$C(A) = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \right) = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right)$

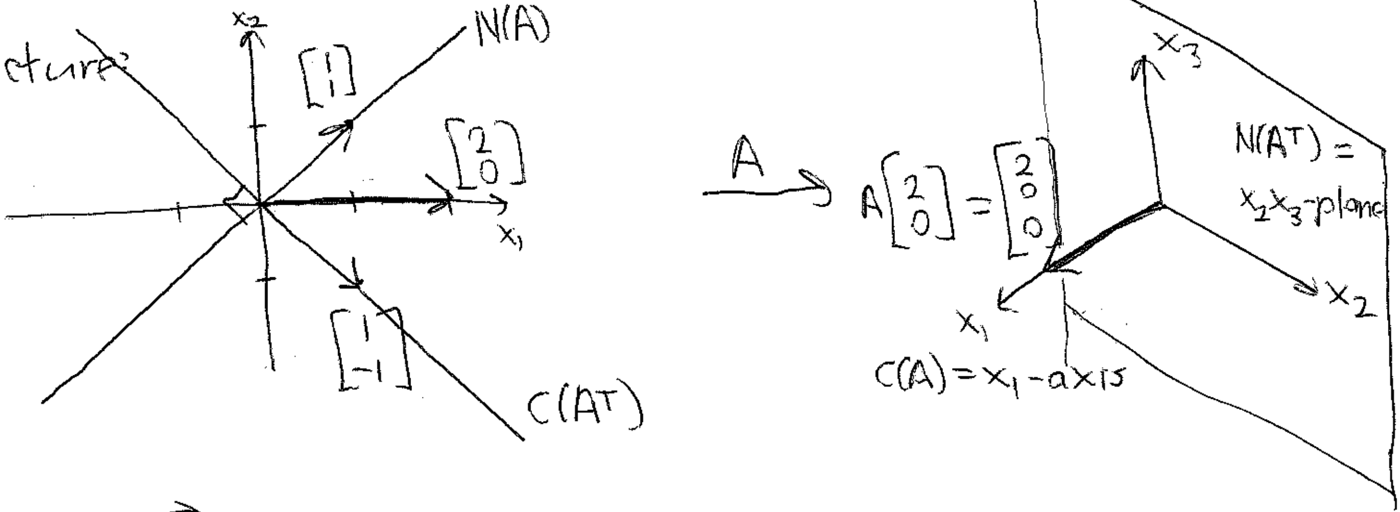
$N(A)$: $\begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ means $x_1 = x_2 \rightsquigarrow N(A) = \text{all } x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$C(A^T) = \text{span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = \text{span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)$ perpen-dicular!

$N(A^T)$: $\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ means $x_1 = 0 \rightsquigarrow$

$N(A^T) = \text{all} \begin{bmatrix} 0 \\ x_2 \\ x_3 \end{bmatrix} = \text{all } x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ $\quad$ ($= x_2x_3$-plane)

The picture:



What are $\vec{x}_r$ and $\vec{x}_n$? Need to write

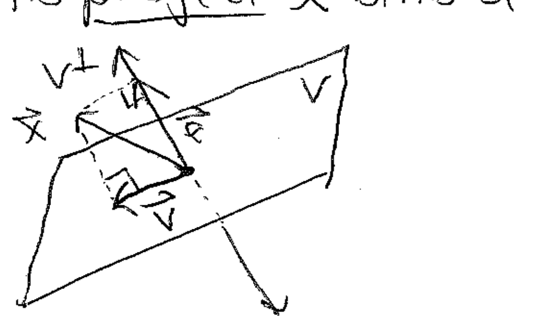
$$\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \underbrace{c \begin{bmatrix} 1 \\ -1 \end{bmatrix}}_{\vec{x}_r} + \underbrace{d \begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\vec{x}_n}$$

Looks like $c = d = 1$ will work.

Section 4.2 Projections

Big linear algebra problem: Figure out how to write $\vec{x} = \vec{x}_r + \vec{x}_n$.

To say another way: Figure out how to <u>project</u> $\vec{x}$ onto a subspace (such as $N(A)$).



This means: Write $\vec{x} = \vec{v} + \vec{e}$

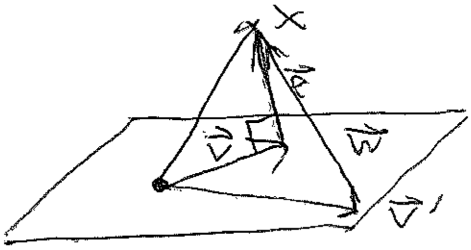
$\vec{v}$ in V $\qquad$ $\vec{e}$ in $V^\perp$

Why call it $\vec{e}$? We are trying to <u>approximate</u> $\vec{x}$ by a vector in V.

$\vec{v}$ is the approximation

$\vec{e}$ is the <u>error</u> of the approximation.

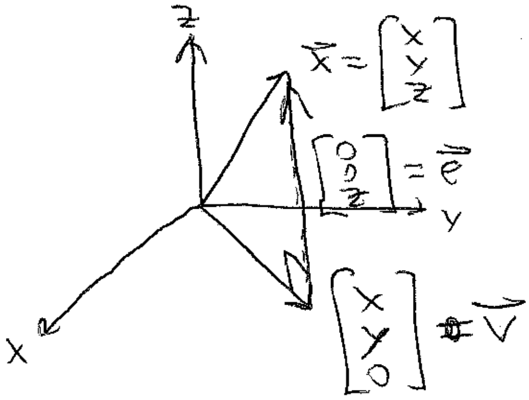
If we did it differently:



We could also write $\vec{x} = \vec{v}' + \vec{w}$
where $\vec{w}$ is <u>not</u> in $V^\perp$, but then
$\|\vec{w}\| > \|\vec{e}\|$, <u>worse</u> error.

So what we want to do is <u>orthogonal</u> projection (minimize error).

<u>Very simple example:</u> Project $\vec{x}$ in $\mathbb{R}^3$ onto the xy-plane.



Error vector $\vec{e}$ is just $\begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$, which is $\perp$ to $\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$ in the xy-plane.

<u>Size</u> of the error is $\left\| \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \right\| = |z|$.

Turns out, there is a <u>matrix</u> that does this projection!

$$\underbrace{\underset{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}{\vec{x}} = \underset{\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}}{\vec{v}}} + \underset{\begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}}{\vec{e}}$$

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\vec{v} = P\vec{x}$$

To project $\vec{x}$ onto $V^\perp$ instead (z-axis) use $I - P$:

$$\begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

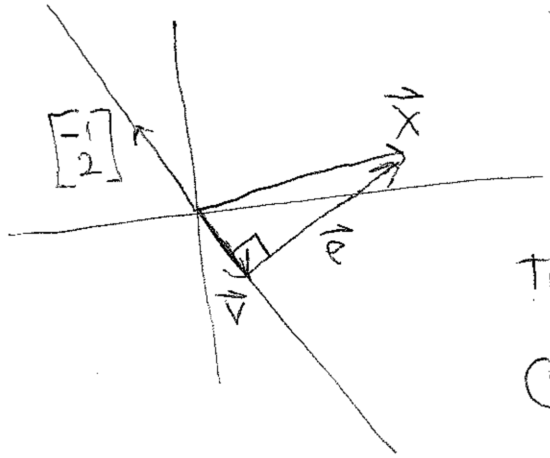
$$\vec{e} = (I - P)\vec{x}$$

P and $I-P$ have two special properties:

<u>Any</u> matrix with these two properties projects vectors onto some subspace!

1. Symmetric: $P = P^T$
2. $P^2 = P$ (projecting twice is the same as projecting once)

Now let's find projection matrix for projecting onto a line in $\mathbb{R}^2$:



$\vec{x} = \vec{v} + \vec{e}$

$\vec{v}$: parallel to line

$\vec{e}$: $\perp$ to line

To find $\vec{v}$ for $\vec{x} = \begin{bmatrix}x\\y\end{bmatrix}$, we know 2 things:

① $\vec{v} = c\begin{bmatrix}-1\\2\end{bmatrix}$ ② $\vec{e} = \vec{x} - \vec{v}$ is $\perp$ to $\begin{bmatrix}-1\\2\end{bmatrix}$

$$c\left(\begin{bmatrix}-1 & 2\end{bmatrix}\begin{bmatrix}-1\\2\end{bmatrix}\right) = \begin{bmatrix}-1 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} \longleftarrow \left(\vec{x} - c\begin{bmatrix}-1\\2\end{bmatrix}\right)\cdot\begin{bmatrix}-1\\2\end{bmatrix} = 0$$

$$\text{So } \vec{v} = c\begin{bmatrix}-1\\2\end{bmatrix} = \boxed{\frac{\begin{bmatrix}-1 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}}{\begin{bmatrix}-1 & 2\end{bmatrix}\begin{bmatrix}-1\\2\end{bmatrix}}}\begin{bmatrix}-1\\2\end{bmatrix}$$

Just a number, c.

Rewrite by putting number $\begin{bmatrix}-1 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$ to the <u>right</u> of $\begin{bmatrix}-1\\2\end{bmatrix}$:

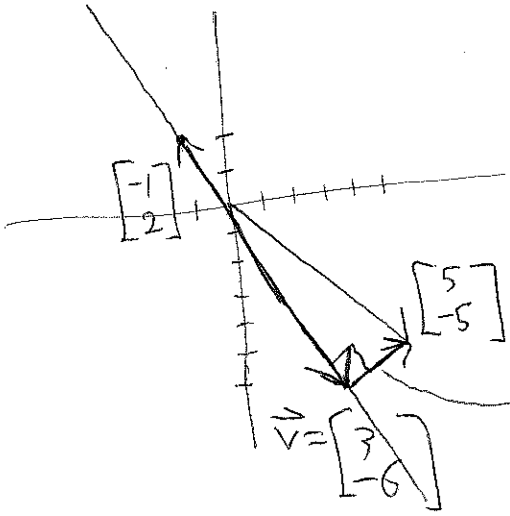
$$\vec{v} = \boxed{\frac{\begin{bmatrix}-1\\2\end{bmatrix}\begin{bmatrix}-1 & 2\end{bmatrix}}{\begin{bmatrix}-1 & 2\end{bmatrix}\begin{bmatrix}-1\\2\end{bmatrix}}}\begin{bmatrix}x\\y\end{bmatrix}$$

$\longleftarrow$ This is a 2×2 matrix, P.

$$P = \frac{1}{1+4}\begin{bmatrix}1 & -2\\-2 & 4\end{bmatrix} = \begin{bmatrix}1/5 & -2/5\\-2/5 & 4/5\end{bmatrix}$$

<u>Example</u>: Project $\vec{x} = \begin{bmatrix}5\\-5\end{bmatrix}$ onto this line:

$$\vec{v} = P\begin{bmatrix}5\\-5\end{bmatrix} = \begin{bmatrix}1/5 & -2/5\\-2/5 & 4/5\end{bmatrix}\begin{bmatrix}5\\-5\end{bmatrix} = \begin{bmatrix}3\\-6\end{bmatrix}$$



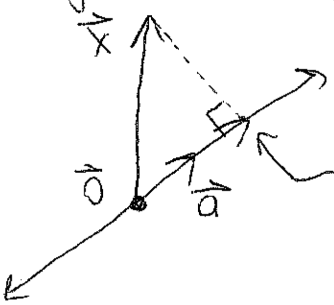
$$\vec{e}=\begin{bmatrix}5\\-5\end{bmatrix}-\begin{bmatrix}3\\-6\end{bmatrix}=\begin{bmatrix}2\\1\end{bmatrix}$$

Size of error: $\left\|\begin{bmatrix}2\\1\end{bmatrix}\right\|=\sqrt{5}$

Check P is a projection matrix: (1) Symmetric, $P=P^T$ ✓

(2) $$P^2=\begin{bmatrix}1/5 & -2/5\\-2/5 & 4/5\end{bmatrix}\begin{bmatrix}1/5 & -2/5\\-2/5 & 4/5\end{bmatrix}=\begin{bmatrix}\frac{1+4}{25} & \frac{-2-8}{25}\\ \frac{-2-8}{25} & \frac{4+16}{25}\end{bmatrix}=\begin{bmatrix}1/5 & -2/5\\-2/5 & 4/5\end{bmatrix}=P$$ ✓

Now project onto *any* line in $\mathbb{R}^n$:



Line $=\operatorname{span}(\vec{a})$

$\vec{v}=c\vec{a}$, and $\vec{e}=\vec{x}-c\vec{a}$ is $\perp$ to $\vec{a}$:

$$\vec{a}^T(\vec{x}-c\vec{a})=0 \longrightarrow \vec{a}^T\vec{x}=c(\vec{a}^T\vec{a})$$

$$\rightarrow c=\frac{\vec{a}^T\vec{x}}{\vec{a}^T\vec{a}}$$

So to project $\vec{x}$ onto $\operatorname{span}(\vec{a})$:

$$\vec{v}=\frac{\vec{a}^T\vec{x}}{\vec{a}^T\vec{a}}\vec{a}$$

Look at $(\vec{a}^T\vec{x})\vec{a}$: $(a_1x_1+a_2x_2+\dots+a_nx_n)\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}=\begin{bmatrix}a_1(a_1x_1+\dots+a_nx_n)\\a_2(a_1x_1+\dots+a_nx_n)\\\vdots\\a_n(a_1x_1+\dots+a_nx_n)\end{bmatrix}$

scalar vector

$$=\begin{bmatrix}a_1a_1 & a_1a_2 & \cdots & a_1a_n\\a_2a_1 & a_2a_2 & \cdots & a_2a_n\\\vdots & \vdots & \ddots & \vdots\\a_na_1 & a_na_2 & \cdots & a_na_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}$$

The "outer product," $\vec{a}\vec{a}^T$

n×n matrix

So $$\vec{v}=\frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}\vec{x}$$

scalar

This means the projectio matrix is: $P = \dfrac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}$

P has the 2 properties of a projection matrix:

(1) symmetric: $P^T = \dfrac{1}{\vec{a}^T\vec{a}}(\vec{a}\,\vec{a}^T)^T = P$ ✓

$(\vec{a}^T)^T\vec{a}^T = \vec{a}\,\vec{a}^T$

(2) $P^2 = \left(\dfrac{1}{\vec{a}^T\vec{a}}\right)^2 \vec{a}\,\underbrace{\vec{a}^T\vec{a}}_{}\,\vec{a}^T = \dfrac{\vec{a}\,\vec{a}^T}{\vec{a}^T\vec{a}} = P$ ✓

scalar, cancels with one in the denominator

Example Projection onto the line $\text{span}\left(\begin{bmatrix}1\\2\\3\end{bmatrix}\right)$ in $\mathbb{R}^3$:

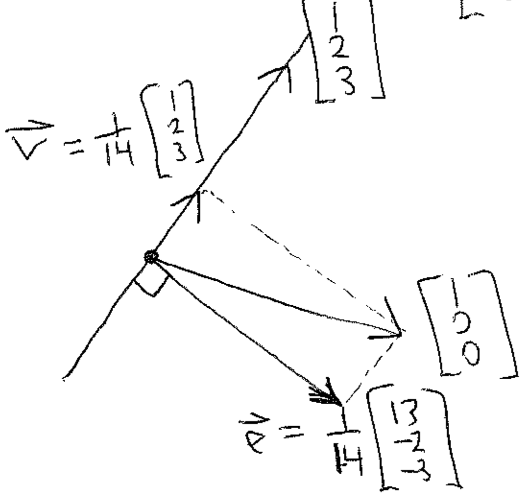
$$P = \frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}} = \frac{\begin{bmatrix}1\\2\\3\end{bmatrix}\begin{bmatrix}1&2&3\end{bmatrix}}{\begin{bmatrix}1&2&3\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}} = \frac{1}{14}\begin{bmatrix}1&2&3\\2&4&6\\3&6&9\end{bmatrix}$$

All rows are multiples of each other
$\longrightarrow \text{rank}(P) = 1$
$\longrightarrow \dim C(P) = 1$
$\longrightarrow C(P)$ is a line

It's the span of $\begin{bmatrix}1\\2\\3\end{bmatrix}$

Let's project $\vec{x} = \begin{bmatrix}1\\0\\0\end{bmatrix}$ onto this line: $P\begin{bmatrix}1\\0\\0\end{bmatrix} = \frac{1}{14}\begin{bmatrix}1\\2\\3\end{bmatrix}$ ← This is $\vec{v}$.

Error: $\vec{e} = \vec{x} - P\vec{x} = \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{1}{14}\begin{bmatrix}1\\2\\3\end{bmatrix} = \frac{1}{14}\begin{bmatrix}13\\-2\\-3\end{bmatrix}$ ← This vector is $\perp$ to $\begin{bmatrix}1\\2\\3\end{bmatrix}$.



Notice that $\vec{e} = \vec{x} - P\vec{x}$

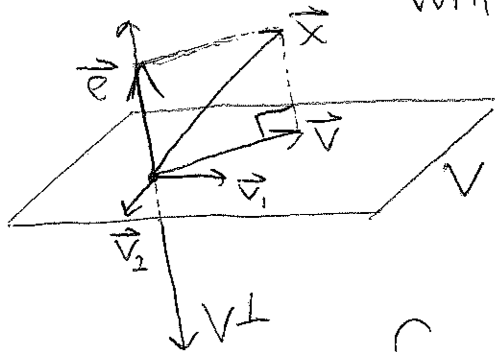
$= (I - P)\vec{x}$

$I - P$ projects $\vec{x}$ onto the orthogonal complement, $V^\perp$.

Now what if V is bigger than a line?

Suppose $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ is a <u>basis</u> for a subspace V in $\mathbb{R}^n$:

Project $\vec{x}$ onto V:



Write $\vec{x} = \vec{v} + \vec{e}$

$\vec{v}$ in V, $\vec{e}$ in $V^\perp$

Know 2 things:

1. $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_m\vec{v}_m$
2. $\vec{x} - \vec{v}$ is $\perp$ to <u>all</u> of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$.

From 1:
$$\vec{v} = \underbrace{\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}}_{\vec{c}} = A\vec{c}$$

From 2:
$$\begin{aligned} \vec{v}_1^T(\vec{x} - A\vec{c}) &= 0 \\ \vec{v}_2^T(\vec{x} - A\vec{c}) &= 0 \\ &\vdots \\ v_m^T(\vec{x} - A\vec{c}) &= 0 \end{aligned} \quad \rightsquigarrow \quad \underbrace{\begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_m^T \end{bmatrix}}_{A^T} (\vec{x} - A\vec{c}) = \vec{0}$$

So $A^T\vec{x} = (A^T A)\vec{c}$

$m \times n$ $n \times m$

We want to solve for $\vec{c}$. But is A^TA invertible?

→ It's $m \times m$, so yes if it's columns are independent

→ Yes if $N(A^TA) = \{\vec{0}\}$.

Need to show $\vec{x} = \vec{0}$.

Let's prove $N(A^TA) = \{\vec{0}\}$: Suppose $\vec{x}$ is in the nullspace:

Then $A^TA\vec{x} = \vec{0}$. This means $\vec{x}^T A^T A \vec{x} = 0$ also

$$\Downarrow$$

$$(A\vec{x})^T(A\vec{x}) = 0$$

$A\vec{x}$ is $\perp$ to itself! Only possible if $A\vec{x} = \vec{0}$, so $\vec{x}$ is in $N(A)$. ~~[illegible]~~

But A has <u>independent</u> columns (basis for V), so its null space is $\{\vec{0}\}$.

So $\vec{x}$ has to be $\vec{0}$. ✓

We showed: $\vec{x}$ in $N(A^TA)$ means $\vec{x}=\vec{0}$, that is, $N(A^TA)=\{\vec{0}\}$,

$\longrightarrow$ A^TA is an invertible $m\times m$ matrix.

Now go back to $(A^TA)\vec{c}=A^T\vec{x}$ (called the "normal equation" for $\vec{c}$)

$\vec{c}=(A^TA)^{-1}A^T\vec{x}$. Now remember that the projection vector was $\vec{v}=A\vec{c}$.

So $\vec{v}=A\vec{c}=\left[A(A^TA)^{-1}A^T\right]\vec{x}$

This is the projection matrix P!

So if you want to project $\vec{x}$ onto V:

1. Find a basis for V: $\{\vec{v}_1,\vec{v}_2,\dots,\vec{v}_m\}$
2. Form $A=\begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m\end{bmatrix}$
3. Project matrix is $P=A(A^TA)^{-1}A^T$
4. Project $\vec{x}$: $\vec{v}=P\vec{x}$

Example $V=$ plane in $\mathbb{R}^3$ with equation $x+2y+3z=0$

1. Find basis: $V=$ all $\begin{bmatrix}x\\y\\z\end{bmatrix}$ with $x=-2y-3z$

$$=\text{all}\begin{bmatrix}-2y-3z\\y\\z\end{bmatrix}=\text{all } y\begin{bmatrix}-2\\1\\0\end{bmatrix}+z\begin{bmatrix}-3\\0\\1\end{bmatrix}$$

linearly independent spanning set, so a basis for V

2. $A=\begin{bmatrix}-2 & -3\\1 & 0\\0 & 1\end{bmatrix}$ Notice: A is not invertible, but A^TA is!

3. $A^TA=\begin{bmatrix}-2 & 1 & 0\\-3 & 0 & 1\end{bmatrix}\begin{bmatrix}-2 & -3\\1 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}5 & 6\\6 & 10\end{bmatrix}$, $(A^TA)^{-1}=\underbrace{\dfrac{1}{50-36}}_{1/14}\begin{bmatrix}10 & -6\\-6 & 5\end{bmatrix}$

$$P = A(A^TA)^{-1}A^T = \begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{14}\begin{bmatrix} 10 & -6 \\ -6 & 5 \end{bmatrix}\begin{bmatrix} -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$= \frac{1}{14}\begin{bmatrix} -2 & -3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -2 & 10 & -6 \\ -3 & -6 & 5 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{bmatrix}$$

Should look familiar!

④ Let's project $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ onto V: $\vec{v} = P\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 13 \\ -2 \\ -3 \end{bmatrix}$

Here, V is the orthogonal complement of the line from earlier example:

So our new P is actually $I - (\text{old } P)$

for the line.

Question: What is the rank of our new P?

$\text{rank}(P) = \dim$ of $C(P)$ ← all vectors like $P\vec{x}$

= all vectors in V

$= \dim V = 2$

In general: rank of a projection matrix = dimension of the subspace you are projecting onto.