Question:

If $Ax=b$ has no solution, we can find the least squares solutions $\hat{x}$ of $Ax=b$ (1) by solving $Ax=P$, where P is the projection of b onto the columnspace(A). However, in practice this is not an effective way to do so. Instead, we can form a new linear system $A^TAx = A^Tb$ (2), and a solution of (2) gives us a least squares solution of (1).

How can we prove a solution of (2) is a least squares solution of (1)?

Solution:

If we assume that $A\hat{x} = P$, $\hat{x}$ is the least squares solution of (1),

ω = columnspace(A),

$W^{\perp}$ = the vector space $\perp$ ω

then $b-p \in \omega^{\perp}$

In fact, we can we prove that $\omega^{\perp} = \text{nullspace}(A^T)$, (*)

and hence,
$$A^T(b-p) = 0$$
$$\Rightarrow A^T(b-A\hat{x}) = 0$$
$$\Rightarrow A^Tb = A^TA\hat{x}, \text{ which is, } A^TA\hat{x} = A^Tb$$

Hence, a solution of (2) is a least squares solution of (1).

So, the rest working is about how to show (*).

① $\text{nullspace}(A^T) \subseteq \omega^T$.

We assume $x \in \text{nullspace}(A^T) \Rightarrow A^Tx = 0$.

We assume $A = (a_1 \ a_2 \ \cdots \ a_n) \Rightarrow A^T = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} \cdot x = 0$$

(where 0 is a vector).

$$\Rightarrow \begin{pmatrix} a_1^T \cdot x \\ a_2^T \cdot x \\ \vdots \\ a_n^T \cdot x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

(where 0 is a number).

$\Rightarrow a_i^T \cdot x = 0, \ i = 1, 2, \cdots, n \ \Rightarrow \langle a_i, x \rangle = 0, \ i = 1, 2, \cdots, n.$

$\Rightarrow x \perp a_i, \ i = 1, 2, \cdots, n$

$\Rightarrow x \perp \text{span}\{a_1, a_2, \cdots a_n\} = \text{columnspace}(A)$

$\Rightarrow x \in (\text{columnspace}(A))^\perp = W^\perp$

Recall x is an arbitrary vector in nullspace(A^T),

$\Rightarrow$ nullspace$(A^T) \subseteq W^T$.

② $W^T \subseteq$ nullspace(A^T).

Assume $\forall y \in W^T$, $\Rightarrow y \perp W = \text{columnspace}(A) = \text{span}\{a_1, \cdots, a_n\}$

$\Rightarrow y \perp a_i, \ i = 1, 2, \cdots, n.$

$\Rightarrow \langle a_i^T, y \rangle = 0, \ i = 1, 2, \cdots, n.$

$$\Rightarrow \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} \cdot y = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow A^T \cdot y = 0$$

$\Rightarrow y \in$ nullspace(A^T)

$\Rightarrow W^T \subseteq$ nullspace(A^T).

From ① and ②, $W^T =$ nullspace(A).