Dot product, length, and distance

We previously defined the length of a vector in $\mathbb{R}^2$ or $\mathbb{R}^3$.
We can generalize this to any vector in $\mathbb{R}^n$.

Def: If $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a vector in $\mathbb{R}^n$, the <u>length</u> of $\vec{x}$ is

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\vec{x} \cdot \vec{x}}.$$

A vector of length 1 is called a <u>unit vector</u>. If $\|\vec{x}\| \neq 0$, then $\frac{1}{\|\vec{x}\|}\vec{x}$ is a unit vector.

Just like with vectors in $\mathbb{R}^3$, we have the following properties:

- $\|\vec{x}\| \geq 0$, and $\|\vec{x}\| = 0$ if and only if $\vec{x} = \vec{0}$
- $\|a\vec{x}\| = |a|\,\|\vec{x}\|$ for all scalars a.

All the same properties of dot product hold as well:

- $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$
- $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$
- $(a\vec{x}) \cdot \vec{y} = a(\vec{x} \cdot \vec{y}) = \vec{x} \cdot (a\vec{y})$ for all scalars a.

Example: If $\|\vec{x}\| = 5$, $\|\vec{y}\| = 2$, and $x \cdot y = -1$, what is $\|\vec{x} + \vec{y}\|$?

First we calculate $\|\vec{x}+\vec{y}\|^2$:

$$\|\vec{x}+\vec{y}\|^2 = (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) = (\vec{x}+\vec{y})\cdot\vec{x} + (\vec{x}+\vec{y})\cdot\vec{y}$$
$$= \vec{x}\cdot\vec{x} + \vec{y}\cdot\vec{x} + \vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{y}$$
$$= \|\vec{x}\|^2 + 2\vec{x}\cdot\vec{y} + \|\vec{y}\|^2$$
$$= 5^2 + 2(-1) + 2^2 = 27.$$

So $\|\vec{x}+\vec{y}\| = \sqrt{27}$

Recall that if $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^3$, then

$\vec{u}\cdot\vec{v} = \|\vec{u}\|\,\|\vec{v}\|\cos\theta$.

Since $|\cos\theta| \le 1$ for any angle, this gives

$$|\vec{u}\cdot\vec{v}| \le \|\vec{u}\|\|\vec{v}\|.$$

This inequality holds more generally for vectors in $\mathbb{R}^n$:

<u>Cauchy Inequality</u>: If $\vec{x}, \vec{y}$ are vectors in $\mathbb{R}^n$, then

$$|\vec{x}\cdot\vec{y}| \le \|\vec{x}\|\,\|\vec{y}\|.$$

Moreover $|\vec{x}\cdot\vec{y}| = \|\vec{x}\|\,\|\vec{y}\|$ if and only if one is a scalar multiple of the other.

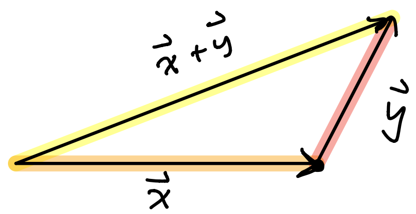
We can combine the Cauchy inequality with the equality from the above example to get:

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

So $\|\vec{x}+\vec{y}\|^2 \le (\|\vec{x}\| + \|\vec{y}\|)^2$. Taking square roots gives us:

The triangle inequality: If $\vec{x}$ and $\vec{y}$ are vectors in $\mathbb{R}^n$, then

$$\|\vec{x}+\vec{y}\| \le \|\vec{x}\| + \|\vec{y}\|$$



Geometrically, the length of one side of a triangle is shorter than the sum of the lengths of the other two sides.

We defined the distance between two vectors in $\mathbb{R}^3$ to be the length of their difference. We have the same def in $\mathbb{R}^n$:

Def: If $\vec{x}$ and $\vec{y}$ are vectors in $\mathbb{R}^n$, then the distance $d(\vec{x},\vec{y})$ between $\vec{x}$ and $\vec{y}$ is

$$d(\vec{x},\vec{y}) = \|\vec{x}-\vec{y}\|.$$

All of the intuitive properties hold:

Properties of $d(\vec{x},\vec{y})$: If $\vec{x}, \vec{y}, \vec{z}$ are vectors in $\mathbb{R}^n$, we have

1.) $d(\vec{x},\vec{y}) \ge 0$.

2.) $d(\vec{x},\vec{y}) = 0 \iff \vec{x} = \vec{y}$.

3) $d(\vec{x},\vec{y}) = d(\vec{y},\vec{x})$.

4.) $d(\vec{x},\vec{z}) \leq d(\vec{x},\vec{y}) + d(\vec{y},\vec{z})$.

Property 4.) is just the triangle inequality:

$$d(\vec{x},\vec{z}) = \|\vec{x}-\vec{z}\| = \|(\vec{x}-\vec{y}) + (\vec{y}-\vec{z})\| \leq \|\vec{x}-\vec{y}\| + \|\vec{y}-\vec{z}\| = d(\vec{x},\vec{y}) + d(\vec{y},\vec{z})$$

Orthogonal sets

We can also define orthogonality for vectors in $\mathbb{R}^n$ using the dot product:

Def:

1.) If $\vec{x}$ and $\vec{y}$ are in $\mathbb{R}^n$, they are orthogonal if $\vec{x} \cdot \vec{y} = 0$.

2.) More generally a set $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\}$ of nonzero vectors is an orthogonal set if every pair is orthogonal i.e. $\vec{x}_i \cdot \vec{x}_j = 0$ for every i and j.

3.) $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\}$ is orthonormal if it is orthogonal and each $\vec{x}_i$ is a unit vector, i.e. has length 1.

Ex: The standard basis is orthonormal.

Ex: If $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\}$ is an orthogonal set of vectors, then

$$\left\{ \frac{\vec{x}_1}{\|\vec{x}_1\|}, \frac{\vec{x}_2}{\|\vec{x}_2\|}, \dots, \frac{\vec{x}_k}{\|\vec{x}_k\|} \right\}$$

is orthonormal. This is called <u>normalizing</u> an orthogonal set.

Ex: If $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$, then

$\vec{v}_1 \cdot \vec{v}_2 = 0$, $\vec{v}_1 \cdot \vec{v}_3 = 0$, $\vec{v}_2 \cdot \vec{v}_3 = 0$, so $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is orthogonal.

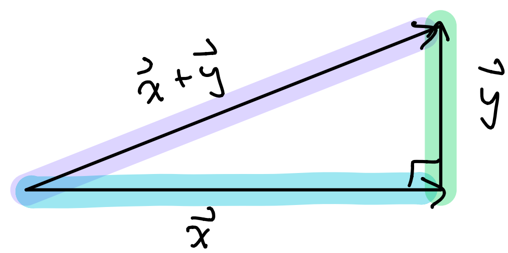
$\|\vec{v}_1\| = 2$, $\|\vec{v}_2\| = \sqrt{2}$, $\|\vec{v}_3\| = \sqrt{2}$, so to normalize, we take

$$\left\{ \frac{1}{2}\vec{v}_1, \frac{1}{\sqrt{2}}\vec{v}_2, \frac{1}{\sqrt{2}}\vec{v}_3 \right\},$$ which is orthonormal.

If we know vectors are orthogonal, what can we say about the length of their sum? In $\mathbb{R}^2$:

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

for $\vec{x}$ and $\vec{y}$ orthogonal, by the Pythagorean Theorem.



This generalizes to an orthogonal set in $\mathbb{R}^n$:

Pythagoras' Theorem: If $\{\vec{x}_1, \ldots, \vec{x}_k\}$ is an orthogonal set in $\mathbb{R}^n$, then

$$\|\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k\|^2 = \|\vec{x}_1\|^2 + \|\vec{x}_2\|^2 + \cdots + \|\vec{x}_k\|^2$$

Why?

$$\|\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k\|^2 = (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \cdot (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k)$$

$$= (\vec{x}_1 \cdot \vec{x}_1) + (\vec{x}_2 \cdot \vec{x}_2) + \cdots + (\vec{x}_k \cdot \vec{x}_k) + \sum_{i \neq j} \underbrace{(\vec{x}_i \cdot \vec{x}_j)}_{= 0}$$

$$= \|\vec{x}_1\|^2 + \cdots + \|\vec{x}_k\|^2.$$

Theorem: Every orthogonal set in $\mathbb{R}^n$ is linearly independent.

Thus, if U is the span of an orthogonal set, then the orthogonal set forms a basis for U, called an <u>orthogonal basis</u>.

The nice thing about orthogonal bases is that if we want to write a vector as a linear combination of the basis vectors, there is an explicit formula for the coefficients:

Expansion Theorem: If $\{\vec{x}_1, \ldots, \vec{x}_m\}$ is an orthogonal basis for U, a subspace of $\mathbb{R}^n$, and $\vec{v}$ is any vector in U, Then

$$\vec{v} = \left(\frac{\vec{v}\cdot\vec{x}_1}{\|\vec{x}_1\|^2}\right)\vec{x}_1 + \dots + \left(\frac{\vec{v}\cdot\vec{x}_m}{\|\vec{x}_m\|^2}\right)\vec{x}_m$$

Notice that this looks like the sum of the projections of $\vec{v}$ onto each $\vec{x}_i$.

This is called the <u>Fourier expansion</u> of $\vec{v}$, and the coefficients are the <u>Fourier coefficients</u>.

<u>Ex</u>:

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$$

is an orthogonal basis for $\mathbb{R}^3$ since

$$\vec{x}_1\cdot\vec{x}_2 = 0, \quad \vec{x}_1\cdot\vec{x}_3 = 0, \text{ and } \vec{x}_2\cdot\vec{x}_3 = 0.$$

In order to write $\vec{v} = \begin{bmatrix} 6 \\ 1 \\ -2 \end{bmatrix}$ as a linear comb. of $\vec{x}_1, \vec{x}_2, \vec{x}_3$, we first find the Fourier coefficients:

$$\frac{\vec{v}\cdot\vec{x}_1}{\|\vec{x}_1\|^2} = \frac{6}{6} = 1$$

$$\frac{\vec{v}\cdot\vec{x}_2}{\|\vec{x}_2\|^2} = \frac{-11}{5}$$

$$\frac{\vec{v}\cdot\vec{x}_3}{\|\vec{x}_3\|^2} = \frac{18}{30} = \frac{3}{5}$$

so $\vec{v} = \vec{x}_1 - \frac{11}{5}\vec{x}_2 + \frac{3}{5}\vec{x}_3$.

Practice Problems: 5.3 : 1, 2, 3bc, 4, 5b, 6