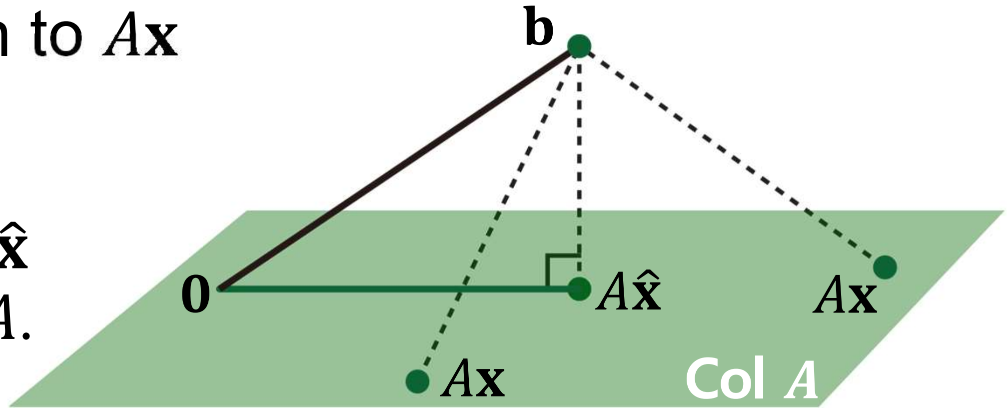
Geometric Interpretation of Least Squares

- The vector $\mathbf{b}$ is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other $\mathbf{x}$.

- To satisfy this, the vector $\mathbf{b} - A\hat{\mathbf{x}}$ should be orthogonal to Col A.

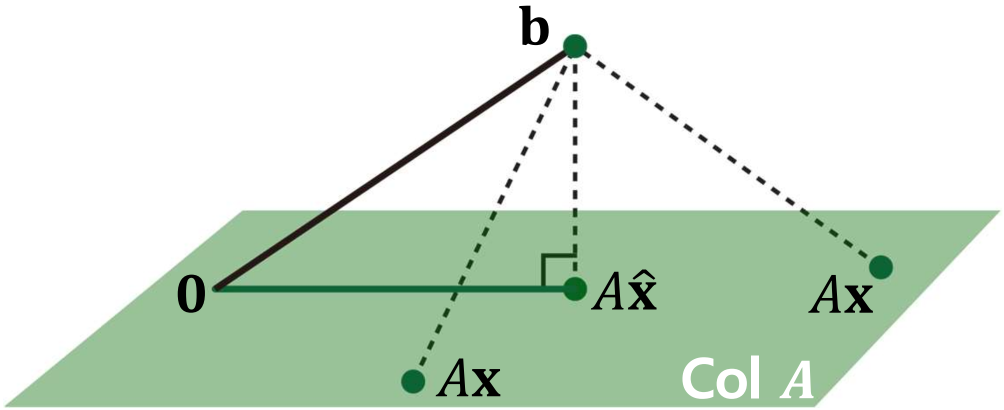


- This means $\mathbf{b} - A\hat{\mathbf{x}}$ should be orthogonal to any vector in Col A:

$$\mathbf{b} - A\hat{\mathbf{x}} \perp (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 \cdots + x_n\mathbf{a}_n) \text{ for any vector } \mathbf{x}$$

Geometric Interpretation of Least Squares

- $\mathbf{b} - A\hat{\mathbf{x}} \perp (x_1\mathbf{a}_1 + x_2\mathbf{a}_2 \cdots + x_n\mathbf{a}_n)$ for any vector $\mathbf{x}$



- Or equivalently,

$$\begin{aligned} &(\mathbf{b} - A\hat{\mathbf{x}}) \perp \mathbf{a}_1 \\ &(\mathbf{b} - A\hat{\mathbf{x}}) \perp \mathbf{a}_2 \\ &\quad\vdots \\ &(\mathbf{b} - A\hat{\mathbf{x}}) \perp \mathbf{a}_n \end{aligned} \Rightarrow \begin{aligned} &\mathbf{a}_1^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \\ &\mathbf{a}_2^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \\ &\quad\vdots \\ &\mathbf{a}_n^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \end{aligned} \Rightarrow A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$$