\vec{v}\cdot\vec{w}\cdot\vec{x}$? No! Scalar · Vector = nonsense.

Intuitively, the dot product $\mathbf{v} \cdot \mathbf{w}$ tells you how much $\mathbf{v}$ points in the direction of $\mathbf{w}$ (or how much $\mathbf{w}$ points in the direction of $\mathbf{v}$).

Example. 2D examples.

$\vec{v} = (1,2)$

$\vec{w}_1 = (3,2)$:

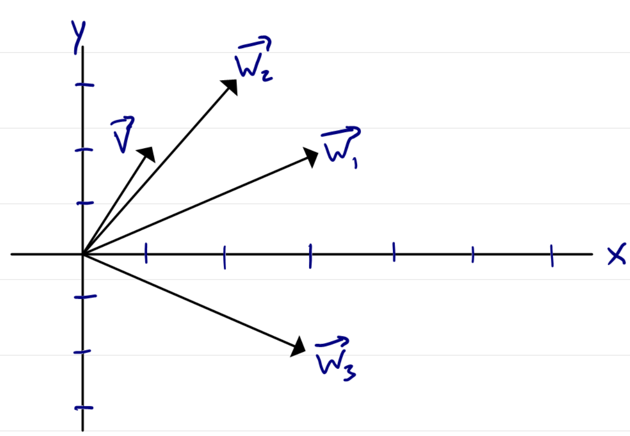
$\vec{v}\cdot\vec{w}_1 = 3+4 = 7$

$\vec{w}_2 = (2,3)$:

$\vec{v}\cdot\vec{w}_2 = 2+6 = 8$

$\vec{w}_3 = (3,-2)$:

$\vec{v}\cdot\vec{w}_3 = 3-4 = -1$



Example. Higher-dimensional examples.

$(3,2,1)\cdot(4,-3,5) = 12-6+5 = 11$

This works in higher dimensions too.

$(3,-2,0,4,1,2)\cdot(0,1,3,-2,-7,3)$
$= 0-2+0-8-7+6 = -11$

We have defined a new mathematical operation, so it's time for another "obvious" theorem telling us what properties it satisfies:

Theorem 2.1 — Properties of the Dot Product

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

a) $\mathbf{v}\cdot\mathbf{w} = \mathbf{w}\cdot\mathbf{v}$ (commutativity)

b) $\mathbf{v}\cdot(\mathbf{w}+\mathbf{z}) = \mathbf{v}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{z}$ (distributivity)

c) $(c\mathbf{v})\cdot\mathbf{w} = c(\mathbf{v}\cdot\mathbf{w})$

Proof. We will prove property (a). You can try the rest on your own (the method is quite similar).

Write $\vec{v} = (v_1, v_2, \ldots, v_n)$ and $\vec{w} = (w_1, w_2, \ldots, w_n)$.
Then

$$\vec{v}\cdot\vec{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n$$
$$= w_1v_1 + w_2v_2 + \cdots + w_nv_n = \vec{w}\cdot\vec{v}.$$

since multiplication on $\mathbb{R}$ is commutative

This completes the proof. ■

Example. Compute $\frac{1}{2}(-1,-3,2)\cdot(6,-4,2)$.

$$= \tfrac{1}{2}(-6+12+4) = \tfrac{1}{2}(10) = 5$$

OR

$$= (-1,-3,2)\cdot(3,-2,1) = -3+6+2 = 5.$$

Example. Show that $(\mathbf{v}+\mathbf{w})\cdot(\mathbf{v}+\mathbf{w}) = \mathbf{v}\cdot\mathbf{v} + 2\mathbf{v}\cdot\mathbf{w} + \mathbf{w}\cdot\mathbf{w}$ for all $\mathbf{v},\mathbf{w}\in\mathbb{R}^n$.

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

Theorem 2.1 (b) and (a)

Just like "FOIL"!

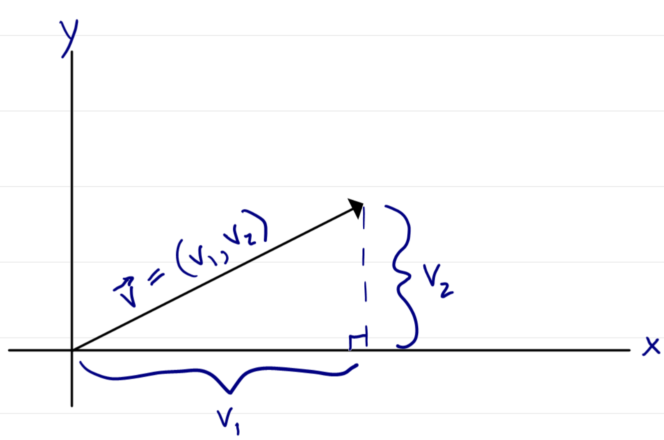
Length of a Vector

We now start making use of the dot product to talk about things like the length of vectors or the angle between vectors (even in high-dimensional spaces).

Example. Length of vectors in $\mathbb{R}^2$.

If $\vec{v} = (v_1, v_2)$ then Pythagoras tells us that

$$\underbrace{\|\vec{v}\|}_{\text{the length of } \vec{v}} = \sqrt{v_1^2 + v_2^2} = \sqrt{\vec{v}\cdot\vec{v}}$$

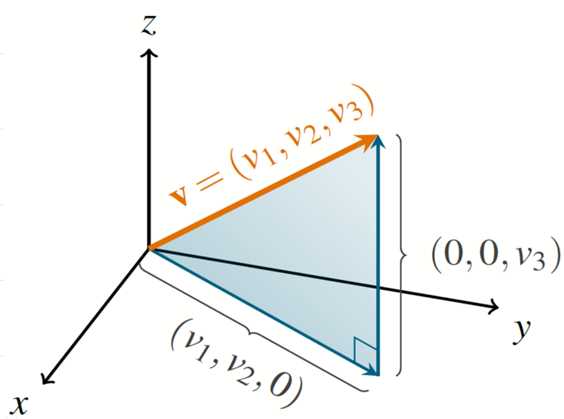


Example. Length of vectors in $\mathbb{R}^3$.

If $\vec{v} = (v_1, v_2, v_3)$ then we can write

$$\vec{v} = (v_1, v_2, 0) + (0, 0, v_3).$$

Pythagoras then says



$$\|\vec{v}\| = \sqrt{\|(v_1, v_2, 0)\|^2 + \|(0, 0, v_3)\|^2}$$

$$= \sqrt{\left(\sqrt{v_1^2 + v_2^2}\right)^2 + |v_3|^2} = \sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{\vec{v} \cdot \vec{v}}$$

In higher dimensions, we *define* the length of a vector so as to continue the pattern that we observed above:

Definition 2.2 — Length of a Vector

The **length** of a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$, denoted by $\|\mathbf{v}\|$, is defined by

$$\|\mathbf{v}\| \stackrel{\text{def}}{=} \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Example. Compute the length of some vectors.

$\vec{v} = (3, 4)$: $\|\vec{v}\| = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5$

$\vec{w} = (\cos(\theta), \sin(\theta))$: $\|\vec{w}\| = \sqrt{\cos^2(\theta) + \sin^2(\theta)} = \sqrt{1} = 1$

$\vec{x} = (5, 1, -4)$: $\|\vec{x}\| = \sqrt{25 + 1 + 16} = \sqrt{42}$

$\vec{y} = (1, 2, 0, -1, -3, 1)$: $\|\vec{y}\| = \sqrt{1 + 4 + 0 + 1 + 9 + 1} = \sqrt{16} = 4$

As always, we have defined a new mathematical object, so we want a theorem that tells us what its properties are.

Theorem 2.2 — Properties of Vector Length

Let $\mathbf{v} \in \mathbb{R}^n$ be a vector and let $c \in \mathbb{R}$ be a scalar. Then

a) $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$

b) $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$

Proof. To prove property (a), we just apply the relevant definitions:

$$\|c\vec{v}\| = \sqrt{(c\vec{v})\cdot(c\vec{v})} = \sqrt{c^2(\vec{v}\cdot\vec{v})}$$
$$= \sqrt{c^2}\sqrt{\vec{v}\cdot\vec{v}} = |c|\,\|\vec{v}\|.$$

To prove property (b), we have to prove two things:

① : "if $\vec{v} = \vec{0}$ then $\|\vec{v}\| = 0$" and ② : "if $\|\vec{v}\| = 0$ then $\vec{v} = \vec{0}$"

Proof of ①: $\|\vec{v}\| = \|\vec{0}\| = \sqrt{0^2 + 0^2 + \dots + 0^2} = 0.$

Proof of ②: $\|\vec{v}\| = 0 \Rightarrow \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} = 0 \Rightarrow v_1^2 + v_2^2 + \dots + v_n^2 = 0$

Each of $v_1^2, v_2^2, \dots, v_n^2$ is ≥ 0, so the only way for this sum to equal 0 is if $v_1^2 = v_2^2 = \dots = v_n^2 = 0$.

But then $v_1 = v_2 = \dots = v_n = 0$, so $\vec{v} = \vec{0}$.

This completes the proof. ■

A vector with length 1 is called a **unit vector**. Every non-zero vector $\mathbf{v} \in \mathbb{R}^n$ can be divided by its length to get a unit vector:

If $\vec{w} = \vec{v}/\|\vec{v}\|$ then

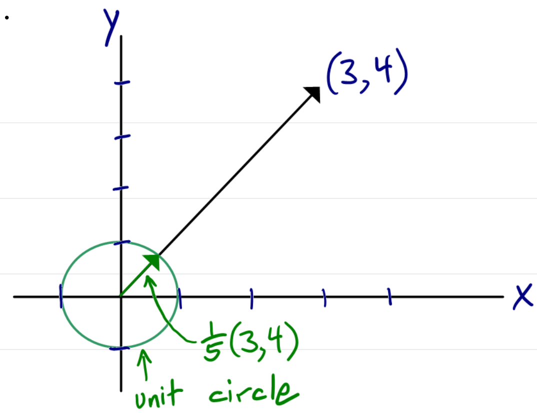
$$\|\vec{w}\| = \left\|\vec{v}/\|\vec{v}\|\right\| = \left\|\left(\frac{1}{\|\vec{v}\|}\right)\vec{v}\right\| = \frac{1}{\|\vec{v}\|}\|\vec{v}\| = 1.$$

Scaling $\mathbf{v}$ to have length 1 like this is called **normalizing v** (and this unit vector $\mathbf{w}$ is called the **normalization** of $\mathbf{v}$).

Example. Normalize the vector $(3,4) \in \mathbb{R}^2$.

We computed $\|(3,4)\| = 5$ earlier.

$\therefore$ The normalization of $(3,4)$ is $\frac{1}{5}(3,4)$.



Example. Show that the standard basis vectors are unit vectors.

Recall that $\vec{e}_j = (0,\ldots,0,\underset{\text{j-th entry}}{1},0,\ldots,0)$.

$\|\vec{e}_j\| = \sqrt{0^2+\cdots+0^2+1^2+0^2+\cdots+0^2} = \sqrt{1} = 1$. Done!

We now start to look at somewhat more interesting properties of the dot product and vector lengths. Our first result in this direction is an inequality that relates the dot product of two vectors to their lengths:

Theorem 2.3 — Cauchy–Schwarz Inequality

Suppose that $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are vectors. Then $|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\|\|\mathbf{w}\|$.

Proof. Define the vector $\mathbf{x} = \|\mathbf{w}\|\mathbf{v} - \|\mathbf{v}\|\mathbf{w}$ and then expand the quantity $\|\mathbf{x}\|^2$ in terms of the dot product:

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

Divide by 2 and rearrange to get

$\cancel{\|\vec{v}\|}\|\vec{w}\|(\vec{v}\cdot\vec{w}) \leq \|\vec{v}\|^{\cancel{2}}\|\vec{w}\|^{\cancel{2}}$, so

$\vec{v}\cdot\vec{w} \leq \|\vec{v}\|\,\|\vec{w}\|$. Done (sorta...)

Two technicalities:

- What if $\|\vec{v}\|=0$ or $\|\vec{w}\|=0$?
- To get $|\vec{v}\cdot\vec{w}|$, replace $\vec{w}$ by $-\vec{w}$. ■

The above theorem is our first example of a theorem with a very non-obvious proof: even though we can follow the steps and see that they are individually true, the choice of $\mathbf{x} = \|\mathbf{w}\|\mathbf{v} - \|\mathbf{v}\|\mathbf{w}$ at the start was something like magic. This particular choice of $\mathbf{x}$ was chosen so that the proof would give us what we wanted. Other choices of $\mathbf{x}$ also result in true inequalities, but ones that are less useful than Cauchy–Schwarz.

Example. Do there exist vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$ with...

$\|\vec{v}\|=2$, $\|\vec{w}\|=3$, and $\vec{v}\cdot\vec{w}=7$?

C-S inequality: $7 = |\vec{v}\cdot\vec{w}| \leq \|\vec{v}\|\,\|\vec{w}\| = 2\times 3 = 6$ No!

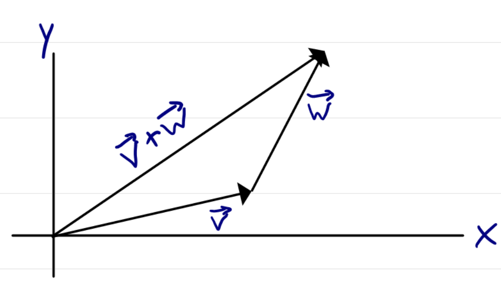
$\|\vec{v}\|=2$, $\|\vec{w}\|=3$, and $\vec{v}\cdot\vec{w}=5$?

$\vec{v}=(2,0)$, $\vec{w}=(5/2, \sqrt{11}/2)$ work.

We have two main uses for the Cauchy–Schwarz inequality. The first is that it helps us prove another geometrically "obvious" fact about vector lengths:

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

The shortest path between two points is a straight line.



Theorem 2.4 — Triangle Inequality

Suppose that $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are vectors. Then $\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$.

Proof. We start by expanding $\|\mathbf{v}+\mathbf{w}\|^2$ in terms of the dot product:

$$\begin{aligned}\|\vec{v}+\vec{w}\|^2 &= (\vec{v}+\vec{w})\cdot(\vec{v}+\vec{w})\\ &= \vec{v}\cdot\vec{v} + 2(\vec{v}\cdot\vec{w}) + \vec{w}\cdot\vec{w}\\ &= \|\vec{v}\|^2 + 2(\vec{v}\cdot\vec{w}) + \|\vec{w}\|^2\\ &\leq \|\vec{v}\|^2 + 2\|\vec{v}\|\|\vec{w}\| + \|\vec{w}\|^2 \qquad \text{(by Cauchy-Schwarz)}\\ &= (\|\vec{v}\| + \|\vec{w}\|)^2.\end{aligned}$$

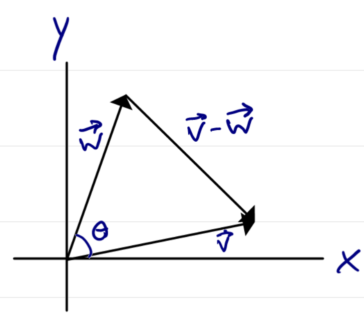
Square root both sides to get $\|\vec{v}+\vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$. ∎

Angle Between Vectors

The second immediate use of the Cauchy–Schwarz inequality is that it helps us define angles in $\mathbb{R}^n$. To get an idea of how this works, let's start by thinking about a triangle with sides given by the vectors $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{v}-\mathbf{w}$:

Angle θ between $\vec{v}$ and $\vec{w}$ can be found by Law of Cosines:

$$\|\vec{v}-\vec{w}\|^2 = \cancel{\|\vec{v}\|^2} + \cancel{\|\vec{w}\|^2} - 2\|\vec{v}\|\|\vec{w}\|\cos(\theta).$$



However, it is also true that

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

Setting these expressions equal and simplifying gives

$$\vec{v}\cdot\vec{w} = \|\vec{v}\|\|\vec{w}\|\cos(\theta), \quad \text{so} \quad \cos(\theta) = \frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\|\vec{w}\|}$$

Our reasoning above gave us a formula for the angle between two vectors in $\mathbb{R}^2$ (and in $\mathbb{R}^3$). We now state it as a definition in higher-dimensional spaces.

Definition 2.3 — Angle Between Vectors

The **angle** θ between two non-zero vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is the quantity

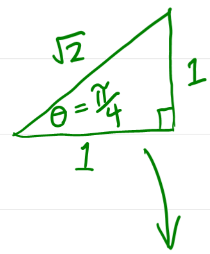
$$\theta = \arccos\left(\frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}\right).$$

$\cos^{-1}$

Example. What is the angle between $\mathbf{v} = (1, 1, 1, 1)$ and $\mathbf{w} = (2, 0, 2, 0)$?

$\vec{v} \cdot \vec{w} = 2 + 0 + 2 + 0 = 4$

Special triangle:



$\|\vec{v}\| = \sqrt{1^2 + 1^2 + 1^2 + 1^2} = \sqrt{4} = 2$

$\|\vec{w}\| = \sqrt{2^2 + 0^2 + 2^2 + 0^2} = \sqrt{8} = 2\sqrt{2}$

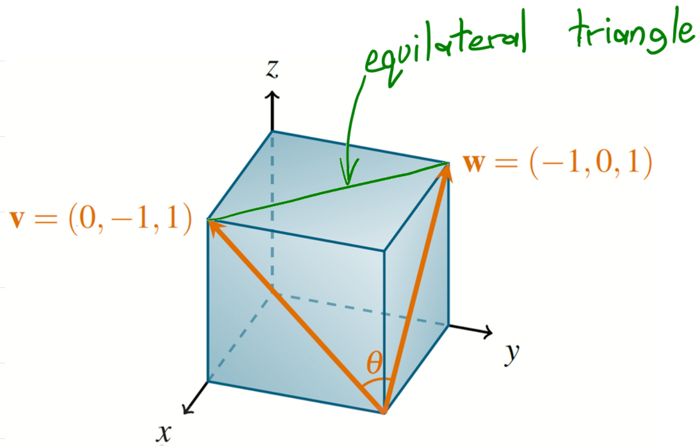
$\therefore\ \theta = \arccos\left(\frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\|\|\vec{w}\|}\right) = \arccos\left(\frac{4}{2(2\sqrt{2})}\right) = \arccos\left(\frac{1}{\sqrt{2}}\right) = \frac{\pi}{4}$

Example. What is the angle between the diagonals of two adjacent faces of a cube?

These vectors point from $(1,1,0)$ to $(1,0,1)$, and from $(1,1,0)$ to $(0,1,1)$, so they are

$\vec{v} = (1,0,1) - (1,1,0) = (0,-1,1)$ and

$\vec{w} = (0,1,1) - (1,1,0) = (-1,0,1)$.



Then

$\vec{v} \cdot \vec{w} = 0 + 0 + 1 = 1, \quad \|\vec{v}\| = \sqrt{0^2 + (-1)^2 + 1^2} = \sqrt{2}, \quad \|\vec{w}\| = \sqrt{(-1)^2 + 0^2 + 1^2} = \sqrt{2},$

so

$\theta = \arccos\left(\frac{1}{\sqrt{2} \cdot \sqrt{2}}\right) = \arccos\left(\frac{1}{2}\right) = \frac{\pi}{3}.$

Recall that $\arccos(x)$ is only defined if $-1 \leq x \leq 1$. How do we know that $-1 \leq \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} \leq 1$?

The Cauchy-Schwarz inequality!

One special case of vector angles that is worth pointing out is the case when $\mathbf{v} \cdot \mathbf{w} = 0$. When this happens...

$$\theta = \arccos\left(\frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\|\vec{w}\|}\right) = \arccos(0) = \frac{\pi}{2} \qquad (\text{i.e., } 90^\circ)$$

That is, $\vec{v}\cdot\vec{w} = 0$ means that $\vec{v}$ is perpendicular to $\vec{w}$.

This special case is important enough that we give it its own name:

Definition 2.4 — Orthogonality

Two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ are called **orthogonal** if $\mathbf{v} \cdot \mathbf{w} = 0$.

Example. Show that the vectors $(1, 1, -2)$ and $(3, 1, 2)$ are orthogonal.

$(1,1,-2)\cdot(3,1,2) = 3 + 1 - 4 = 0$ ✓

Example. Find a non-zero vector orthogonal to $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$

(Side note: $\vec{0}$ is orthogonal to everything!)

$\vec{w} = (v_2, -v_1)$ works:

$$\vec{v}\cdot\vec{w} = (v_1, v_2)\cdot(v_2, -v_1) = v_1v_2 - v_2v_1 = 0.$$