Orthogonality

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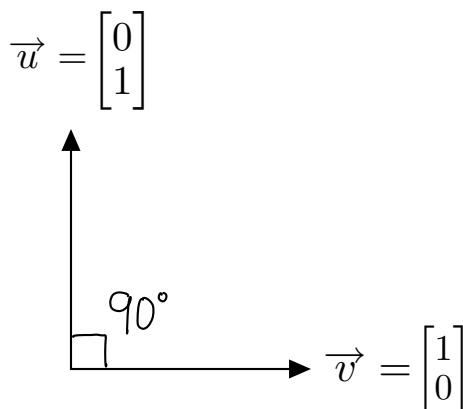
where:

$$\vec{u} \qquad \|\vec{u} - \vec{v}\|$$

$$\vec{v}$$
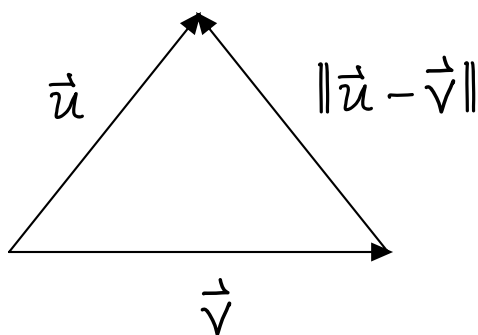
note that this is one of many ways to measure distance, see ML notes

*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* $\|\vec{u}, \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$

⚠ can only happen when
● cosine similarity = 0 ⚠

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

$$\langle \vec{u}, \vec{v} \rangle = 0 \Leftrightarrow \vec{u} \perp \vec{v}$$
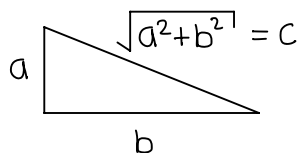
$$\vec{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$90°$$

$$\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- 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

→ independence is a very convenient property of orthogonality

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

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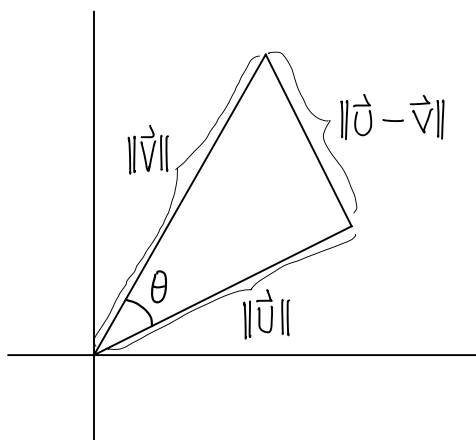

$\sqrt{a^2+b^2} = c$

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:

$\star$
- The derivation for **cosine similarity** is:

$\underset{=}{\text{def.}}$ 0 when $\vec{U}$ and $\vec{V}$ are **orthogonal**

$$\cos\theta = \frac{\|\vec{U}-\vec{V}\|^2 - \|\vec{U}\|^2 + \|\vec{V}\|^2}{-2*\|\vec{U}\|*\|\vec{V}\|} = \frac{(-1)\|\vec{U}-\vec{V}\|^2 - \|\vec{U}\|^2 + \|\vec{V}\|^2}{(-1) - 2*\|\vec{U}\|*\|\vec{V}\|}$$

!!!

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

ex) $\vec{a} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}\begin{matrix} X \\ Y \\ Z \end{matrix} \perp \vec{b} = \begin{bmatrix} -5 \\ 5 \\ 7 \end{bmatrix}\begin{matrix} X \\ Y \\ Z \end{matrix} \longrightarrow \langle\vec{a}\ \vec{b}\rangle = \begin{bmatrix} 5 & 5 & 0 \end{bmatrix}\begin{bmatrix} -5 \\ 5 \\ 7 \end{bmatrix} = (5)(-5) + (5)(5) + (0)(7) = 0$

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

"dot"

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

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

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

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

$\underbrace{\qquad}_{\|\vec{u}_1\|^2}$ $\quad 0 \qquad 0 \qquad 0$

$\underbrace{\qquad\qquad\qquad\qquad}$
goes to zero by the definition
of **orthogonality**

∴ 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 $\vec{u}_1 \neq \vec{0}$

∴ $\vec{u}_1 \vec{0} = 0 \iff c_1 = 0$

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

∴ 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 itself.

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

$c_i \underline{\overline{\text{STANDARD BASIS}}}$?

* 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!) *

$c_i \underline{\overline{\text{ORTHONORMAL BASIS}}} \dfrac{\langle y, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle} = \dfrac{\langle y, \vec{u}_i \rangle}{\|\vec{u}_i\|^2}$

☺

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

$c_i \underline{\overline{\text{ORTHOGONAL BASIS}}} \dfrac{\langle y, \vec{u}_i \rangle}{\langle \vec{u}_i, \vec{u}_i \rangle}$ but w/ even more computational efficiency than orthogonality...

WOOHOO!!!

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