Last time:

- an orthonormal basis is a basis consisting of unit vectors which are pairwise orthogonal.
- If a square matrix $U$ has orthonormal columns, then $U$ is invertible and $U^{-1} = U^T$.

  Such a matrix is called an orthogonal matrix.

Today: understanding orthogonal projection.

Recall that if $u \neq 0$ and $v$ are vectors in $\mathbb{R}^n$,

$$\text{proj}_u v = \frac{v \cdot u}{u \cdot u} u$$

is the orthogonal projection of $v$ onto $u$.

We can write any vector $v$ uniquely! as the sum of

- a vector parallel to $u$, and
- a vector orthogonal to $u$.

$\text{proj}_u v$ is the summand parallel to $u$.

Today: we'll define $\text{proj}_W v$, the projection of $v$ onto an arbitrary subspace $W$.

# Example

Suppose $\{u_1, u_2, u_3\}$ is an orthogonal basis for $\mathbb{R}^3$, and let $W = \text{span}\{u_1, u_2\}$.

Write $y = \hat{y} + z$, for some $\hat{y} \in W$, and $z \in W^\perp$.

Recall for any orthogonal basis we have

$$y = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 + \frac{y \cdot u_3}{u_3 \cdot u_3} u_3.$$

It follows that we should take

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

and $$z = \frac{y \cdot u_3}{u_3 \cdot u_3} u_3.$$

Since $u_3$ is orthogonal to both $u_1$ and $u_2$, $z$ is orthogonal to $W$, as desired.

Clearly $\hat{y} \in W$.

# Orthogonal decomposition theorem

Let $W$ be a subspace of $\mathbb{R}^n$.

Each $y \in \mathbb{R}^n$ can be written uniquely as

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

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

If $\{u_1, \ldots, u_p\}$ is an orthogonal basis of $W$, then

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

The vector $\hat{y}$ is called the orthogonal projection of $y$ onto $W$.

## Example

Given

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

and $W = \text{span}\{u_1, u_2, u_3\}$, write $y = \begin{bmatrix} 2 \\ -3 \\ 4 \\ 1 \end{bmatrix}$ as the sum of a vector in $W$ and a vector in $W^\perp$.

Observe that $\{u_1, u_2, u_3\}$ is an orthogonal basis for $W$.

The orthogonal projection of $y$ onto $W$ is

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{y \cdot u_2}{u_2 \cdot u_2} u_2 + \frac{y \cdot u_3}{u_3 \cdot u_3} u_3$$

$$= \frac{-2}{3}\begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} + \frac{7}{3}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \frac{6}{3}\begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{5}{3} \\ -\frac{8}{3} \\ \frac{13}{3} \\ 1 \end{bmatrix}$$

Also $z = y - \hat{y} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ 0 \end{bmatrix}.$

We could do this problem because we were given an orthogonal basis.

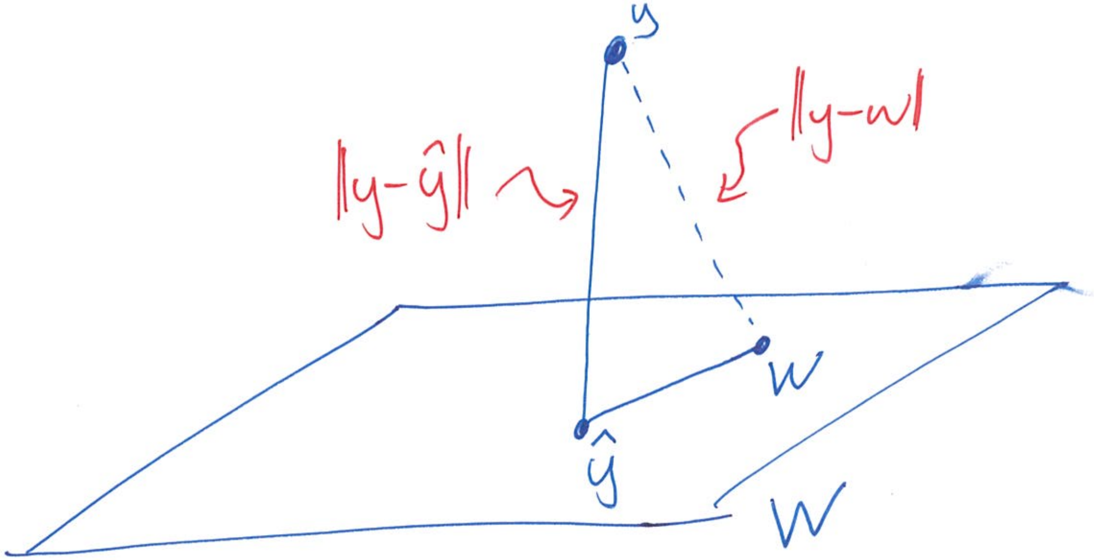
What would we do otherwise?

# The best approximation theorem

Let $W$ be a subspace of $\mathbb{R}^n$, and $y$ any vector in $\mathbb{R}^n$. Then $\hat{y}$, the orthogonal projection of $y$ onto $W$, is the closest vector in $W$ to $y$, that is:

$$\|y-\hat{y}\| < \|y-w\|$$

for every $w \in W$ with $w \neq \hat{y}$.



<u>Proof</u> Let $w \in W$, $w \neq \hat{y}$. Then $\hat{y} - w \in W$.

By definition $y - \hat{y}$ is orthogonal to $W$, and so orthogonal to $\hat{y} - w$.

We can write $y - w = (y - \hat{y}) - (\hat{y} - w)$

and by Pythagoras

$$\|y - w\|^2 = \|y - \hat{y}\|^2 + \|\hat{y} - w\|^2.$$

However $\|\hat{y} - w\|^2 \geq 0$ (with equality only if $\hat{y} = w$),

so $\|y - w\|^2 > \|y - \hat{y}\|^2$.

$\square$

<u>Definition</u> the distance from a vector $y$ to a subspace $W$ is this least distance, i.e. $\|y - \hat{y}\|$, where $\hat{y}$ is the projection of $y$ onto $W$.

## Example

Let $y=\begin{bmatrix}3\\-1\\1\\13\end{bmatrix}$, $u_1=\begin{bmatrix}1\\-2\\-1\\2\end{bmatrix}$, $u_2=\begin{bmatrix}4\\1\\0\\3\end{bmatrix}$.

Find the distance from $y$ to $\text{span}\{u_1, u_2\}$.

The projection $\hat{y}$ of $y$ onto $\text{span}\{u_1, u_2\}$ is

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

$$= \frac{30}{10}\begin{bmatrix}1\\-2\\-1\\2\end{bmatrix} + \frac{26}{26}\begin{bmatrix}-4\\1\\0\\3\end{bmatrix} = \begin{bmatrix}-1\\-5\\-3\\9\end{bmatrix}.$$

The distance is thus

$$\|y-\hat{y}\| = \left\|\begin{bmatrix}4\\4\\4\\4\end{bmatrix}\right\| = \sqrt{4^2+4^2+4^2+4^2} = \sqrt{64} = 8.$$

<u>Theorem</u> if $\{u_1, \dots, u_k\}$ is an orthonormal basis for a subspace $W$ of $\mathbb{R}^n$, then

$$\text{proj}_W y = (y \cdot u_1)u_1 + (y \cdot u_2)u_2 + \dots + (y \cdot u_k)u_k.$$

<u>Proof</u>: since it's an orthonormal basis, all the denominators in our previous formula vanish!

<u>Theorem</u> if $\{u_1, \dots, u_k\}$ is an orthonormal basis for a subspace $W$ of $\mathbb{R}^n$, and $U = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_k \\ | & & | \end{bmatrix}$, then

$$\text{proj}_W y = UU^T y \quad \text{for all } y \in \mathbb{R}^n.$$

## Example

Let $W = \operatorname{span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$,

and $x = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$.

What is the closest vector to $x$ in $W$?

Notice that ~~although~~ the spanning set for $W$ is orthogonal, although not orthonormal.

Let's normalise it!

$$\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = 9, \quad \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \cdot 9,$$

so define

$$u_1 = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \quad u_2 = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}.$$

Now $W = \operatorname{span}\{u_1, u_2\}$, and $\{u_1, u_2\}$ is an orthonormal basis for $W$.

Write $U = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$.

We could verify that the columns of $U$ really are orthonormal, by verifying to $U^TU = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

The closest vector is

$$\mathrm{proj}_W x = UU^T x = \begin{bmatrix} \frac{2}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$$

$$= \frac{1}{9} \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}.$$

We can also compute the distance from $x$ to $W$:

$$\|x - \mathrm{proj}_W x\| = \left\| \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix} \right\| = \sqrt{4+16+16} = 6.$$

This example showed that the standard ~~s~~ matrix for ~~proje~~ projection onto $W$ is

$$\frac{1}{9}\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}.$$

What if we worked in $\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 2 \end{bmatrix} \right\}$ coordinates?

Here I've carefully chosen $b_1$ and $b_2$ spanning $W$, and ~~$b_3$~~ $b_3$ orthogonal to $W$.

Thus each basis vector is an eigenvector for projection, and the matrix for projection in $\mathcal{B}$ coordinates ~~is~~ is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.