Orthogonality & Projections

"orthogonal" = "perpendicular"

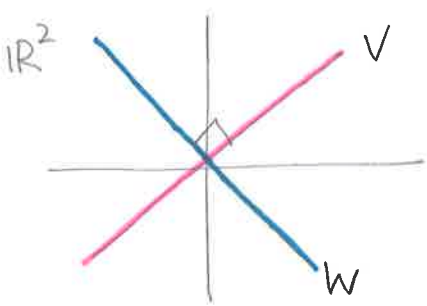
Two vectors $\vec{v}, \vec{w}$ in $\mathbb{R}^n$ are orthogonal if

$$\underset{n\times 1}{\vec{v}} \cdot \underset{n\times 1}{\vec{w}} = 0 \iff \underset{1\times n}{\vec{v}^T} \underset{n\times 1}{\vec{w}} = 0$$

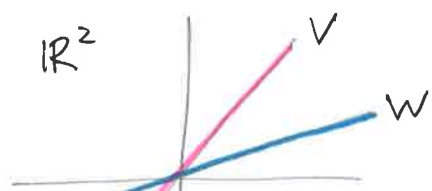
Two subspaces $V, W \subset \mathbb{R}^n$ are orthogonal if

$$\vec{v} \cdot \vec{w} = 0 \quad \text{for all } \vec{v} \text{ in } V, \vec{w} \text{ in } W$$

We also write $V \perp W$ when V, W are orthogonal.



$V \perp W$



~~$V \perp W$~~

Useful criterion:

if V has basis $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$
W has basis $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_\ell$
then $V \perp W \iff \vec{v}_i \cdot \vec{w}_j = 0$ for $1 \le i \le k$, $1 \le j \le \ell$

Given $V \subset \mathbb{R}^n$ define its orthogonal complement:

$$V^\perp = \{ \vec{w} \text{ in } \mathbb{R}^n \mid \vec{w} \perp \vec{v} \text{ for all } \vec{v} \text{ in } V \}$$

$V^\perp \subset \mathbb{R}^n$ is a subspace of $\mathbb{R}^n$

Proof:
- has the zero vector since $\vec{0} \cdot \vec{v} = 0$ always
- if $\vec{w}_1, \vec{w}_2$ are in $V^\perp$ then $\vec{w}_1 \cdot \vec{v} = 0$ and $\vec{w}_2 \cdot \vec{v} = 0$ for all $\vec{v}$ in V. Then $(\vec{w}_1 + \vec{w}_2) \cdot \vec{v} = \vec{w}_1 \cdot \vec{v} + \vec{w}_2 \cdot \vec{v} = 0$ for all $\vec{v}$ in V, so we conclude $\vec{w}_1 + \vec{w}_2$ is in $V^\perp$.
  Therefore $V^\perp$ is closed under $+$.
- Closure under scalar multiplication is a similar arg.

QED

$V, W \subset \mathbb{R}^n$ subspaces. If $V \perp W$ then $V \cap W = \{\vec{0}\}$

Proof: Suppose $\vec{v}$ is in $V \cap W$, so $\vec{v}$ is in V and W.

$V \perp W$ means every vector in V is orthogonal to every vector in W.

In particular, $\vec{v} \cdot \vec{v} = 0$.

But $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2 = 0$ so it must be that $\vec{v} = \vec{0}$.

QED

Let A be an $m \times n$ matrix.

The rowspace of A, denoted $R(A) \subset \mathbb{R}^n$, is orthogonal to the null space $N(A) \subset \mathbb{R}^n$, i.e.

$$R(A) \perp N(A).$$

Proof: Let $\vec{v}_1, \ldots, \vec{v}_m$ be the rows of A

$$A = \begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix}$$

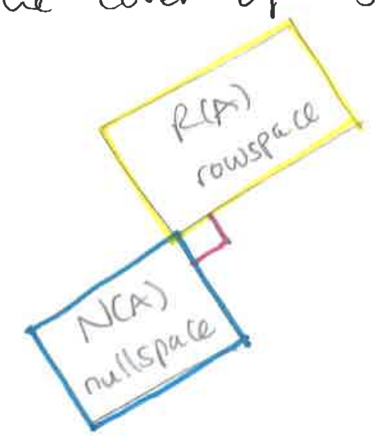
Let $\vec{w}$ be in $N(A)$, i.e. $A\vec{w} = \vec{0}$. Then

$$\vec{0} = A\vec{w} = \begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix} \vec{w} = \begin{bmatrix} \vec{v}_1 \cdot \vec{w} \\ \vdots \\ \vec{v}_m \cdot \vec{w} \end{bmatrix}$$

So $\vec{w} \perp \vec{v}_i$ for $i = 1, \ldots, m$.

$\vec{w}$ was an arbitrary vector in $N(A)$
and the $\vec{v}_i$'s form a basis of $R(A)$
so we've shown $R(A) \perp N(A)$. QED

This last proposition explains the "right angle" on the cover of Strang's textbook:



$$R(A) + N(A) = \mathbb{R}^n$$

$$R(A) \perp N(A)$$

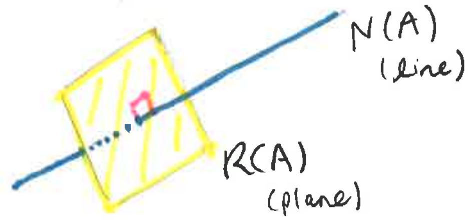
$$R(A) \cap N(A) = \{\vec{0}\}$$

Example $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 5 \end{bmatrix} \xrightarrow{\text{elim.}} \begin{bmatrix} \boxed{1} & 0 & 2 \\ 0 & \boxed{1} & -3 \\ 0 & 0 & 0 \end{bmatrix}$ RREF

rowspace $R(A) = \text{span}\left(\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right)$ (note $A = A^T$ so rowspace = column space)

solving $A\vec{x} = \vec{0}$ gives $\vec{x} = t\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$ for any $t \in \mathbb{R}$

thus $N(A) = \text{span}\left(\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}\right)$.



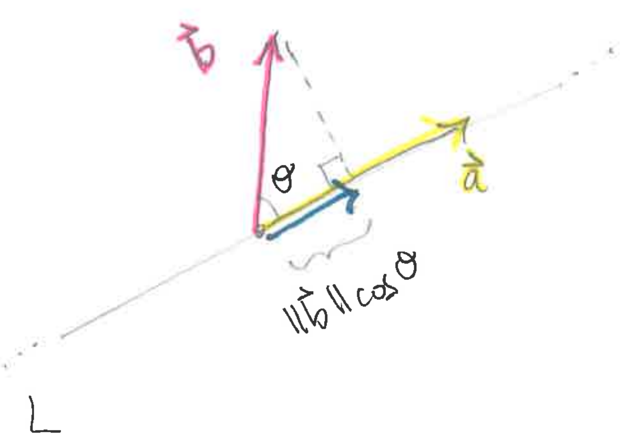
We check $\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 0$, $\begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 0$ verifying $R(A) \perp N(A)$.

Projections

Let $\vec{b}$ be any vector in $\mathbb{R}^n$

Suppose we're given a line $L = \text{span}(\vec{a})$, where $\vec{a} \neq \vec{0}$.

What is the projection of $\vec{b}$ onto L?



$$\|\vec{a}\| \|\vec{b}\| \cos\theta = \vec{a} \cdot \vec{b}$$

$$\|\vec{b}\| \cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|} \quad (*)$$

The projected vector is

$$\overset{\text{scalar}}{\left(\|\vec{b}\|\cos\theta\right)}\overset{\text{vector}}{\left(\text{unit vector in direction of } \vec{a}\right)}$$

$$= \left(\|\vec{b}\|\cos\theta\right)\left(\frac{\vec{a}}{\|\vec{a}\|}\right)$$

$$= \left(\frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|}\right)\left(\frac{\vec{a}}{\|\vec{a}\|}\right) = \overset{\text{scalar}}{\left(\frac{\vec{a}\cdot\vec{b}}{\vec{a}\cdot\vec{a}}\right)}\overset{\text{vector}}{\vec{a}}$$

where we used $(*)$ and also $\|\vec{a}\|^2 = \vec{a}\cdot\vec{a}$.

Define a matrix P as follows: "The projection matrix associated to $\vec{a}$"

$$\boxed{P = \frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}} = \underbrace{\left(\frac{1}{\vec{a}\cdot\vec{a}}\right)}_{\text{scalar}}\vec{a}\vec{a}^T$$

Note $\underset{n\times 1}{\vec{a}}\,\underset{1\times n}{\vec{a}^T}$ is an $n\times n$ matrix,

$\underset{1\times n}{\vec{a}^T}\,\underset{n\times 1}{\vec{a}} = \vec{a}\cdot\vec{a}$ is a scalar, so P is $n\times n$.

For any $\vec{b}$ in $\mathbb{R}^n$ we compute:

$$P\vec{b} = \left(\frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}\right)\vec{b} = \underbrace{\left(\frac{1}{\vec{a}^T\vec{a}}\right)}_{\text{scalar}}\underbrace{(\vec{a}\vec{a}^T)}_{\text{matrix}}\vec{b}$$

$$= \left(\frac{1}{\vec{a}^T\vec{a}}\right)\vec{a}\,(\vec{a}^T\vec{b}) = \left(\frac{\vec{a}^T\vec{b}}{\vec{a}^T\vec{a}}\right)\vec{a} =$$ projection of $\vec{b}$ onto $L = \mathrm{span}(\vec{a})$.

Can thus view P as a linear transformation:

$P: \mathbb{R}^n \to \mathbb{R}^n$

input any vector $\vec{b}$ in $\mathbb{R}^n$ $\xrightarrow{P}$ output $P(\vec{b})$ is the projection of $\vec{b}$ onto $\mathrm{span}(\vec{a})$

Note $P^2 = P$:

$$P^2 = PP = \left(\frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}\right)\left(\frac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}\right) = \frac{\vec{a}\,\cancel{(\vec{a}^T\vec{a})}\,\vec{a}^T}{\cancel{(\vec{a}^T\vec{a})}(\vec{a}^T\vec{a})} = \frac{\vec{a}\vec{a}^T}{\vec{a}^T a} = P$$

"Projecting twice gives the same answer."

<u>Ex.</u> Give the projection matrix P that projects onto the line $\mathrm{span}\left(\begin{bmatrix}1\\2\\-1\end{bmatrix}\right) \subset \mathbb{R}^3$. Use it to project $\vec{v} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ onto this line.

Here $\vec{a} = \begin{bmatrix}1\\2\\-1\end{bmatrix}$. Compute $P = \dfrac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}}$.

$$\vec{a}^T\vec{a} = \vec{a}\cdot\vec{a} = \begin{bmatrix}1\\2\\-1\end{bmatrix}\cdot\begin{bmatrix}1\\2\\-1\end{bmatrix} = 1^2 + 2^2 + (-1)^2 = 6.$$

$$\vec{a}\vec{a}^T = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{bmatrix}$$

Thus $P = \dfrac{\vec{a}\vec{a}^T}{\vec{a}^T\vec{a}} = \dfrac{1}{6}\begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{bmatrix}$.

Project $\vec{v}$ onto $\text{span}(\vec{a})$:

$$P(\vec{v}) = \frac{1}{6}\begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\vec{v}} = \frac{1}{6}\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/3 \\ -1/3 \end{bmatrix}.$$

Projection onto a subspace

Consider a subspace $V \subset \mathbb{R}^m$ with basis $\vec{a}_1, \dots, \vec{a}_n$.

In particular $V = \text{span}(\vec{a}_1, \dots, \vec{a}_n)$, $\dim V = n$,

Let $A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{bmatrix}$, so A is $m \times n$

Projection matrix: (associated to basis $\vec{a}_1, \dots, a_n$)

$$\boxed{P = A(A^TA)^{-1}A^T}$$

note P is $m \times m$:

$$\underset{n\times m}{A^T}\underset{m\times n}{A} \rightarrow A^TA \text{ is } n\times n, \text{ so is } (A^TA)^{-1}$$

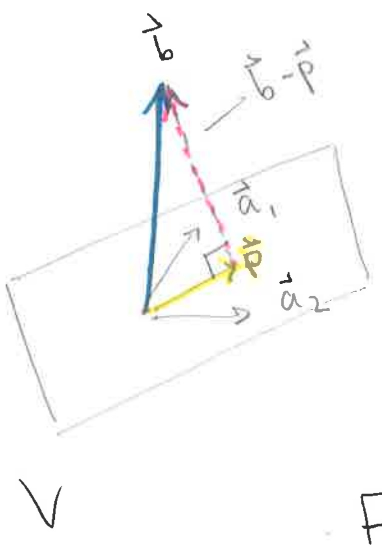
then $$P = \underset{m\times \cancel{n}}{A}\ \underset{\cancel{n}\times \cancel{n}}{(A^TA)^{-1}}\ \underset{\cancel{n}\times m}{A^T} \quad \text{is } m\times m.$$

Why is A^TA invertible?

Fact: $\boxed{A^TA \text{ is invertible} \iff \text{columns of } A \text{ are independent.}}$

For us the columns of A are a basis, so all is well.

Claim: $P: \mathbb{R}^m \to \mathbb{R}^m$ is the linear transformation that takes any vector $\vec{b}$ in $\mathbb{R}^m$ and outputs the vector $P(\vec{b})$ which is the projection of $\vec{b}$ onto V.



$\vec{p}$ = proj. of $\vec{b}$ onto V

$\vec{p} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = A\vec{x}$

where $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ to be determined.

From picture we see we should have $\vec{b}-\vec{p} \perp V$

$$\Leftrightarrow \quad \vec{b} - \vec{p} \perp \vec{a}_i \quad \text{for} \quad i = 1, \dots, n$$

$$\Leftrightarrow \quad (\vec{b} - \vec{p}) \cdot \vec{a}_i = 0 \quad \text{for} \quad i = 1, \dots, n$$

$$\Leftrightarrow \quad a_i^T (\vec{b} - \vec{p}) = 0 \quad \text{for} \quad i = 1, \dots, n$$

$$\Leftrightarrow \quad \underbrace{\begin{bmatrix} \text{---} \ \vec{a}_1 \ \text{---} \\ \vdots \\ \text{---} \ \vec{a}_n \ \text{---} \end{bmatrix}}_{A^T} \begin{bmatrix} \vec{b} - A\vec{x} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \quad A^T(\vec{b} - A\vec{x}) = \vec{0} \quad \Leftrightarrow \quad A^T\vec{b} = A^T A \vec{x}.$$

Multiply both sides by $(A^T A)^{-1}$ to get

$$(A^T A)^{-1} A^T \vec{b} = \vec{x}$$

So we obtain projected vector $\vec{p} = Ax$ is given by

$$A(A^T A)^{-1} A \vec{b} = P\vec{b},$$

as claimed!