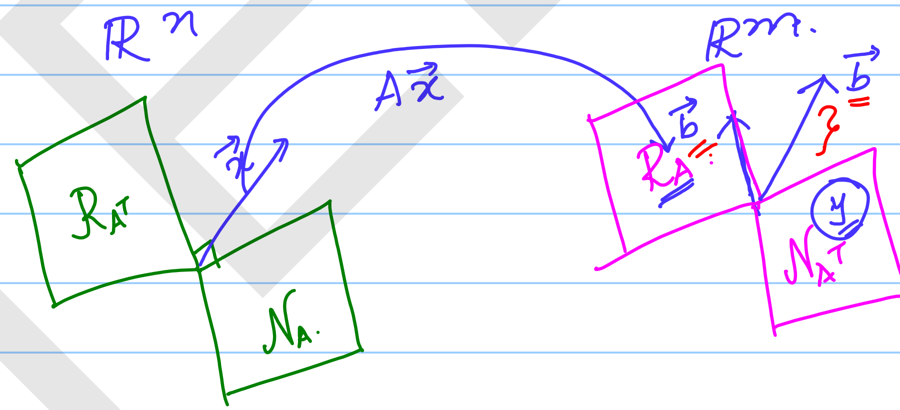
Recall: $R_{A^T} \perp N_A$ in $\mathbb{R}^n$

$R_A \perp N_{A^T}$ in $\mathbb{R}^m$.

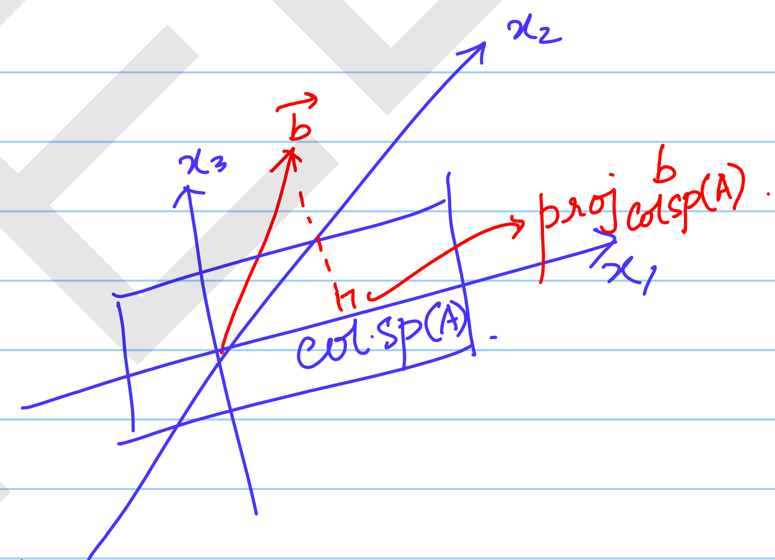
for any matrix $A \in \mathbb{R}^{m \times n}$.



Suppose $\vec{b}$ is not in the Col. Sp (A). $\Rightarrow$ No Scalars $x_1, x_2 \cdots x_n$ exist s.t

$$\vec{b} = x_1 A_1 + \cdots + x_n A_n$$

A_i: i^{th} col of $A^{m \times n}$



$\vec{b}$ does not belong to the Col. Sp(A).

$\vec{b} \in \mathbb{R}^m$, $\vec{b} \notin$ Col. Sp(A).

Since $\vec{b}$ does not belong to the Col. Sp(A), we would like to approx $\vec{b}$ in the Col. Sp(A)

$\Rightarrow$ We want to find that vector b^* which best approximates $\vec{b}$ in the Col. Sp(A).

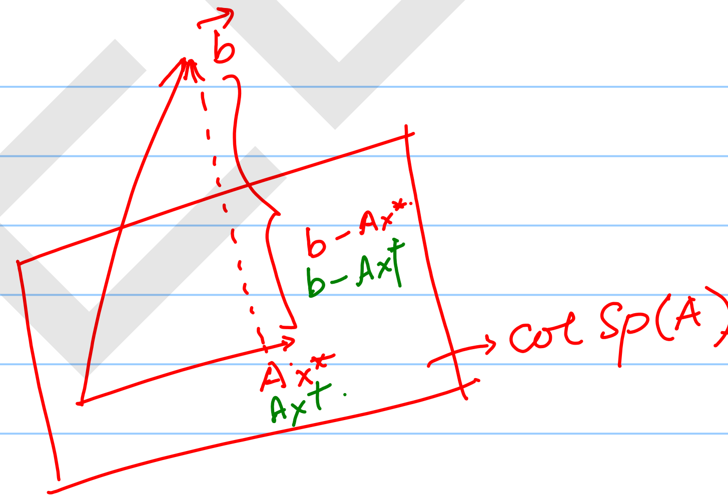
Note that for whatever x, we have

$$\|b - Ax\| \text{ will never be } 0.$$

because no $(x_1, \ldots, x_n)$ exists s.t $Ax = b$.

So we want to make $\|b - Ax\|$ as small as possible.

If we can find a unique vector x^* that minimizes the length of $b - Ax$, we call x^* to be the best optimal soln to $Ax = b$ in the sense of LEAST SQUARES.



$\|b - Ax\|$ is the smallest when Ax^* is the orthogonal proj of $\vec{b}$ onto the col. sp(A).

$(b - Ax^*)$ is orthogonal to the col Sp(A).

We want $(b - Ax^*)$ to be orthog. to every vector in the col. Sp(A).

Any vector in the col. sp$(A) = \underline{\underline{Ay}}$ for $y \in \mathbb{R}^n$.

$\Rightarrow \widetilde{(b - Ax^*)} \cdot \tilde{Ay} = 0$.

$= (Ay)^T (b - Ax^*) = 0$

$= y^T A^T (b - Ax^*) = 0$.

$= y^T \underbrace{(A^T b - A^T A x^*)}_{\uparrow} = 0$.

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

Since the only vector which is orthogonal to every vector in $\mathbb{R}^n$ is the zero vector, we have

$A^T b - A^T A x^* = 0.$

$\Rightarrow \quad A^T A x^* = A^T b.$

$A \in \mathbb{R}^{m \times n} \qquad A^T \in \mathbb{R}^{n \times m}$

$A^T A$: Square matrix of dim. $n \times n$.

Suppose the cols of A are all linearly indep $\Rightarrow$ Rank of $A = n$.

$A^T A$ will be invertible.

$\Rightarrow \boxed{x^* = (A^T A)^{-1} A^T b.}$

$x^* = (A^T A)^{-1} A^T b$ is the best optimal soln to $Ax = b$.

$\boxed{x^{\dagger} = (A^T A)^{-1} A^T b} \qquad x^{\dagger} =$ x dagger

We wanted to Solve

$Ax = b.$

But if $b \notin$ col sp(A) no soln exists. Therefore we solve for $Ax^{\dagger} = b^*$, where

b^* = Orthog Projection of b onto col sp(A).

$$x^{\dagger} = (A^TA)^{-1}A^T b. \longrightarrow (1)$$

Suppose we have A which is an invertible matrix then given $Ax = b$, we get

$$x = A^{-1} b. \longrightarrow (2)$$

$(A^T A)^{-1} A^T \longrightarrow$ Pseudo inverse of A.

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

Look at

$$\underbrace{A x^\dagger}_{\text{Projection of } \vec{b} \text{ onto the column space of } A.} = \Big(A \underbrace{(A^T A)^{-1} A^T}\Big) b$$

$\therefore$ The matrix $A(A^T A)^{-1} A^T$ is the matrix that projects every $\vec{b}$ onto the col. sp(A).

Question: How do we prove that $A(A^TA)^{-1}A^T$ is actually a projection matrix?

Question 2: What is a projection matrix?

Question 3: What are the properties of a projection matrix?