A $n\times m$ $\quad A: \mathbb{R}^m \longrightarrow \mathbb{R}^n$

A^T $m\times n$ $\quad A^T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$

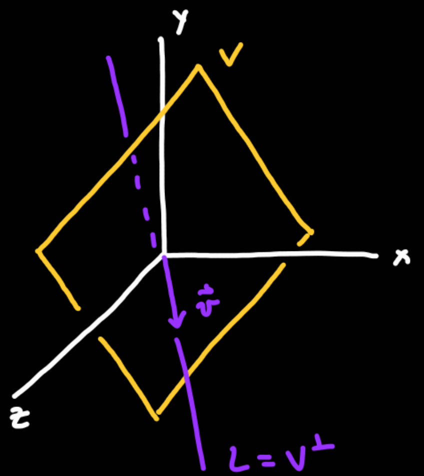
$$(MN)^T = N^T M^T$$
$$(MN)^{-1} = N^{-1} M^{-1}$$

$A^T A: \mathbb{R}^m \longrightarrow \mathbb{R}^m$

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

Example: $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ projection onto $V = \mathrm{span}\left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right)$.

$$A = \begin{bmatrix} 2/3 & 1/3 & -1/3 \\ 1/3 & 2/3 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix} \qquad A^T = A$$



V is given by $x - y + z = 0$

$\vec{v} = \begin{bmatrix}1\\-1\\1\end{bmatrix}$

$V = \mathrm{im}(A)$

$L = V^\perp = (\mathrm{im}(A))^\perp$

$\ker(A^T) = (\mathrm{im}(A))^\perp = V^\perp = L$

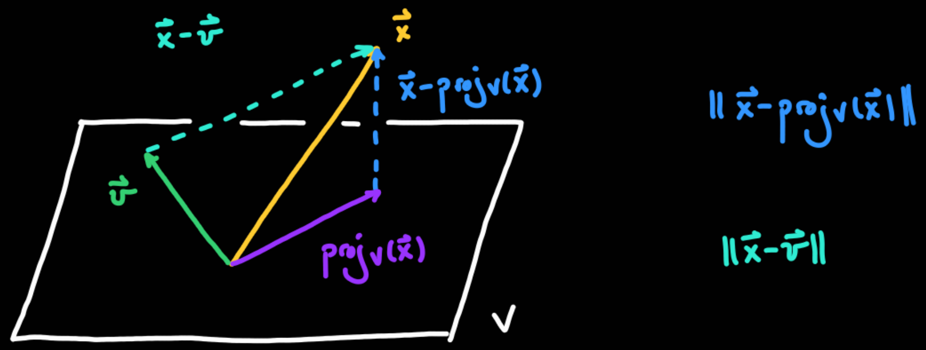
To compute $\ker(A^T)$ we solve the equation $A\vec{x} = \vec{0}$.

$$\begin{bmatrix} 2/3 & 1/3 & -1/3 \\ 1/3 & 2/3 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix} \qquad \vec{x} = t\cdot\begin{bmatrix}1\\-1\\1\end{bmatrix} = t\cdot\vec{v}$$

So $\ker(A^T) = \mathrm{span}(\vec{v}) = L = V^\perp = (\mathrm{im}(A))^\perp$.

If A is an orthogonal matrix, then $A^T A = I$

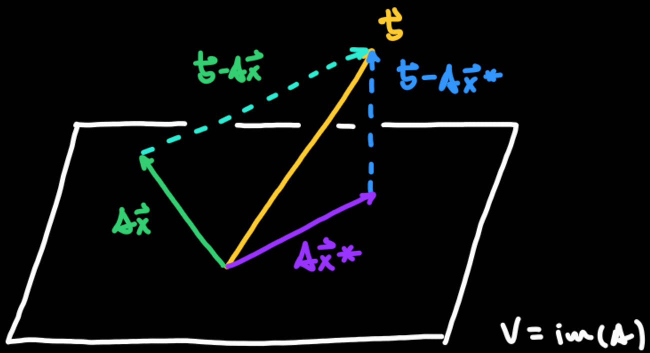
Recall:



Projection onto a subspace is solving a minimization problem: let $\vec{v}$ in V, compute the distance from $\vec{v}$ to $\vec{x}$, choose the one with minimum distance.

$$\|\vec{x}-\text{proj}_V(\vec{x})\| \leq \|\vec{x}-\vec{v}\| \quad \text{for } \vec{v} \text{ in } V.$$

Given A an $n\times m$ matrix, $\vec{b}$ a vector in $\mathbb{R}^n$, we want to solve $A\vec{x}=\vec{b}$.



If $\vec{b}$ is in $\text{im}(A)$ then there exists $\vec{x}$ in $\mathbb{R}^m$ such that $A\vec{x}=\vec{b}$.

A vector $\vec{x}^*$ in $\mathbb{R}^m$ is called a <u>least-squares solution</u> to $A\vec{x}=\vec{b}$ if:

$$\|\vec{b}-A\vec{x}^*\| \leq \|\vec{b}-A\vec{x}\| \quad \text{for all } \vec{x} \text{ in } \mathbb{R}^m.$$

Now, from $A\vec{x}=\vec{b}$ we can produce the equation $A^TA\vec{x}=A^T\vec{b}$. This will always be a consistent system, it is known as the <u>normal equation</u> of $A\vec{x}=\vec{b}$.

Theorem: A $n\times m$, $\vec{b}$ $\mathbb{R}^n$, if $\ker(A) = \{\vec{0}\}$ then the system $A\vec{x} = \vec{b}$

has a unique least squares solution: $\vec{x}^* = (A^TA)^{-1}A^T\vec{b}$

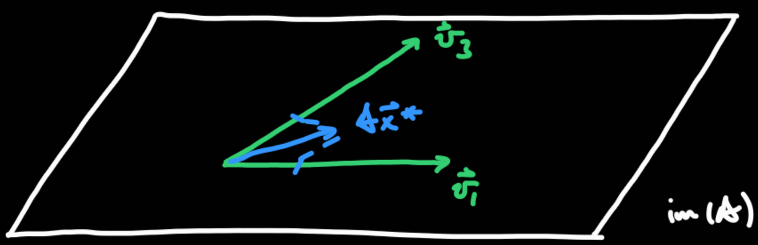
Example: $A = \begin{array}{c} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \\ \vec{v}_1 \quad \vec{v}_3 \end{array}$ $\quad \text{im}(A) = \text{span}(\vec{v}_1, \vec{v}_3) \quad$ $\vec{b} = \begin{array}{c} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ \vec{v}_2 \end{array}$ $\quad \ker(A) = \{\vec{0}\}$

$$\vec{x}^* = \left(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 1 \end{bmatrix} =$$

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

$$A\vec{x}^* = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \frac{1}{3}\vec{v}_1 + \frac{1}{3}\vec{v}_3$$

$$P\vec{b} = \underbrace{\begin{bmatrix} 2/3 & 1/3 & -1/3 \\ 1/3 & 2/3 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix}}_{\text{matrix encoding the projection onto } V}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{3}\vec{v}_1 + \frac{1}{3}\vec{v}_3$$



Theorem: V has basis $\vec{v}_1, \dots, \vec{v}_m$ then the matrix encoding the orthogonal

projection onto V is:

$$P = A(A^TA)^{-1}A^T = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}\left(\begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix}\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}\right)^{-1}\begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix}.$$

If $\vec{v}_1, \dots, \vec{v}_m$ are an orthonormal basis then $\underbrace{\begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix}}_{A^T}\underbrace{\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}}_{A} = I_m$

so $(A^T A) = I_n$ so $P = A A^T$.