A an $m \times$ matrix
Orthogonal Decomposition Theorem Applied to subspace $W = \text{Col}\, A$

Suppose $\boldsymbol{b}$ is in $\mathbb{R}^m$; $W = \text{Col}\, A$ is a subspace of $\mathbb{R}^m$.

Let $\{\boldsymbol{v_1}, \boldsymbol{v_2}, ..., \boldsymbol{v_p}\}$ be an orthogonal basis for Col A.

We can write $\boldsymbol{b} = \widehat{\boldsymbol{b}} + \boldsymbol{z}$ where $\begin{cases} \widehat{\boldsymbol{b}} \in \text{Col}\, A \text{ and} \\ \boldsymbol{z} \in (\text{Col}\, A)^{\perp} \end{cases}$

using

$$\begin{cases} \widehat{\boldsymbol{b}} = \frac{\boldsymbol{b}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\boldsymbol{v_1} + ... + \frac{\boldsymbol{b}\cdot\boldsymbol{v_i}}{\boldsymbol{v_i}\cdot\boldsymbol{v_i}}\boldsymbol{v_i} + ... + \frac{\boldsymbol{b}\cdot\boldsymbol{v_p}}{\boldsymbol{v_n}\cdot\boldsymbol{v_p}}\boldsymbol{v_p} \\ \quad = \text{proj}_{\boldsymbol{v_1}}\, \boldsymbol{b} + ... + \text{proj}_{\boldsymbol{v_i}}\, \boldsymbol{b} + ,,, + \text{proj}_{\boldsymbol{v_p}}\, \boldsymbol{b} \\ \\ \boldsymbol{z} = \boldsymbol{b} - \widehat{\boldsymbol{b}} \end{cases}$$

$\widehat{\boldsymbol{b}} = \text{proj}_{\text{Col}\, A}\, \boldsymbol{b} =$ projection of $\boldsymbol{b}$ on Col A
$\boldsymbol{z} = \boldsymbol{b} - \widehat{\boldsymbol{b}} =$ component of $\boldsymbol{b}$ orthogonal to Col A

and $\boldsymbol{z}$ are unique: meaning: if "somehow" you write

$$\boldsymbol{b} = \boldsymbol{v_1} + \boldsymbol{v_2} \text{ where } \begin{cases} \boldsymbol{v_1} & \text{is in Col } A \\ \boldsymbol{v_2} & \text{is in } (\text{Col}\, A)^{\perp} \end{cases}$$

then it must be that $\boldsymbol{v_1} = \widehat{\boldsymbol{b}}$ and $\boldsymbol{v_2} = \boldsymbol{z}$
(as given by the formulas above)

Example Find the least squares solution of

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 3 \\ -1 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

$$A \quad \boldsymbol{x} = \quad \boldsymbol{b}$$

(*an inconsistent system*)

We solve $A\boldsymbol{x} = \widehat{\boldsymbol{b}}$ instead; because A happens to have orthogonal columns, it's easy to find $\widehat{\boldsymbol{b}}$.

Let $W = \operatorname{Col} A$. The columns of $A = [\boldsymbol{v_1}\ \boldsymbol{v_2}]$ are an <u>orthogonal basis</u> for W. Orthogonal Decomposition Theorem gives:

$$\widehat{\boldsymbol{b}} = \frac{\boldsymbol{b}\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1 + \frac{\boldsymbol{b}\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2$$

$$= \frac{0}{2}\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} + \left(-\frac{3}{13}\right)\begin{bmatrix} 0 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{6}{13} \\ -\frac{9}{13} \\ 0 \end{bmatrix}$$

Then just solve $A\boldsymbol{x} = \widehat{\boldsymbol{b}}$ directly:

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 3 \\ -1 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ -\frac{6}{13} \\ -\frac{9}{13} \\ 0 \end{bmatrix}$$

But the work for this is already done! in the solution

$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, x_1 and x_2 are the weights needed to

write $\begin{bmatrix} 0 \\ -\frac{6}{13} \\ -\frac{9}{13} \\ 0 \end{bmatrix}$ as a linear combination of the

columns of $A = [\boldsymbol{v_1}\ \boldsymbol{v_2}]$ – and we have already found those weights:

$$\boldsymbol{x} = \begin{bmatrix} 0 \\ -\frac{3}{13} \end{bmatrix} = \begin{bmatrix} \frac{\boldsymbol{b}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}} \\ \frac{\boldsymbol{b}\cdot\boldsymbol{v_2}}{\boldsymbol{v_2}\cdot\boldsymbol{v_2}} \end{bmatrix}$$

For $A\boldsymbol{x} = \boldsymbol{b}$, the least squares solution is

$$\boldsymbol{x} = \widehat{\boldsymbol{x}} = \begin{bmatrix} 0 \\ -\frac{3}{13} \end{bmatrix}$$

The least squares error is $||\, A\,\widehat{\boldsymbol{x}} - \boldsymbol{b}\,|| = ||\,\widehat{\boldsymbol{b}} - \boldsymbol{b}\,||$

$$= || \begin{bmatrix} 0 \\ -\frac{6}{13} \\ -\frac{9}{13} \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} ||$$

$$= \sqrt{(0-1)^2 + (-\tfrac{6}{13} - 0)^2 + (-\tfrac{9}{13} + 1)^2 + (0-1)^2}$$

$$\approx 1.5191$$

For any $\boldsymbol{x} \neq \begin{bmatrix} 0 \\ -\frac{3}{13} \end{bmatrix}$, $\|A\boldsymbol{x} - \boldsymbol{b}\| > 1.5191$

(" $>$ " here rather than " $\geq$ " because,

having solved, we know in this example that
$\boldsymbol{x} = \widehat{\boldsymbol{x}} = \begin{bmatrix} 0 \\ -\frac{3}{13} \end{bmatrix}$ is the <u>unique</u> least squares solution. (*Sometimes there is more than one!*)

For example: if we used $\boldsymbol{x} = \mathbf{0}$, instead of

$\boldsymbol{x} = \widehat{\boldsymbol{x}} = \begin{bmatrix} 0 \\ -\frac{3}{13} \end{bmatrix}$ then we get a larger

error $= ||A\boldsymbol{x} - \boldsymbol{b}|| = ||\,\mathbf{0} - \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}|| = \sqrt{3}$

$\approx 1.7321 > 1.5191$

To find leat squares solution for $A\boldsymbol{x} = \boldsymbol{b}$
(brute force method)

1) Find a basis for $W = \text{Col } A$: the set of pivot columns of A

2) Apply the Gram-Schmidt Process (*if necessary*) to
get an orthogonal basis for $W = \text{Col } A$

3) Use the Orthogonal Decomposition Theorem to find $\widehat{\boldsymbol{b}}$

4) Solve $A\boldsymbol{x} = \widehat{\boldsymbol{b}}$ (*using row reduction*; *always consistent since* $\widehat{\boldsymbol{b}}$ *is in Col* A)

Easier (justified in class): The solutions of $A\boldsymbol{x} = \widehat{\boldsymbol{b}}$ are eactly the same as the solutions of

$$A^T A\boldsymbol{x} = A^T\boldsymbol{b}$$

Example Find least squares solutions for $A\boldsymbol{x} = \boldsymbol{b}$ (*inconsistent*)

$$\begin{bmatrix} -5 & 2 & -3 & 2 & -1 \\ 1 & 1 & 2 & 1 & 0 \\ 3 & 1 & 4 & 1 & 1 \\ -1 & 1 & 0 & 1 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 0 \\ 3 \\ 6 \\ 2 \end{bmatrix}$$

Solve $A^T A\,\boldsymbol{x} = A^T \boldsymbol{b}$

$$\begin{bmatrix} 36 & -7 & 29 & -7 & 8 \\ -7 & 7 & 0 & 7 & -1 \\ 29 & 0 & 29 & 0 & 7 \\ -7 & 7 & 0 & 7 & -1 \\ 8 & -1 & 7 & -1 & 2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 19 \\ 11 \\ 30 \\ 11 \\ 6 \end{bmatrix}$$

The augmented matrix row reduces to

$$\begin{bmatrix} \mathbf{1} & 0 & 1 & 0 & 0 & \frac{19}{34} \\ 0 & \mathbf{1} & 1 & 1 & 0 & \frac{41}{17} \\ 0 & 0 & 0 & 0 & \mathbf{1} & \frac{67}{34} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

There are 2 free variables, $\boldsymbol{x_3}$ and $\boldsymbol{x_4}$. The general solution is

$$\widehat{\boldsymbol{x}} = \boldsymbol{x} = \begin{bmatrix} \frac{19}{34} - x_3 \\ \frac{41}{17} - x_3 - x_4 \\ x_3 \\ x_4 \\ \frac{67}{34} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{19}{34} \\ \frac{41}{17} \\ 0 \\ 0 \\ \frac{67}{34} \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

All of these solutions are "least squares solutions."

For any two of them, say $\widehat{\boldsymbol{x}}_1$ and $\widehat{\boldsymbol{x}}_2$:

$$A\widehat{\boldsymbol{x}}_1 = \widehat{\boldsymbol{b}} \quad \text{and } A\widehat{\boldsymbol{x}}_2 = \widehat{\boldsymbol{b}}$$

so $\| A\widehat{\boldsymbol{x}}_1 - \boldsymbol{b} \| = \| A\widehat{\boldsymbol{x}}_2 - \boldsymbol{b} \| = \| \widehat{\boldsymbol{b}} - \boldsymbol{b} \|$

$=$ least square error

<u>If there are infinitely many least squares solutions, then they all produce the same least square error</u>.

To compute least square error, pick any particular least squares solution: say, with $x_3 = 1$, $x_4 = 0$.

$$\widehat{\boldsymbol{x}} = \begin{bmatrix} \frac{19}{34} \\ \frac{41}{17} \\ 0 \\ 0 \\ \frac{67}{34} \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{15}{34} \\ \frac{24}{17} \\ 1 \\ 0 \\ \frac{67}{34} \end{bmatrix}$$

Then $A\widehat{\boldsymbol{x}} = A \begin{bmatrix} -\frac{15}{34} \\ \frac{24}{17} \\ 1 \\ 0 \\ \frac{67}{34} \end{bmatrix} = \begin{bmatrix} \frac{1}{17} \\ \frac{101}{34} \\ \frac{103}{17} \\ \frac{63}{34} \end{bmatrix} = \widehat{\boldsymbol{b}}$

so least squares error $=$

$$\| A\widehat{\boldsymbol{x}} - \boldsymbol{b} \| = \| \widehat{\boldsymbol{b}} - \boldsymbol{b} \|$$

$$= \| \begin{bmatrix} \frac{1}{17} \\ \frac{101}{34} \\ \frac{103}{17} \\ \frac{63}{34} \end{bmatrix} - \begin{bmatrix} 0 \\ 3 \\ 6 \\ 2 \end{bmatrix} \| = \frac{1}{\sqrt{34}}$$