Last time

- Orthogonal matrices $\qquad Q^{-1} = Q^T$
- Transposes

Warning If Q is $m \times n$ with orthonormal columns, then $Q^T Q = I_m$, but Q is not invertible unless square.

The transpose also permits a clean formula for orthogonal projection.

Given an ONB $\vec{u}_1, \ldots, \vec{u}_r$ for V, we have

$$\operatorname{proj}_V(\vec{x}) = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_r \cdot \vec{x})\vec{u}_r$$

$$= \vec{u}_1 \vec{u}_1^T \vec{x} + \cdots + \vec{u}_r \vec{u}_r^T \vec{x}$$

$$= (\vec{u}_1 \vec{u}_1^T + \cdots + \vec{u}_r \vec{u}_r^T)\vec{x}$$

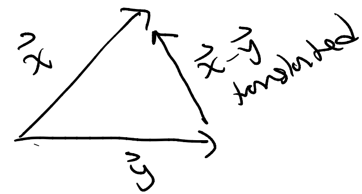
$$= QQ^T \vec{x},$$

where $Q = [\vec{u}_1 \cdots \vec{u}_r]$.

$$\boxed{\operatorname{proj}_V = QQ^T \qquad Q = \begin{bmatrix} | & & | \\ \vec{u}_1 & \cdots & \vec{u}_r \\ | & & | \end{bmatrix}}$$

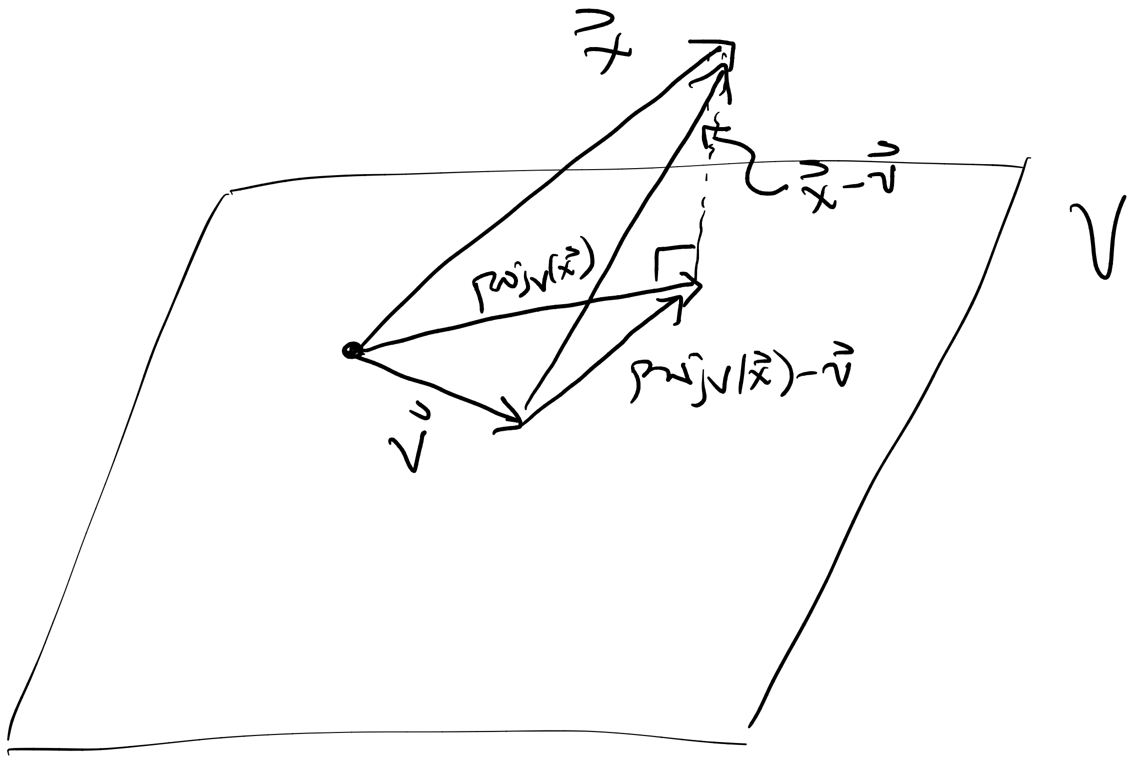
Orthogonal projection is important because it minimizes distance.

Def The *distance* between two vectors $\vec{x}$ and $\vec{y}$ is the length of their difference $|\vec{x}-\vec{y}|$.



Pythagoras $|\vec{x}-\vec{y}|^2 = |\vec{x}|^2 + |\vec{y}|^2$ if $\vec{x}$ and $\vec{y}$ are orthogonal.

Then $\text{proj}_V(\vec{x})$ is the vector in V <u>closest</u> to $\vec{x}$, i.e., $|\vec{x} - \text{proj}_V(\vec{x})| < |\vec{x} - \vec{v}|$ for every $\vec{v}$ in V different from $\text{proj}_V(\vec{x})$.



By Pythagoras,

$$|\vec{x}-\vec{v}|^2 = |\vec{v}-\text{proj}_V(\vec{x})|^2 + |\vec{x}-\text{proj}_V(\vec{x})|^2$$
$$\geq |\vec{x}-\text{proj}_V(\vec{x})|^2$$

with equality if and only if $|\vec{v}-\text{proj}_V(\vec{x})|^2=0$.

Q What to do if $A\vec{x}=\vec{b}$ has no solution?

A Minimize the error $|A\vec{x}-\vec{b}|$.

Def Let A be an $m\times n$ matrix. A vector $\vec{x}^*$ in $\mathbb{R}^n$ is called a least-squares solution to the system $A\vec{x}=\vec{b}$ if

$$|\vec{b}-A\vec{x}^*| \leq |\vec{b}-A\vec{x}|$$

for every vector $\vec{x}$ in $\mathbb{R}^n$.

The collection of vectors $A\vec{x}$ as $\vec{x}$ varies is just the image of A, so we want to minimize the distance from $\vec{b}$ to $\text{im}(A)$. In other words, $\vec{x}^*$ is a least-squares solution precisely when

$$A\vec{x}^* = \text{proj}_{\text{im}(A)}(\vec{b}),$$

i.e., when $\vec{b} - A\vec{x}^*$ is orthogonal to $\text{im}(A)$.

Def The <u>orthogonal complement</u> of the subspace V of $\mathbb{R}^n$ is the collection $V^\perp$ of all vectors $\vec{x}$ in $\mathbb{R}^n$ orthogonal to every vector in V.

<u>Facts about $V^\perp$</u>

(1) $V^\perp$ is a subspace

(2) The only vector in V and in $V^\perp$ is $\vec{0}$

(3) $\dim V + \dim V^\perp = n$

(4) $(V^\perp)^\perp = V$

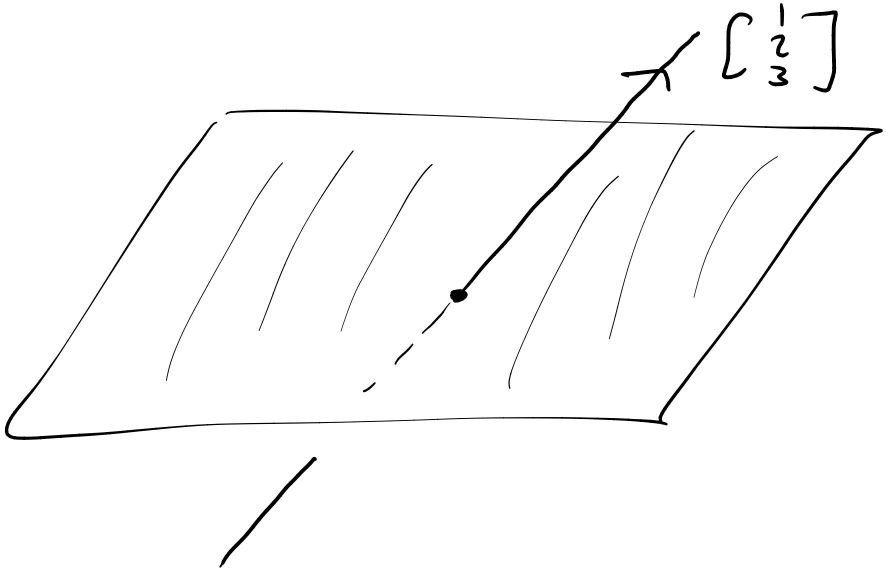
(5) Every vector $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$ with $\vec{x}^{\parallel}$ in V and $\vec{x}^{\perp}$ in $V^\perp$, uniquely.

So $\vec{x}^*$ is a LSS to $A\vec{x}=\vec{b}$ precisely when $\vec{b}-A\vec{x}^*$ is in $\text{im}(A)^\perp$. Since the columns of A span $\text{im}(A)$, this condition is equivalent to being orthogonal to every column of A, which is to say dotting to 0 with every row of A^T.

$$\boxed{\text{im}(A)^\perp = \ker(A^T)}$$

<u>Ex</u> If V is the line in $\mathbb{R}^3$ parallel to $\begin{bmatrix}1\\2\\3\end{bmatrix}$, then $V=\text{im}(A)$, where $A=\begin{bmatrix}1\\2\\3\end{bmatrix}$,

while $\ker(A^T) = \ker([1\ 2\ 3])$ is the plane described by the equation $x+2y+3z=0$



In summary, $\vec{x}^*$ is a LSS if and only if $A^T(\vec{b}-A\vec{x}^*)=\vec{0}$.

Normal equation: The least-squares solutions of the system $A\vec{x} = \vec{b}$ are the solutions of the system

$$\boxed{A^T A\vec{x} = A^T\vec{b}}$$