DATE
PAGE

AI61003 Linear Algebra for AI & ML
Assignment 02 - Problem 4

Consider $A \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$.
A has linearly independent column vectors.

$$Ax = x_1 A_1 + x_2 A_2 \ldots\ldots + x_n A_n$$

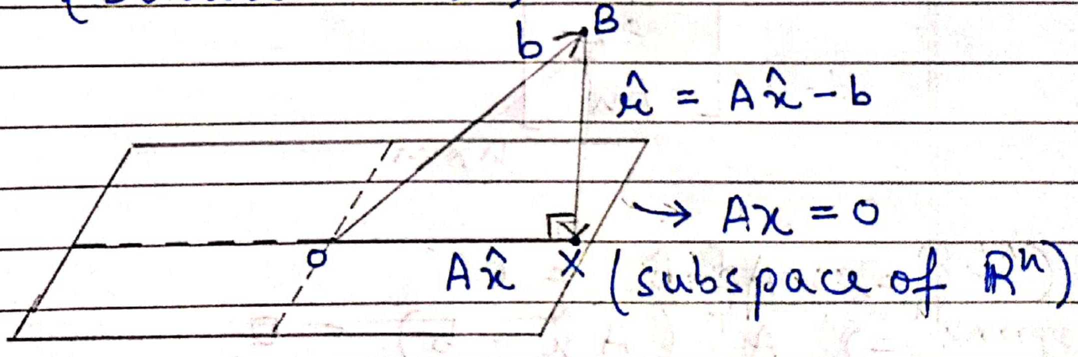
$$= \sum_{i=1}^{n} x_i A_i$$

where $A_i \in \mathbb{R}^m$ is the i^{th} column of A.

The LS solution to the eq^n $Ax - b$ is if $\hat{x}$, then $J = \| A\hat{x} - b \|_2^2$ is minimized.

$$J = \| \hat{x}_1 A_1 + \hat{x}_2 A_2 \ldots\ldots + \hat{x}_n A_n - b \|_2^2 = \|\hat{u}\|_2^2$$

Now $Ax = 0$ is a hyperplane (subspace) in $\mathbb{R}^n$ (shown below).



So geometrically, in order to minimize $\| A\hat{x} - b \|_2^2$ or $\| \hat{u} \|_2^2$, drop a perpe-ndicular from B onto the plane $Ax = 0$. The $\perp^r$ meets the plane at pt X where $A\hat{x} = \overrightarrow{OX}$. Now since $A\hat{x}$ is known, $\hat{x}$ can be derived uniquely because A has linearly independent columns.

Now since $\hat{r}$ is $\perp^r$ to the plane $Ax=0$ it is also $\perp^r$ to all the vectors that lie on the plane.

$\therefore \hat{r} \perp A_i \;, \; 1 \le i \le n$

$\Rightarrow A_i^T \hat{r} = 0 \;, \; 1 \le i \le n$

$$\Rightarrow \begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_n^T \end{bmatrix} \hat{r} = 0$$ (system of n linear eq^{n}s)

But $$\begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_n^T \end{bmatrix}_{n \times m} = A^T$$

$\Rightarrow A^T \hat{r} = 0$

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

$\Rightarrow A^T A \hat{x} = A^T b$ (normal eq^n).

Hence we have geometrically derived the normal eq^n. This is called "normal" because this eq^n is derived from the fact that the residual vector for the LS sol^n

$\hat{r}$ must be normal (perpendicular) to the plane $Ax = 0$.

Now if A does not have linearly independent columns consider the normal eq^n. Here $A^T A$ is the gram matrix of A. Gram matrix of a matrix is invertible if and only if the matrix has linearly independent columns.
If A doesnt have linearly independent columns, $A^T A$ will be non-invertible. Hence, the normal eq^ns will not have a unique sol^n.
★ So ∴ in this case there can be multiple least squares solutions, all of which minimize their residuals.