

Orthogonal bases are the best:

Thm (expansion thm for coordinates relative to an orthogonal basis)

Suppose $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthogonal basis for \mathbb{R}^n . Then for any vector \vec{v} in \mathbb{R}^n ,

we have:

$$\vec{v} = \frac{\vec{v} \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 + \dots + \frac{\vec{v} \cdot \vec{v}_n}{\|\vec{v}_n\|^2} \vec{v}_n$$

that is, the coordinates of \vec{v} with respect to the basis B are

$$[\vec{v}]_B = \left(\frac{\vec{v} \cdot \vec{v}_1}{\|\vec{v}_1\|^2}, \frac{\vec{v} \cdot \vec{v}_2}{\|\vec{v}_2\|^2}, \dots, \frac{\vec{v} \cdot \vec{v}_n}{\|\vec{v}_n\|^2} \right).$$

In other words: to express \vec{v} as a linear combination of elements of an orthogonal basis, you don't have to do any row reduction — you just calculate some dot products.

* In function spaces for waveforms (periodic fcts), this theorem is the Fourier series expansion of your waveform

Proof: Actually, we basically already did the work.

So let $\{\vec{v}_1, \dots, \vec{v}_n\}$ be an orthogonal basis of \mathbb{R}^n .
Since it is a basis of \mathbb{R}^n , any $\vec{v} \in \mathbb{R}^n$ can be written as

$$\vec{v} = a_1 \vec{v}_1 + \dots + a_n \vec{v}_n$$

in a unique way.

Take the dot product of both sides with \vec{v}_i :

$$\begin{aligned} \vec{v}_i \cdot \vec{v} &= a_1 (\vec{v}_i \cdot \vec{v}_1) + \dots + a_i (\vec{v}_i \cdot \vec{v}_i) + \dots + a_n (\vec{v}_i \cdot \vec{v}_n) \\ &= a_1 (0) + \dots + a_i \|\vec{v}_i\|^2 + \dots + a_n (0) \end{aligned}$$

all dot products are zero except $\vec{v}_i \cdot \vec{v}_i$, by orthogonality

$$= a_i \|\vec{v}_i\|^2.$$

$$\therefore a_i = \frac{\vec{v}_i \cdot \vec{v}}{\|\vec{v}_i\|^2} = \frac{\vec{v} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} \quad (\text{we used } \vec{v} \cdot \vec{u} = \vec{u} \cdot \vec{v})$$

So the coordinates are the given formulae.

□

example: Express $(1, 2, 3)$ as a linear combination of $B = \{(1, 2, 1), (1, 0, -1), (1, -1, 1)\}$.

Option 1: row reduce $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 0 & -1 & 2 \\ 1 & -1 & 1 & 3 \end{array} \right]$.

Option 2: We notice that B is an orthogonal basis of \mathbb{R}^3 (actually, we just notice it is orthogonal & has 3 vectors, therefore is a basis)

\therefore we calculate:

$$\vec{v} = (1, 2, 3) \quad \vec{v}_1 = (1, 2, 1) \Rightarrow \vec{v}_1 \cdot \vec{v} = 1 + 4 + 3 = 8$$

$$\|\vec{v}_1\|^2 = 1^2 + 2^2 + 1^2 = 6$$

$$\therefore a_1 = \frac{8}{6} = \frac{4}{3}$$

$$\vec{v}_2 = (1, 0, -1) \quad \therefore \vec{v} \cdot \vec{v}_2 = 1 - 3 = -2$$

$$\|\vec{v}_2\|^2 = 2$$

$$\therefore a_2 = \frac{-2}{2} = -1$$

$$\vec{v}_3 = (1, -1, 1) \quad \therefore \vec{v} \cdot \vec{v}_3 = 1 - 2 + 3 = 2$$

$$\|\vec{v}_3\|^2 = 1 + (-1)^2 + 1 = 3$$

$$\therefore a_3 = \frac{2}{3}$$

$$\therefore \vec{v} = \frac{4}{3}\vec{v}_1 - \vec{v}_2 + \frac{2}{3}\vec{v}_3.$$

We check:

$$\begin{aligned} & \frac{4}{3}(1, 2, 1) - (1, 0, -1) + \frac{2}{3}(1, -1, 1) \\ &= \left(\frac{4}{3}, \frac{8}{3}, \frac{4}{3}\right) + \left(-\frac{3}{3}, 0, \frac{3}{3}\right) + \left(\frac{2}{3}, -\frac{2}{3}, \frac{2}{3}\right) \end{aligned}$$

$$= \left(\frac{3}{3}, \frac{6}{3}, \frac{9}{3}\right) = (1, 2, 3) = \vec{v} \quad \checkmark$$

□

example: Consider $W = \{(x, y, z, w) \in \mathbb{R}^4 \mid x - y - z + w = 0\}$

So this is Null($\underbrace{[1 \ -1 \ -1 \ 1]}_A$)
a matrix of rank 1

$$\therefore \dim W = 4 - \text{rank } A = 3.$$

Say we have found the following 3 vectors:

$$\vec{w}_1 = (1, 1, 1, 1) \quad \vec{w}_2 = (1, -1, 1, -1), \quad \vec{w}_3 = (1, 1, -1, -1).$$

These are each in W (because $x - y - z + w = 0$)
and they are orthogonal (do their pairwise dot products)

\therefore this is an LI ~~set~~ of 3 vectors in W
 \Rightarrow a basis of W . (in fact an orthogonal basis!)

So we have $\vec{w} = (4, 3, 2, 1) \in W$ (check).

Write \vec{w} in terms of this basis:

$$\left. \begin{aligned} \frac{\vec{w} \cdot \vec{w}_1}{\|\vec{w}_1\|^2} &= \frac{4+3+2+1}{4} = \frac{10}{4} = \frac{5}{2} \\ \frac{\vec{w} \cdot \vec{w}_2}{\|\vec{w}_2\|^2} &= \frac{4-3+2-1}{4} = \frac{2}{4} = \frac{1}{2} \\ \frac{\vec{w} \cdot \vec{w}_3}{\|\vec{w}_3\|^2} &= \frac{4+3-2-1}{4} = \frac{4}{4} = 1 \end{aligned} \right\} \therefore \vec{w} = \frac{5}{2} \vec{w}_1 + \frac{1}{2} \vec{w}_2 + \vec{w}_3$$

as we can readily check.

Next up: orthogonal projection.

Mar 21: orthogonal projections and Gram-Schmidt Ch 19

update. Test 4 covers Ch 14-19.2

19.3 orthogonal projections

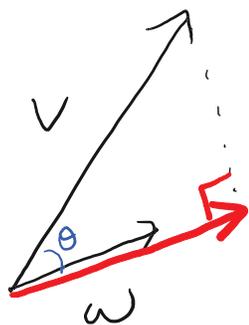
So that formula for the Fourier coefficients of a vector

$$\frac{w_1 \cdot v}{w_1 \cdot w_1} w_1 + \dots + \frac{w_n \cdot v}{w_n \cdot w_n} w_n$$

makes sense even if $v \notin \text{span}\{w_1, \dots, w_n\}$.

What does it give us in this case? (The answer can't be v , because v is in $\text{span}\{w_1, \dots, w_n\}$.)

Recall: the projection of a vector v onto a vector w is given by



$$\text{proj}_w(v) = \frac{w \cdot v}{w \cdot w} w = \frac{v \cdot w}{\|w\|^2} w$$

Proof: $w \cdot v = \|w\| \underbrace{\|v\| \cos \theta}_{\text{length of projection}}$

and a unit vector in the direction of w is $\frac{w}{\|w\|}$. $\therefore \text{proj} = \|v\| \cos \theta \cdot \frac{\bar{w}}{\|w\|}$

$$= \frac{w \cdot v}{w \cdot w} \bar{w}$$

□

So our formula is just the sum of the orthogonal projections of v onto each basis vector of the subspace.

Def 1 (Orthogonal projection onto a subspace)

Suppose W is a subspace of \mathbb{R}^n and

$\{\vec{w}_1, \dots, \vec{w}_k\}$ is an **orthogonal basis** of W .

Then for any $\vec{v} \in \mathbb{R}^n$, the orthogonal projection of \vec{v} onto W is defined by

$$\text{proj}_W(\vec{v}) = \frac{\vec{v} \cdot \vec{w}_1}{\|\vec{w}_1\|^2} \vec{w}_1 + \dots + \frac{\vec{v} \cdot \vec{w}_k}{\|\vec{w}_k\|^2} \vec{w}_k$$

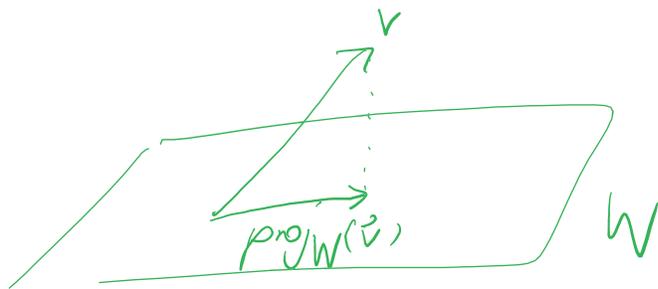
eg) Recall that the subspace $W = \{(x, y, z, w) \in \mathbb{R}^4 \mid x - y - z + w = 0\}$ had orthogonal basis $\{(1, 1, 1, 1), (1, -1, 1, -1), (1, 1, -1, -1)\}$.

So using our formula on $\vec{v} = (1, 1, 1, 0)$ we

get $\text{proj}_W(\vec{v}) = \frac{3}{4} \vec{w}_1 + \frac{1}{4} \vec{w}_2 + \frac{1}{4} \vec{w}_3 = \left(\frac{5}{4}, \frac{3}{4}, \frac{3}{4}, \frac{1}{4}\right)$

which is a vector in W that's kind of close to \vec{v} .

In fact, $\vec{v} - \text{proj}_W(\vec{v}) = \left(-\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$ is $\perp W$.



eg) Consider $W = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = 0\}$.
 An orthogonal basis for this plane is $\{(1, -1, 0), (0, 0, 1)\}$
 (since it is orthogonal, it is LI, and has 2 elements
 and $\dim W = 2 \therefore$ basis)

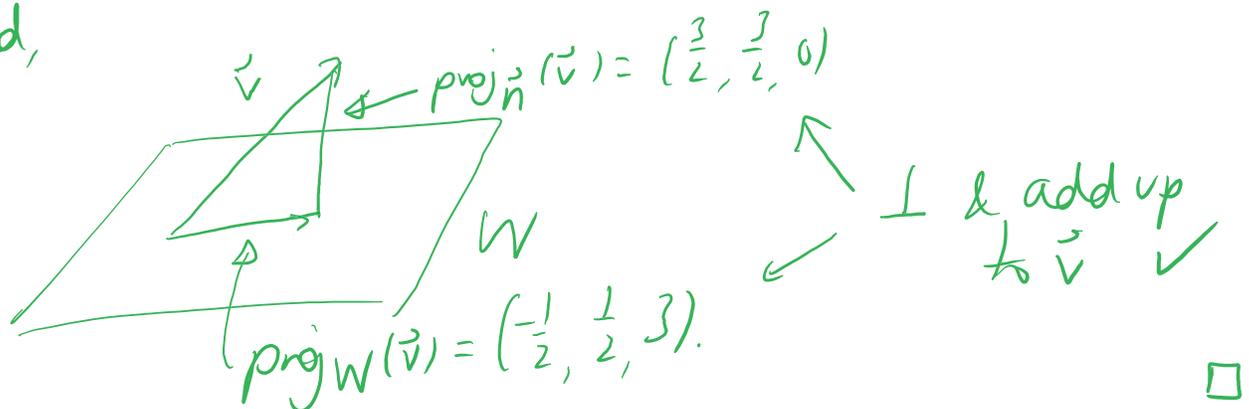
Let $v = (1, 2, 3)$. We calculate:

$$\frac{v \cdot w_1}{\|w_1\|^2} w_1 + \frac{v \cdot w_2}{\|w_2\|^2} w_2 = \frac{1-2}{2} w_1 + \frac{3}{1} w_2 = -\frac{1}{2} (1, -1, 0) + 3(0, 0, 1) \\ = \left(-\frac{1}{2}, \frac{1}{2}, 3\right)$$

A normal vector to W is $\vec{n} = (1, 1, 0)$.

The projection of \vec{v} onto \vec{n} is $\frac{v \cdot \vec{n}}{\|\vec{n}\|^2} \vec{n} = \frac{3}{2} (1, 1, 0) = \left(\frac{3}{2}, \frac{3}{2}, 0\right)$.

So indeed,



So in our examples, we observe that:

$$\text{proj}_W(\vec{v}) \in W$$

$\vec{v} - \text{proj}_W(\vec{v})$ is orthogonal to W

$\Rightarrow \text{proj}_W(\vec{v})$ is the closest point in W to \vec{v} .

Thm (Best Approximation Theorem)

Let W be a subspace of \mathbb{R}^n . Then for any $\vec{v} \in \mathbb{R}^n$:

- 1) $\text{proj}_W(\vec{v}) \in W$
- 2) $\vec{v} - \text{proj}_W(\vec{v})$ is orthogonal to every vector in W
- 3) The vector $\vec{w} = \text{proj}_W(\vec{v})$ is the vector in W which is closest to \vec{v} , i.e. for which $\|\vec{v} - \vec{w}\|$ is minimum.

In particular, $\text{proj}_W(\vec{v})$ is uniquely defined by these conditions, so you get the same answer even if you use a different orthogonal basis.

Proof: Let $\{w_1, \dots, w_k\}$ be an orthogonal basis for W and use it to define $\text{proj}_W(\vec{v})$.

(1) is clear, since $\text{proj}_W(\vec{v})$ is a linear combination of $w_1, \dots, w_k \dots$ in W .

(2) Let's show that $\vec{v} - \text{proj}_W(\vec{v})$ is \perp to w_i for each basis vector w_i , first.

$$\text{Well: } w_i \cdot \left(\vec{v} - \frac{\vec{v} \cdot w_1}{\|w_1\|^2} w_1 - \dots - \frac{\vec{v} \cdot w_k}{\|w_k\|^2} w_k \right)$$

$$= w_i \cdot \vec{v} - \frac{\vec{v} \cdot w_1}{\|w_1\|^2} \underbrace{w_i \cdot w_1}_{=0} - \dots - \frac{\vec{v} \cdot w_i}{\|w_i\|^2} w_i \cdot w_i - \dots - \frac{\vec{v} \cdot w_k}{\|w_k\|^2} \underbrace{w_i \cdot w_k}_{=0}$$

$$= w_i \cdot \vec{v} - \frac{\vec{v} \cdot w_i}{w_i \cdot w_i} w_i \cdot w_i$$

$$= 0 \quad \text{since } w_i \cdot \vec{v} = \vec{v} \cdot w_i.$$

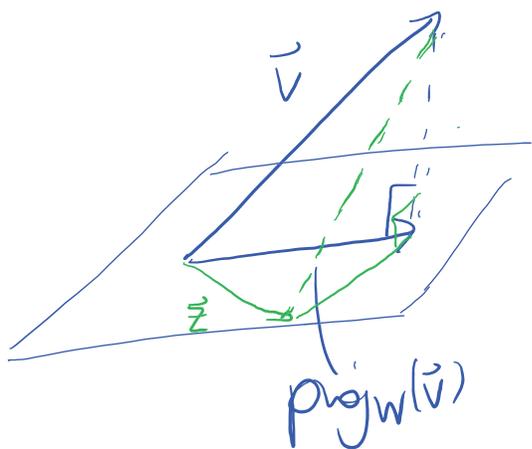
So $\vec{x} = \vec{v} - \text{proj}_W(\vec{v})$ is \perp to each w_i .

If $\vec{w} \in W$, then $\vec{w} = a_1 \vec{w}_1 + \dots + a_k \vec{w}_k$ for some a_i .

$$\begin{aligned} \therefore \vec{x} \cdot \vec{w} &= \vec{x} \cdot (a_1 \vec{w}_1 + \dots + a_k \vec{w}_k) \\ &= a_1 (\vec{x} \cdot \vec{w}_1) + \dots + a_k (\vec{x} \cdot \vec{w}_k) \\ &= 0. \quad \text{since } \vec{x} \cdot \vec{w}_i = 0 \text{ for each } i. \end{aligned}$$

$\therefore \vec{x}$ is orthogonal to every vector in W .

(3) So let \vec{z} be some other vector in W , and let's show that



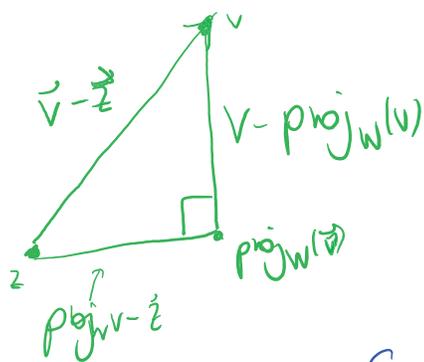
$$\|\vec{v} - \vec{z}\| \geq \|\vec{v} - \text{proj}_W(\vec{v})\|.$$

That would prove that the projection is the closest vector in W to \vec{v}

("the best approximation")

So:

$$\vec{v} - \vec{z} = \underbrace{\vec{v} - \text{proj}_W(\vec{v})}_{\text{this vector is } \perp \text{ to } W} + \underbrace{\text{proj}_W(\vec{v}) - \vec{z}}_{\text{this vector is in } W}$$



\therefore by Pythagoras:

$$\|\vec{v} - \vec{z}\|^2 = \|\vec{v} - \text{proj}_W(\vec{v})\|^2 + \underbrace{\|\text{proj}_W(\vec{v}) - \vec{z}\|^2}_{\geq 0}$$

So $\|\vec{v} - \vec{z}\|^2$ is always a little bit more than $\|\vec{v} - \text{proj}_W(\vec{v})\|^2$ (unless $\vec{z} = \text{proj}_W(\vec{v})$).
So yes, the projection minimizes the distance.

So (3) tells us that $\text{proj}_W(\vec{v})$ is the unique closest point in W to \vec{v} — and that is a geometric fact. So no matter what formula we use, it has to give the same point.

□

eg) Another orthogonal basis for the plane $x+y=0$ is $\{(2, -2, 4), (-1, 1, 1)\}$, as you can readily check. Using this basis we calculate the projection of $(1, 2, 3)$ onto W as:

$$\frac{v \cdot w_1}{\|w_1\|^2} w_1 + \frac{v \cdot w_2}{\|w_2\|^2} w_2 = \frac{2 - 4 + 12}{4 + 4 + 16} w_1 + \frac{-1 + 2 + 3}{1 + 1 + 1} w_2$$

$$= \frac{10}{24} w_1 + \frac{4}{3} w_2$$

$$= \frac{5}{12} (2, -2, 4) + \frac{4}{3} (-1, 1, 1)$$

$$= \left(\frac{5}{6} - \frac{8}{6}, -\frac{5}{6} + \frac{8}{6}, \frac{5}{3} + \frac{4}{3} \right)$$

$$= \left(-\frac{1}{2}, \frac{1}{2}, 3 \right) \quad \checkmark \text{ same as before.}$$

□

This theorem is extremely practical, and used in image processing and in error correcting codes.

Error correcting codes: so the signals you send out lie in certain subspaces, but noise on the communication channel means they aren't in the subspace when they arrive. So we project, correcting the errors the noise introduced

To make all this work (Fourier coefficients, orthogonal projection): we need orthogonal bases for our subspaces. How do we find them?

Answer: Gram-Schmidt Algorithm

- Suppose $\{u_1, \dots, u_m\}$ is any basis of a vector space U .

Define:

$$w_1 = u_1 \quad \text{and} \quad V_1 = \text{span}\{w_1\}.$$

$$w_2 = u_2 - \text{proj}_{V_1}(u_2) \quad \text{and} \quad V_2 = \text{span}\{w_1, w_2\}.$$

$$w_3 = u_3 - \text{proj}_{V_2}(u_3) \quad \text{and} \quad V_3 = \text{span}\{w_1, w_2, w_3\}$$

$$\vdots$$
$$w_m = u_m - \text{proj}_{V_{m-1}}(u_m) \quad \text{and} \quad V_m = \text{span}\{w_1, \dots, w_m\}.$$

Then $V_m = U$ and $\{w_1, \dots, w_m\}$ is an orthogonal basis of U .

So why does this work? each vector w_i is a linear combination of u_1, \dots, u_i , so lies in U . Plus, each new vector w_i is \perp to all w_1, \dots, w_{i-1} .

example: Say we want an orthogonal basis of
 $U = \{(x, y, z, w) \in \mathbb{R}^4 \mid x + z + w = 0\}$.

We start with one basis:

$$\left\{ \underbrace{(1, 1, -1, 1)}_{u_1}, \underbrace{(0, 1, 1, -2)}_{u_2}, \underbrace{(0, 0, 1, -2)}_{u_3} \right\}.$$

So $w_1 = u_1 = (1, 1, -1, 1)$.

$$w_2 = u_2 - \text{proj}_{w_1} u_2 = u_2 - \frac{u_2 \cdot w_1}{w_1 \cdot w_1} w_1$$

$$= (0, 1, 1, -2) - \frac{1-1-2}{1+1+1+1} (1, 1, -1, 1)$$

$$= (0, 1, 1, -2) + \frac{1}{2} (1, 1, -1, 1)$$

$$= \left(\frac{1}{2}, \frac{3}{2}, \frac{1}{2}, -\frac{3}{2} \right)$$

(check: $w_1 \perp w_2$ ✓)
 $w_2 \in U$ ✓)

and so

$$w_3 = u_3 - \text{proj}_{w_1}(u_3) - \text{proj}_{w_2}(u_3)$$

$$= (0, 0, 1, -2) - \frac{u_3 \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{u_3 \cdot w_2}{w_2 \cdot w_2} w_2$$

$$= (0, 0, 1, -2) - \frac{-3}{4} (1, 1, -1, 1) - \frac{7/2}{\frac{1}{4} + \frac{9}{4} + \frac{1}{4} + \frac{9}{4}} \left(\frac{1}{2}, \frac{3}{2}, \frac{1}{2}, -\frac{3}{2} \right)$$

$$= (0, 0, 1, -2) + \left(\frac{3}{4}, \frac{3}{4}, -\frac{3}{4}, \frac{3}{4} \right) - \frac{7}{10} \left(\frac{1}{2}, \frac{3}{2}, \frac{1}{2}, -\frac{3}{2} \right)$$

$$= \left(0, 0, \frac{20}{20}, -\frac{40}{20} \right) + \left(\frac{15}{20}, \frac{15}{20}, -\frac{15}{20}, \frac{15}{20} \right) + \left(-\frac{7}{20}, -\frac{21}{20}, -\frac{7}{20}, \frac{21}{20} \right)$$

$$= \left(\frac{8}{20}, -\frac{6}{20}, -\frac{2}{20}, -\frac{4}{20} \right) = \frac{1}{10} (4, -3, -1, -2) \quad \begin{array}{l} \in W \\ \perp w_1, w_2 \end{array} \checkmark$$

Thus an orthogonal basis is

$$\left\{ (1, 1, -1, 1), \left(\frac{1}{2}, \frac{3}{2}, \frac{1}{2}, \frac{-3}{2}\right), \left(\frac{4}{10}, \frac{-3}{10}, \frac{-1}{10}, \frac{-2}{10}\right) \right\}$$

We could scale the vectors if we wanted a nicer-looking basis; or we could have started with a nicer basis of U (one with more zeros...) to get a different choice.

□

eg) If we had instead started with

$$\left\{ (0, 1, 0, 0), (-2, 0, 1, 0), (-1, 0, 0, 1) \right\}$$

(the basic solutions to the equation $x + 2z + w = 0$)
then by Gram-Schmidt we'd have gotten:

$$w_1 = u_1 = (0, 1, 0, 0)$$

$$w_2 = u_2 - \text{proj}_{w_1}(u_2) = u_2 = (-2, 0, 1, 0) \quad \text{since already } \perp$$

$$w_3 = u_3 - \text{proj}_{w_1}(u_3) - \text{proj}_{w_2}(u_3)$$

$$= u_3 - 0 - \frac{u_3 \cdot w_2}{w_2 \cdot w_2} w_2$$

$$= (-1, 0, 0, 1) - \frac{2}{4+1} (-2, 0, 1, 0)$$

$$= (-1, 0, 0, 1) - \frac{2}{5} (-2, 0, 1, 0)$$

$$= \left(-\frac{1}{5}, 0, \frac{-2}{5}, 1\right)$$

check: in W ✓
 $\perp w_1$ ✓ $\perp w_2$ ✓

So we get the orthogonal basis

$$\{(0, 1, 0, 0), (-2, 0, 1, 0), (-\frac{1}{5}, 0, -\frac{2}{5}, 1)\}$$

We sometimes want an orthonormal basis (because it makes the projection formula easier) so scale each vector by its norm.

$$\{(0, 1, 0, 0), (-\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}, 0), \frac{\sqrt{5}}{\sqrt{6}}(-\frac{1}{5}, 0, -\frac{2}{5}, 1)\}$$

□

Another approach to orthogonality (which we repeatedly skip for time) is using the observation that

$$\text{Null}(A) \perp \text{Row}(A).$$

So when we want to find a vector orthogonal to $\{w_1, \dots, w_n\}$, we can create the matrix $\begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$ and find its nullspace.