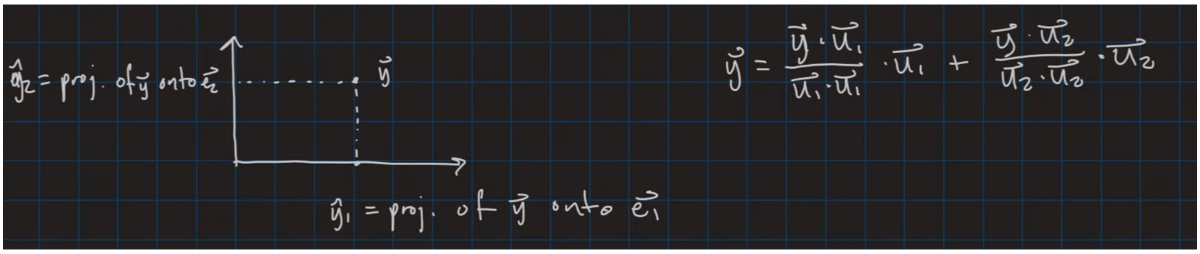
Last time we left off with an example of orthogonal projections.




Here, we're basically breaking down the vector $\vec{y}$ into a linear combination of projections and the basis vectors. This is true as long as the basis vectors are orthogonal.

# 1 Orthonormal Set

An orthonormal set is a special case of an orthogonal space. We say $\{\vec{u_1}, \cdots, \vec{u_n}\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors. We can take any orthogonal set and apply the norm to make every vector a unit vector.

If $W$ is the subspace spanned by such a set, then it is called an orthonormal basis.

**Theorem 1** *An $m \times n$ matrix $U$ has orthonormal columns if, and only if, $U^T U = I$*

**Theorem 2** *Let $U \in \mathbb{R}^{m\times n}$ with orthonormal columns, and $\vec{x}$ and $\vec{y} \in \mathbb{R}^n$. Then*

*(a)* $||U\vec{x}|| = ||\vec{x}||$

*(b)* $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$

*(c)* $(U\vec{x}) \cdot (U\vec{y}) = 0$ *if, and only if,* $\vec{x} \cdot \vec{y} = 0$

(a) and (c) state that the linear mapping $\vec{x} \mapsto U\vec{x}$ preserves both length and direction.

An **orthogonal matrix** is a square invertible matrix $U$ such that $U^{-1} = U^T$.

# 2 Orthogonal Projections

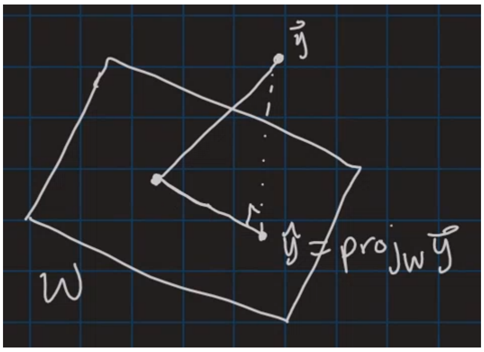
In the previous section, we talked about an orthogonal projection

$$\vec{y} = \hat{y} + \vec{z}$$

where $\hat{y}$ is a scaled basis vector and $\vec{z}$ is a vector orthogonal to $\hat{y}$.

There is a key property: the relationship between $\vec{y}$ and $\hat{y}$. Given $\vec{y}$ and a subspace $W \in \mathbb{R}^n$, there exists $\hat{y} \in W$ such that

1. $\hat{y}$ is the unique vector in $W$ for which $\vec{y} - \hat{y}$ is orthogonal to $W$.
2. $\hat{y}$ is the unique vector in $W$ closest to $\vec{y}$.




## Example

Let $\{\vec{u_1}, \cdots, \vec{u_n}\}$ be an orthogonal basis for $\mathbb{R}^5$ and let

$$\vec{y} = c_1\vec{u_1} + \cdots + c_5\vec{u_5}$$

Consider $W = \text{Span}\{\vec{u_1}, \vec{u_2}\}$ and write $\vec{y}$ as a linear combination of coordinates and basis vectors.

Remember that $\vec{y} \in \mathbb{R}^5$. We can write out $\vec{y}$ as a linear combination and take a vector $\vec{z}$ from it

$$\vec{y} = \underbrace{c_1\vec{u_1} + c_2\vec{u_2}}_{\vec{z_1}} + \underbrace{c_3\vec{u_3} + c_4\vec{u_4} + c_5\vec{u_5}}_{\vec{z_2}}$$

$$\vec{z_1} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \qquad \vec{z_2} = \begin{bmatrix} c_3 \\ c_4 \\ c_5 \end{bmatrix}$$

We know right off the bat that $\vec{z_1}$ is in $W$. What we need to show is that $\vec{z_2} \in W^\perp$. We can show this by showing that the two $z$ vectors are orthogonal to each other.

$$\vec{z_2} \cdot \vec{u_1} = (c_3\vec{u_3} + c_4\vec{u_4} + c_5\vec{u_5}) \cdot \vec{u_1}$$

$$= c_3\vec{u_3} \cdot \vec{u_1} + c_4\vec{u_4} \cdot \vec{u_1} + c_5\vec{u_5} \cdot \vec{u_1} = 0$$

Because $\vec{u_1}$ is orthogonal to $\vec{u_{3 \ldots 5}}$, the dot product will be 0. We can do the same for $\vec{z_2} \cdot \vec{u_2} = 0$.

What we just did is break up our vector $\vec{y}$ into two vectors: one in $W$ and one orthogonal to that, in $W^\perp$.

**Theorem 3** *Let $W$ be a subspace of $\mathbb{R}^n$. Then each $\vec{y} \in \mathbb{R}^n$ can be written uniquely as*

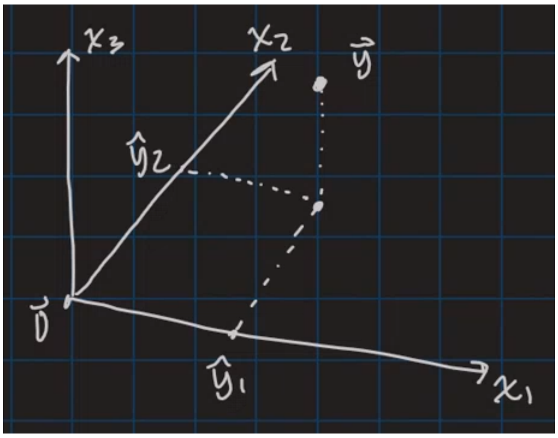
$$\vec{y} = \hat{y} + \vec{z}$$

*where $\vec{y} \in W$ and $\vec{z} \in W^{\perp}$.*

*If $\{\vec{u_1}, \cdots, \vec{u_n}\}$ is any orthogonal basis of $W$, then*

$$\vec{y} = \frac{\vec{y} \cdot \vec{u_1}}{\vec{u_1} \cdot \vec{u_1}}\vec{u_1} + \cdots + \frac{\vec{y} \cdot \vec{u_p}}{\vec{u_p} \cdot \vec{u_p}}\vec{u_p}$$

*and $\vec{z} = \vec{y} - \hat{y}$.*




Here's a visualization of a vector $\hat{y}$ (the dot below $\vec{y}$ on the $x_1$–$x_2$ plane) as a linear combination of $\hat{y_1}$ and $\hat{y_2}$, which are scaled basis vectors ($\vec{x_1}$ and $\vec{x_2}$).

$$\hat{y} = \frac{\vec{y} \cdot \vec{u_1}}{\vec{u_1} \cdot \vec{u_1}}\vec{u_1} + \frac{\vec{y} \cdot \vec{u_2}}{\vec{u_2} \cdot \vec{u_2}}\vec{u_2}$$

$$\hat{y} = \hat{y_1} + \hat{y_2}$$

We're taking the vector $\vec{y}$ and projecting it down onto the plane $W = \text{Span}\{\vec{x_1}, \vec{x_2}\}$.

The vector from $\hat{y}$ to $\vec{y}$ (directly up, orthogonal to $W$) is the $\vec{z}$ vector. It's length is

$$||\vec{z}|| = ||\vec{y} - \hat{y}||$$

## Properties of Orthogonal Projections

If $\vec{y} \in W = \text{Span}\{\vec{u_1}, \cdots, \vec{u_p}\}$, then $proj_W\vec{y} = \vec{y}$. This means that the vector $\vec{y}$ is already in the subspace $W$.

**Theorem 4 (Best Approximation Theorem)** *Let $W$ be a subspace of $\mathbb{R}^n$, $\vec{y}$ can be any vector in $\mathbb{R}^n$, and let $\hat{y}$ be the orthogonal projection of $\vec{y}$ onto $W$*

$$\hat{y} = proj_W\vec{y}$$

*Then $\hat{y}$ is the closest point in $W$ to $\vec{y}$, in the sense that*

$$||\vec{y} - \hat{y}|| < ||\vec{y} - \vec{v}||$$

*for all $\vec{v} \in W$ distinct from $\vec{y}$.*

## Example

Given $W = \text{Span}\{\vec{u}, \cdots, \vec{u_p}\}$ and a vector $\vec{y}$, find the closest point in $W$ to $\vec{y}$.

The closest point is the projection of $\vec{y}$ onto $W$

$$\hat{y} = proj_W \vec{y}$$

What if we want to find the distance between $\vec{y}$ and the subspace $W$?

First we could calculate $\hat{y}$ like above, then take the norm

$$||\vec{y} - \hat{y}||$$

If $S$ is an orthonormal set, $proj_S \vec{y}$ becomes simpler.

**Theorem 5** *If $\{\vec{u_1}, \cdots, \vec{u_n}\}$ is an orthonormal basis for a subspace $W \in \mathbb{R}^n$, then*

$$proj_W \vec{y} = (\vec{y} \cdot \vec{u_1})\vec{u_1} + \cdots + (\vec{y} \cdot \vec{u_p})\vec{u_p}$$

*(Notice where there is a dot product and where there is multiplication)*

*If $U = \begin{bmatrix} \vec{u_1} & \cdots & \vec{u_p} \end{bmatrix}$, then*

$$proj_W \vec{y} = UU^T \vec{y} \qquad \text{for all } \vec{y} \in \mathbb{R}^n$$

---

If calculating by hand, it's generally more convenient to use the equation

$$\vec{y} = \sum_i \frac{\vec{y} \cdot \vec{u_i}}{\vec{u_i} \cdot \vec{u_i}} \vec{u_i}$$