Orthonogonal Bases

(Section 5.3)

Recall Q: How do we solve a best-fit polynomial question, written as $A\vec{x} = \vec{b}$ with $\vec{b}$ not in Ran(A), if A^tA is not invertible?

Since the process of solution is to replace $\vec{b}$ with its orthogonal projection onto Ran(A) and solve that equation, we'd be OK if we could (figure out a formula for the orthogonal projection.)

Goal →

Example 1: (a line in $\mathbb{R}^n$) Any line through the origin in $\mathbb{R}^n$ is the span of a vector $\vec{v} \neq \vec{0}$.

Let $\vec{u} = \frac{1}{\|\vec{v}\|_2} \vec{v}$. I

Claim that the orthogonal projection onto $\text{span}(\vec{v})$ is

$$\boxed{P = \vec{u}\,\vec{u}^t.}$$

First, observe that ($\vec{v}$ is a column vector)

$$\vec{u}^t \cdot \vec{u} = \left(\frac{\vec{v}}{\|\vec{v}\|_2}\right)^t \left(\frac{\vec{v}}{\|\vec{v}\|_2}\right)$$

$$= \frac{1}{\|\vec{v}\|_2} \vec{v}^t \cdot \frac{\vec{v}}{\|\vec{v}\|_2}$$

$$= \frac{1}{\|\vec{v}\|_2^2} \vec{v}^t \cdot \vec{v}$$

$$= \frac{1}{\cancel{\|\vec{v}\|_2^2}} \cancel{\|\vec{v}\|_2^2}$$

$$= 1.$$

$\vec{u}$ is a vector of Euclidean length 1, called a unit vector.

Then we check that $P = \vec{v}\,\vec{v}^{\,t}$ is an orthogonal projection:

$$P^2 = (\vec{v}\,\vec{v}^{\,t})(\vec{v}\,\vec{v}^{\,t})$$

$$= \vec{v}\,(\underbrace{\vec{v}^{\,t}\vec{v}}_{=1})\,\vec{v}^{\,t}$$

$$= \vec{v}\cdot\vec{v}^{\,t} = P \checkmark$$

$$P^t = (\vec{v}\cdot\vec{v}^{\,t})^t = (\vec{v}^{\,t})^t\,\vec{v}^{\,t}$$

$$= \vec{v}\,\vec{v}^{\,t}$$

$$= P \checkmark$$

Moreover, if $\vec{x} \in \mathbb{R}^n$,

$$P\vec{x} = \vec{u}\,\underbrace{\vec{u}^t \cdot \vec{x}}_{\text{scalar}}$$

$$= (\vec{u}^t \cdot x) \cdot \vec{u}$$

$$= (\vec{u}^t \cdot x) \cdot \frac{\vec{v}}{\|\vec{v}\|_2}$$

$$= \left(\frac{\vec{u}^t \cdot x}{\|\vec{v}\|_2} \right) \cdot \vec{v}$$

$=$ a multiple of $\vec{v}$

So $\mathrm{Ran}(P) = \mathrm{span}(\vec{v}) =$ a line.

Example 2: (a plane in $\mathbb{R}^3$)

Let W be the plane through the origin in $\mathbb{R}^3$ determined by

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \;\middle|\; 2x - 3y + z = 0 \right\}.$$

Find an orthogonal projection onto W.

We need 2 vectors that are not multiples of eachother to get all of W.

Note if $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in W$, then

$$2x - 3y + z = 0, \text{ so}$$

$$z = 3y - 2x$$

Replacing z, vectors in W look like

$$\begin{bmatrix} x \\ y \\ 3y-2x \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ -2x \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ 3y \end{bmatrix}$$

$$= x\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

Let $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$,

$$\vec{u}_1 = \frac{\vec{v}_1}{\|v_1\|_2}, \quad \vec{u}_2 = \frac{\vec{v}_2}{\|v_2\|_2}.$$

Q: Does $P = \vec{u}_1 \vec{u}_1^{\,t} + \vec{u}_2 \vec{u}_2^{\,t}$ work?

Check whether $P = P^t$ and $P = P^2$.

$$P^t = \left((\vec{v}_1 \vec{v}_1^{\,t}) + (\vec{v}_2 \vec{v}_2^{\,t}) \right)^t$$

$$= (\vec{v}_1 \vec{v}_1^{\,t})^t + (\vec{v}_2 \vec{v}_2^{\,t})^t$$

$$= (v_1^{\,t})^t \vec{v}_1^{\,t} + (\vec{v}_2^{\,t})^t v_2^{\,t}$$

$$= \vec{v}_1 \vec{v}_1^{\,t} + \vec{v}_2 \vec{v}_2^{\,t}$$

$$= P \checkmark$$

$$p^2 = p \cdot p$$

$$= (\vec{v}_1 \vec{v}_1^{\,t} + \vec{v}_2 \vec{v}_2^{\,t}) \cdot (\vec{v}_1 \vec{v}_1^{\,t} + \vec{v}_2 \vec{v}_2^{\,t})$$

$$= \vec{v}_1 \vec{v}_1^{\,t} (\vec{v}_1 \vec{v}_1^{\,t}) + (\vec{v}_1 \vec{v}_1^{\,t})(\vec{v}_2 \vec{v}_2^{\,t}) + (\vec{v}_2 \vec{v}_2^{\,t})(\vec{v}_1 \vec{v}_1^{\,t}) + (\vec{v}_2 \vec{v}_2^{\,t})(\vec{v}_2 v_2^{t})$$

$\vec{v}_1, \vec{v}_2$ are unit vectors

$$= \vec{v}_1 (\underbrace{\vec{v}_1^{\,t} \vec{v}_1}_{=1}) \vec{v}_1^{\,t} + \vec{v}_1 (\vec{v}_1^{\,t} \vec{v}_2) v_2^{t} + \vec{v}_2 (v_2^{t} \vec{v}_1) \vec{v}_1^{\,t} + \vec{v}_2 (\underbrace{\vec{v}_2^{\,t} \vec{v}_2}_{=1}) v_2^{t}$$

$$= \vec{v}_1 \vec{v}_1^{\,t} + \vec{v}_2 \vec{v}_2^{\,t} + \vec{v}_1 (\vec{v}_1^{\,t} \vec{v}_2) \vec{v}_2^{\,t} + \vec{v}_2 (\vec{v}_2^{\,t} \vec{v}_1) \vec{v}_1^{\,t}$$

So $P^2 = \underbrace{\vec{v}_1 \vec{v}_1^{\,t} + \vec{v}_2 \vec{v}_2^{\,t}}_{P}$
$$+ \vec{v}_1 (\vec{v}_1^{\,t} \vec{v}_2) \vec{v}_2^{\,t}$$
$$+ \vec{v}_2 (\vec{v}_2^{\,t} \vec{v}_1) \vec{v}_1^{\,t}$$

$$P^2 = P + \vec{v}_1 (\vec{v}_1^{\,t} \vec{v}_2) \vec{v}_2^{\,t} + \vec{v}_2 (\vec{v}_2^{\,t} \vec{v}_1) \vec{v}_1^{\,t}$$

Should be zero for P to be an orthogonal projection.

But you can check that this isn't zero!

Our problems would disappear if we knew

$$\boxed{\vec{v}_1^{\,t} \cdot \vec{v}_2 = 0}.$$

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

$$\|\vec{v}_1\|_2 = \sqrt{\vec{v}_1^{\,t}\ \vec{v}_1} = \sqrt{\begin{bmatrix} 1 & 0 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}}$$

$$\|\vec{v}_1\|_2 = \sqrt{5}$$

$$\|\vec{v}_2\|_2^2 = \sqrt{\vec{v}_2^{\,t} \cdot \vec{v}_2} = \sqrt{\begin{bmatrix} 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}}$$

$$\|\vec{v}_2\|_2 = \sqrt{10} = 2\sqrt{5}$$

Calculate

$$\vec{v}_1(\vec{v}_1^{\,t}\,\vec{v}_2)\,\vec{v}_2^{\,t} :$$

$$\vec{v}_1 (\vec{v}_1^t \vec{v}_2) v_2^t$$

$$= \frac{\vec{v}_1}{\|\vec{v}_1\|_2} \left(\left(\frac{\vec{v}_1}{\|\vec{v}_1\|_2}\right)^t \frac{\vec{v}_2}{\|\vec{v}_2\|_2} \right) \left(\frac{\vec{v}_2}{\|\vec{v}_2\|_2}\right)^t$$

$$= \frac{1}{\|\vec{v}_1\|_2^2} \cdot \frac{1}{\|\vec{v}_2\|_2^2} \left(\vec{v}_1 (\vec{v}_1^t \vec{v}_2) \vec{v}_2^t \right)$$

$$= \frac{1}{50} \left(\vec{v}_1 (\vec{v}_1^t \vec{v}_2) v_2^t \right)$$

$$\vec{v}_1^t \vec{v}_2 = [1 \quad 0 \quad -2] \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

$$\vec{v}_1^t \vec{v}_2 = -6$$

So $\vec{v}_1(\vec{v}_1{}^t \vec{v}_2)\vec{v}_2{}^t$

$$= \frac{-6}{50} \vec{v}_1 \vec{v}_2{}^t$$

$$= \frac{-6}{50} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \end{bmatrix}$$

$$= \frac{-6}{50} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -2 \\ 3 & 0 & -6 \end{bmatrix}$$

$$= \frac{-3}{25} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -2 \\ 3 & 0 & -6 \end{bmatrix}$$

Then $\vec{v}_2(\vec{v}_2^{\,t} \cdot \vec{v}_1)\vec{v}_1^{\,t}$

$$= \left(\vec{v}_1(\vec{v}_1^{\,t}\ \vec{v}_2)\vec{v}_2^{\,t}\right)^t$$

$$= -\frac{3}{25}\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -2 \\ 3 & 0 & -6 \end{bmatrix}^t$$

$$= -\frac{3}{25}\begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & -2 & -6 \end{bmatrix}.$$

Adding the two,

$$\vec{v}_1(\vec{v}_1^{\,t} \cdot \vec{v}_2)\vec{v}_2^{\,t} + \vec{v}_2(\vec{v}_2^{\,t} \cdot \vec{v}_1)\vec{v}_1^{\,t}$$

$$= -\frac{3}{25}\left(\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & -2 \\ 3 & 0 & -6 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & -2 & -6 \end{bmatrix}\right)$$

$$= -\frac{3}{25}\left(\begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & -2 \\ 3 & -2 & -12 \end{bmatrix}\right)$$

$$\neq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Orthogonality

A collection of vectors S in $\mathbb{R}^n$, thought of as columns, is said to be **orthogonal** if $\boxed{\vec{v}^{\,t} \cdot \vec{w} = 0}$ for all $\vec{v}, \vec{w} \in S$ $(\vec{v} \neq \vec{w})$.

Example 3: ($\mathbb{R}^2$ orthogonality) In $\mathbb{R}^2$, if $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ is a column vector, let $\vec{w} = \begin{bmatrix} -y \\ x \end{bmatrix}$. Then $\vec{w}$ is always orthogonal to $\vec{v}$

Since

$$\vec{v}^t \cdot \vec{w} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} -y \\ x \end{bmatrix}$$

$$\vec{v}^t \cdot \vec{w} = -yx + xy$$

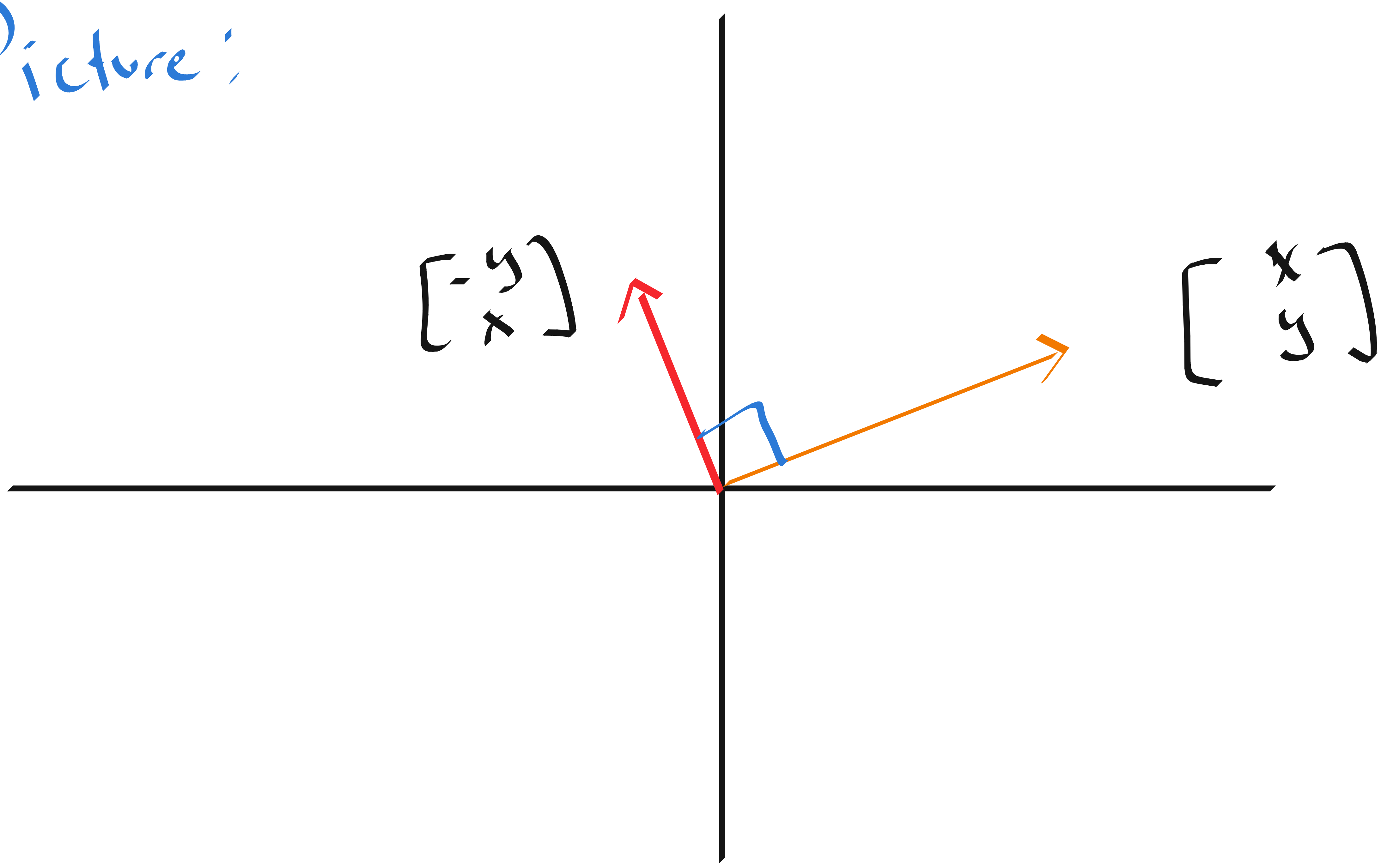
$$\vec{v}^t \cdot \vec{w} = xy - xy$$

$$\vec{v}^t \cdot \vec{w} = 0 \quad ✓$$

If $\begin{bmatrix} x \\ y \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, then $\begin{bmatrix} -y \\ x \end{bmatrix}$ is a vector on the line perpendicular to $\operatorname{span}\left\{\begin{bmatrix} x \\ y \end{bmatrix}\right\}$.

In $\mathbb{R}^2$, orthogonality = perpendicularity.

Picture:



Orthogonal / Orthonormal Bases

A subset $\mathcal{B}$ of nonzero vectors in $\mathbb{R}^n$ is an orthogonal basis for a subspace W of $\mathbb{R}^n$ if

1) $\boxed{\mathcal{B} \subseteq W}$

2) $\text{span}(\mathcal{B}) = W$

3) $\mathcal{B}$ is orthogonal.

If, in addition, every vector in $\mathcal{B}$ is a <u>unit</u> vector, we say $\mathcal{B}$ is an orthonormal basis.

Back to Example 2:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$ are

vectors in W, but not orthogonal,

since

$$\vec{v}_2^{\,t} \cdot \vec{v}_1 = [0 \;\; 1 \;\; 3] \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = -6 \neq 0$$

Replace $\vec{v}_1$ and $\vec{v}_2$ with orthonormal vectors, as follows:

let $\vec{w}_1 = \vec{v}_1$.

let $\vec{w}_2 = \vec{v}_2 - \dfrac{\vec{v}_1^{\,t}\vec{v}_2}{\|\vec{v}_1\|_2^2}\vec{v}_1$.

Here, $\|\vec{v}_1\|_2^2 = \vec{v}_1^{\,t}\vec{v}_1$

$$= \begin{bmatrix} 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

$$= 5$$

So $\vec{w}_2 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} - \dfrac{(-6)}{5} \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$

$$= \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} + \begin{bmatrix} 6/5 \\ 0 \\ -12/5 \end{bmatrix}$$

$$= \begin{bmatrix} 6/5 \\ 1 \\ 3/5 \end{bmatrix}$$

Observe that

$$\vec{w}_2^{\,t}\,\vec{w}_1 = \begin{bmatrix} 6/5 & 1 & 3/5 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

$$= 6/5 - 6/5 = 0$$

$\vec{w}_2 \in \mathrm{span}(\{\vec{v}_1, \vec{v}_2\}) = W$, so

set

$$\vec{u}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|_2} = \frac{\vec{v}_1}{\|\vec{v}_1\|_2}$$

$$\vec{u}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|_2}.$$

Then $\boxed{P = \vec{u}_1\vec{u}_1^{\,t} + \vec{u}_2\vec{u}_2^{\,t}}$ is

the orthogonal projection onto W!

Orthogonal Projections from Orthonormal Bases

If $W \subseteq \mathbb{R}^n$ is a subspace and $W \neq \{\vec{0}\}$, take an orthonormal basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ of W, where $\mathcal{B}$ is written as column vectors. Then

$$\boxed{P = \sum_{i=1}^{k} \vec{v}_i \vec{v}_i^{\,t}}$$

is the orthogonal projection onto W.

Choosing $W = \mathrm{Ran}(A)$ solves our best-fit polynomial issue if A^tA is not invertible, but ...

Q: How to get an orthonormal basis?

A: Later ...