Consider the following example in order to get some intuition (due to Strang):

ex) Solve the LSQ, for $\tilde{x}$ the least squares estimate solution to $Ax=b$, where

$$A=\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad b=\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

We note that $\text{row}(A)$ is spanned by first two rows of A.

First, let's project b to $\text{col}(A)$: $Pb=\begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}$, by inspection.

So $b=Pb+r$, where $r=\begin{pmatrix} 0 \\ 0 \\ b_3 \end{pmatrix}$ and see that $r \perp Pb$.

$A\tilde{x}=Pb$, where $\tilde{x}=\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \\ \tilde{x}_4 \end{pmatrix}$, is

$$\begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \\ \tilde{x}_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}$$

So $\hat{x}_1 = b_1/\sigma_1 \quad \hat{x}_2 = b_2/\sigma_2$

but $\hat{x}_3$ & $\hat{x}_4$ are arbitrary.

$$\|\hat{x}\| = \sqrt{\tilde{x}_1^2 + \tilde{x}_2 + \tilde{x}_3^2 + \tilde{x}_4^2}$$

The smallest we can make $\hat{x}$ is by setting $\hat{x}_3 = \tilde{x}_4 = 0$.

Let $\boxed{\min \|\hat{x}\| \equiv x^+}$

which means that, if the general solution $\tilde{x} = x_r + x_n$

where $x_r \in \text{row}(A)$ and $Ax_n = 0$

then $\hat{x} = x^+ = x_r$ with x_n set to 0.

The minimum norm solution of $A\hat{x} = Pb$

is x^+

$$x^+ = \begin{bmatrix} b_1/\sigma_1 \\ b_2/\sigma_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sigma_1 & 0 & 0 \\ 0 & 1/\sigma_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \equiv A^+ b$$

A^+ is the pseudoinverse of A

if $A \in F^{m \times n}$ then $A^+ \in F^{n \times m}$

$$x^+ = A^+ b$$

Note: $(A^+)^+ = A$, where A is invertible or not.

Another example, the case when A^{-1} is invertible:

let $Ax = b$, $Pb = b$

and $x = A^{-1} b$

in this case $x^+ = x$ (smallest norm solution)

$A^+ = A^{-1}$

//

General Case: x^+ is always in row(A)

Recall that $\hat{x} = x_r + x_h$

where
$x_r \in \text{row}(A)$
$x_h \in \text{Null}(A)$, generally

for $A\hat{x} = Pb$

1. Row space component solves $Ax_r = Pb$
2. $x_r \perp x_h$ so

$$\|\hat{x}\|^2 = \|x_r\|^2 + \|x_h\|^2$$

and smallest $\|\hat{x}\|^2$ when $\|x_h\|^2 = 0$

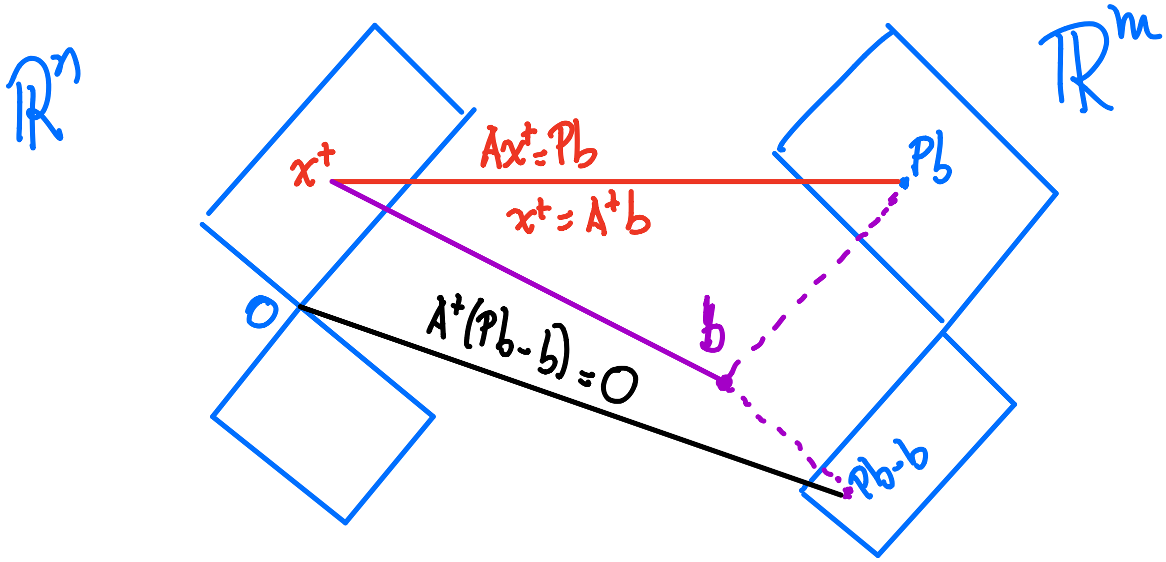
i.e. $x_h = 0$.

3. All solutions of $A\hat{x} = Pb$ have same row space component x_r.

$$x_r = x^+$$

//

The action of the Pseudo Inverse:



A^+ inverts "whatever is invertible", i.e. the part associated with the rowspace of A.

A^+ knocks out the $\text{null}(A^T)$ by sending it to zero & knocks out $\text{null}(A)$ by choosing $x^+ = x_r$ //

For $A \in \mathbb{F}^{m \times n}$ $\quad A^+ \in \mathbb{F}^{n \times m}$

because it takes $b \in \mathbb{F}^m$ back to $x \in \mathbb{F}^n$

such that $\dim(x^+) = \text{rank}(A)$.

ex) $Ax=b$ is $-x_1+2x_2+2x_3=18$

so $A=[-1\ 2\ 2] \in \mathbb{R}^{1\times 3}$

$Pb=18$

$A^+[18]=x^+=\begin{bmatrix}-2\\4\\4\end{bmatrix}$

So $A^+=[-1\ 2\ 2]^+=\begin{bmatrix}-1/9\\2/9\\2/9\end{bmatrix}$

$col(A^+)=row(A)$

//

Using SVD to find A^+: Take $A\in\mathbb{R}^{m\times n}$

Use SVD to find

$$A=Q_1\Sigma Q_2^T \qquad \Sigma\in\mathbb{R}^{m\times n}$$

then $A^+=Q_2\Sigma^+Q_1^T \qquad \Sigma^+\in\mathbb{R}^{n\times m}$

Σ has singular values $\sigma_1, \sigma_2, \sigma_3 \dots, \sigma_r$ on diagonal

Σ^+ has diagonal entries
$1/\sigma_1, 1/\sigma_2, \dots 1/\sigma_r$

Properties for $A \in \mathbb{F}^{m \times n}$

(A) If A is full (column) rank $[\operatorname{rank}(A) = n \leq m]$

A^+ is LEFT inverse of A

$\therefore \; A^+A = I \in \mathbb{R}^{n \times n}$

and $A^+ = (A^TA)^{-1}A^T$

(B) If A is full (row) rank $[\operatorname{rank}(A) = m \leq n]$

A^+ is RIGHT inverse of A

$AA^+ = I \in \mathbb{R}^{m \times m}$

and $A^+ = A^T(AA^T)^{-1}$

(c) If A is invertible (square)

$$A^+ = A^{-1}$$

and for $Ax = b$ $\quad x = x^+$

//