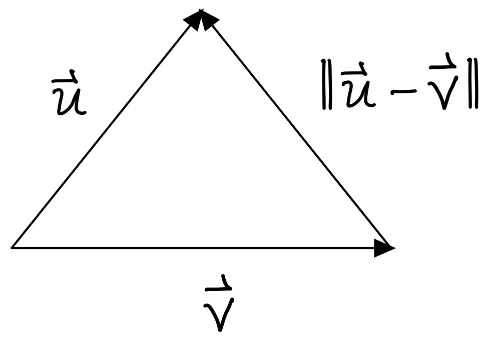
Orthogonality

def: $\text{dist}(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\|$

where:



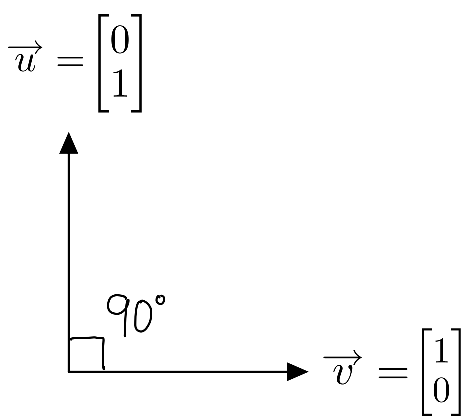
** note that this is one of many ways to measure distance, see ML notes **

★ **def:** $\vec{u}$ and $\vec{v}$ are **orthogonal** *iff* $\langle \vec{u},\vec{v}\rangle = 0 \equiv$ $\vec{u}$ and $\vec{v}$ are **orthogonal** *iff* $\underbrace{\|\vec{u},\vec{v}\|^2 = \|\vec{u}\|^2+\|\vec{v}\|^2}$

! can only happen when cosine similarity = 0 !

- "orthogonal" generalizes the notion of perpendicular to higher dimensions:

$$\langle \vec{u}, \vec{v} \rangle = 0 \Leftrightarrow \vec{u} \perp \vec{v}$$



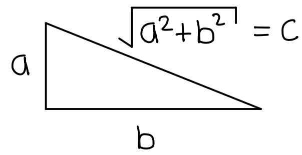
- note that the zero vector $\vec{0}$ is the only one **orthogonal to all other vectors in $\mathbb{R}^n$**

★ • orthogonal $\Rightarrow$ linearly independent, but linearly independent $\not\Rightarrow$ orthogonal

$\longrightarrow$ independence is a very convenient property of orthogonality

Recall: the pythagorean theorem, the law of cosines, and cosine similarity

- PATHAGOREAN THM: $a^2 + b^2 = c^2$

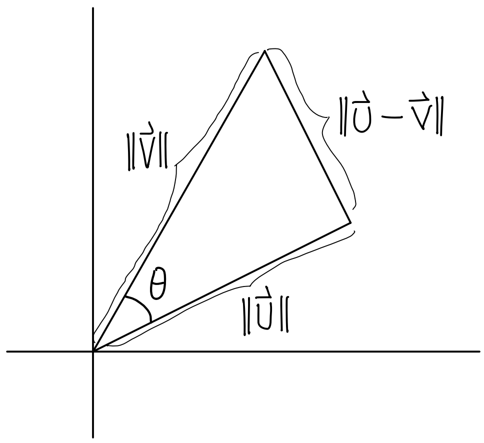


but the Pythagorean Theorem only works on <u>right</u> triangles

- for non-right triangles, we need the **Law of Cosines**, which says:

$$c^2 = a^2 + b^2 - 2ab\cos\theta \quad \equiv \quad \|\vec{u}-\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2 * \|\vec{u}\| * \|\vec{v}\| * \cos\theta$$

where:



★ • The derivation for cosine similarity is:

def. 0 when $\vec{u}$ and $\vec{v}$ are **orthogonal**

$$\cos\theta = \frac{\|\vec{u}-\vec{v}\|^2 - \|\vec{u}\|^2 + \|\vec{v}\|^2}{-2 * \|\vec{u}\| * \|\vec{v}\|} = \frac{(-1)\,\|\vec{u}-\vec{v}\|^2 - \|\vec{u}\|^2 + \|\vec{v}\|^2}{(-1)\ \ -2 * \|\vec{u}\| * \|\vec{v}\|}$$

!!!

$$= \frac{\|\vec{u}-\vec{v}\|^2 + \|\vec{u}\|^2 - \|\vec{v}\|^2}{2 * \|\vec{u}\| * \|\vec{v}\|} = \frac{2\,\vec{u}^T v}{2 * \|\vec{u}\| * \|\vec{v}\|} = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| * \|\vec{v}\|} \propto \rho_{\vec{u},\vec{v}}$$

ex) $\vec{a} = \begin{array}{cc} \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix} & \begin{array}{c} x \\ y \\ z \end{array} \end{array} \perp \vec{b} = \begin{array}{cc} \begin{bmatrix} -5 \\ 5 \\ 7 \end{bmatrix} & \begin{array}{c} x \\ y \\ z \end{array} \end{array} \longrightarrow \langle \vec{a}\ \vec{b} \rangle = \begin{bmatrix} 5 & 5 & 0 \end{bmatrix} \begin{bmatrix} -5 \\ 5 \\ 7 \end{bmatrix} = (5)(-5) + (5)(5) + (0)(7) = 0$

so, $\vec{a}$ and $\vec{b}$ are "orthogonal" to each other

def: Let W be a subspace of $\mathbb{R}^n$, then $W^{\perp} = \{\vec{x} \in \mathbb{R}^n \mid \langle \vec{x}, \vec{w} \rangle = 0 \ \forall \ \vec{w} \in W\}$.
and W is called the **perpendicular space**

★**Thm:** If $S = \{\vec{u}_1, \dots, \vec{u}_k\}$ is an orthogonal set of NON-ZERO vectors in $\mathbb{R}^n$, then S is a linnearly independent set and hence forms a **basis** for span$\{S\}$

pf: let

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_k\vec{u}_k = \vec{0}$$

then

"dot"

$$[c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_k\vec{u}_k] \cdot u_1$$

$$= \vec{u}_1 \vec{0} = 0 \qquad (1)$$

any vector times $\vec{0}$ equals zero

$$= \underbrace{c_1 \langle \vec{u}_1, \vec{u}_1 \rangle}_{\|\vec{u}_1\|^2} + \underbrace{c_2 \langle \vec{u}_1, \vec{u}_2 \rangle}_{0} + \underbrace{\cdots}_{0} + \underbrace{c_k \langle \vec{u}_1, \vec{u}_k \rangle}_{0} \qquad (2)$$

goes to zero by the definition of **orthogonality**

$\therefore$ either $c_1 = 0$, $\vec{u}_1 = \vec{0}$ to force $\|\vec{u}_1\|^2 = 0$, or both

BUT WE SAID THAT ALL $\vec{u}_i$'s are nonzero, so $\vec{u}_1 \neq \vec{0}$

$\therefore \quad \vec{u}_1 \vec{0} = 0 \iff c_1 = 0$

$\boxed{\circlearrowright \text{ pf. for each and every } u_i \text{ to see that this holds } \forall \vec{u}_i \in S}$

$\therefore$ all vectors $u_i \in S$ are linearly independent b/c all c_i's must be 0

★**def:** an **orthogonal basis** for the **perpendicular space** is a basis that is also an orthogonal set spanning W

Que: why does this make orthogonal bases superior to standard bases?

*Thm: let $\{\vec{u}_1, \cdots, \vec{u}_k\}$ be an orthogonal basis for subspace W of $\mathbb{R}^n$. Unlike a standard basis for the subspace, $\{\vec{u}_1, \cdots, \vec{u}_k\}$ being an orthogonal basis, then we automatically know the values for every coef. in the linear combination: $\vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_k\vec{u}_k$ instead of requiring gaussian ellim. to find them.

$$c_i \overset{\text{STANDARD BASIS}}{=\!=\!=\!=\!=} \; ?$$

*We just don't know the values for C; w/o performing a painstaking gaussian ellim. process (involing an augmented matrix [Ac=u] and row reduction, which is computationally expensive!) *

$$c_i \overset{\text{ORTHONORMAL BASIS}}{=\!=\!=\!=\!=} \frac{\langle y, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle} = \frac{\langle y, \vec{u}_i \rangle}{\|\vec{u}_i\|^2}$$

* still have to iteratively compute the Ci's (one at a time), but now have a much faster way to do it *

$$c_i \overset{\text{ORTHOGONAL BASIS}}{=\!=\!=\!=\!=} \frac{\langle y, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle}$$

but w/ even more computational efficiency than orthogonality...

WOOHOO!!!

! * see section 8 of notes (Fourier Expansion) to understand why everyone goes bonkers for orthonormality * !