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),

$W=$ columnspace (A),

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

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

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

and hence, $A^T(b-p)=0$

$\Rightarrow A^T(b-A\hat{x})=0$

$\Rightarrow A^Tb=A^TA\hat{x}$, 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 nullspace$(A^T) \subseteq W^{\perp}$.

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

We assume $A=(a_1 \ a_2 \ \cdots \ a_n) \Rightarrow A^T=\begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix}$ $\Bigg\} \Rightarrow \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} \cdot x = 0$

(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 $(A^T) \subseteq W^{\perp}$.

② $W^{\perp} \subseteq$ nullspace (A^T).

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

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

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

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

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

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

From ① and ②, $W^{\perp} =$ nullspace (A).