Lecture 16

Revisiting definitions

Orthogonal complement of subspace W

- denoted $W^{\perp}$

$W^{\perp} = \{$ all $\vec{z}$ s.t. $\vec{z} \perp W\}$

or

$\{$ all $\vec{z}$ s.t. $\vec{z} \perp \vec{u}$ for any $\vec{u} \in W\}$

$\vec{z} \perp \vec{u} \longrightarrow \vec{u} \cdot \vec{z} = 0$

- $proj_W \vec{z} = 0$ for all $z \in W^{\perp}$
- $W^{\perp}$ can be thought of as all area that cannot be described by W

$\Rightarrow \dim(W) + \dim(W^{\perp}) = n$

where $W, W^{\perp} \in \mathbb{R}^n$

TH 1

$$W = \text{Span}\left\{ \begin{array}{cc} \vec{u}_1 & \vec{u}_2 \\ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{array} \right\} \qquad \vec{b} = \begin{bmatrix} 3 \\ 5 \\ 8 \end{bmatrix}$$

$$\vec{b} = \underset{\substack{\downarrow \\ W}}{\hat{b}} + \underset{\substack{\downarrow \\ W^{\perp}}}{\vec{z}}$$

$$\text{Proj}_W \vec{b} = \frac{\vec{b} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \cdot \vec{u}_1 + \frac{\vec{b} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \cdot \vec{u}_2$$

$$= 3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 5 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 0 \end{bmatrix}$$

TH2

$$W = \text{span}\left\{ \overset{\vec{u}_1}{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}}, \overset{\vec{u}_2}{\begin{bmatrix} -3 \\ -6 \\ 0 \end{bmatrix}} \right\} \quad \dim(W^\perp) = ?$$

$$\dim(W) + \dim(W^\perp) = \dim(\mathbb{R}^3)$$
$$1 + 2 = 3$$

Find $W^\perp$

Create matrix $A = \begin{bmatrix} 1 & -3 \\ 2 & -6 \\ 0 & 0 \end{bmatrix} \Rightarrow A^T\vec{x} = 0$

nullspace of A^T

Find $\vec{x}$ s.t.

$\vec{x} \cdot \vec{u}_1 = 0$, $\vec{x} \cdot \vec{u}_2 = 0$

$$A^T x = \begin{bmatrix} 1 & 2 & 0 \\ -3 & -6 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \overbrace{1 \cdot x_1 + 2 \cdot x_2 + 0 \cdot x_3}^{\vec{x} \cdot \vec{u}_1} \\ \end{bmatrix}$$

Theorem: An orthogonal set of non-zero vectors is linearly independent

Theorem: Suppose $\dim(W) = p$ and $\{\vec{u}_1, \dots, \vec{u}_p\}$ is an orthogonal set of non-zero vectors in W

$\rightarrow \{\vec{u}_1, \dots, \vec{u}_p\}$ is a basis for W

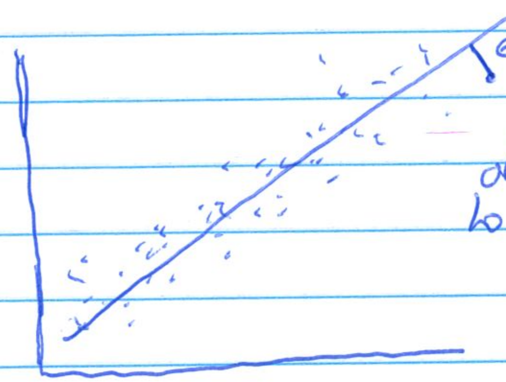
(3)

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \qquad b = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

Want to solve $A\vec{x} = \vec{b}$

$$\left[\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 5 & 2 \\ 1 & 7 & 3 \\ 1 & 8 & 3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 1/3 \\ 0 & 0 & 0 \end{array}\right]$$

- Still interested in finding a best solution



minimize total distance of all points to the line

Def A least-square solution of $A\vec{x} = \vec{b}$ is an $\hat{\vec{x}}$

$$\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{x}\|$$

for all $\vec{x} \in \mathbb{R}^n$

$A\vec{x} = \vec{b}$ is inconsistent. Find $\hat{x}$ that brings us closest to $\vec{b}$ s.t. we minimize the distance between $\vec{b}$ and $\hat{b} = A\hat{x}$

$\hat{b} \in \text{colsp}(A)$

So we want to find $\hat{x}$

$$\hat{x} = \min_{\vec{x} \in \mathbb{R}^n} \| \vec{b} - A\vec{x} \|$$

<u>Best Approximation Theorem</u>

- Let W be a subspace of $\mathbb{R}^n$
- Let $\vec{y}$ be a vector of $\mathbb{R}^n$

$\rightarrow \| \vec{y} - \text{proj}_W \vec{y} \| < \| \vec{y} - \vec{v} \|$
for any $v \neq \text{proj}_W \vec{y}$

vector in W "closest" to $\vec{y}$
is $\text{proj}_W \vec{y}$

So want to change
$$\| \vec{b} - A\hat{\vec{x}} \| \leq \| \vec{b} - A\vec{x} \| \quad \forall x \in \mathbb{R}^n$$
to fit language above

Let $W = \text{colsp}(A)$

$\rightarrow W = \{ A\vec{x} \mid \vec{x} \in \mathbb{R}^n \}$

$\rightarrow$ let $\vec{v} \in W$

$\rightarrow$ Let $\hat{b} = \text{proj}_W \vec{b}$

$\rightarrow \| \vec{b} - \hat{\vec{b}} \| \leq \| \vec{b} - \vec{v} \|$ for all $v \in W$
by best approximation theorem

We have defined our problem

$$\Rightarrow \quad \|\vec{b} - A\hat{\vec{x}}\| \leq \|\vec{b} - A\vec{x}\| \text{ for all } \vec{x} \in \mathbb{R}^n$$

where $\quad A\hat{\vec{x}} = \hat{\vec{b}} = \operatorname{proj}_W \vec{b}$

$$W = \operatorname{Colsp}(A)$$

How do we calculate $\hat{\vec{b}}$ ~~of~~ or $A\hat{\vec{x}}$

~ Use orthogonal decomposition

$$\vec{b} = \underbrace{\hat{\vec{b}}}_{\text{in } W} + \underbrace{\vec{z}}_{\in W^{\perp}}$$

$$\vec{z} = (\vec{b} - \hat{\vec{b}}) \in W^{\perp}$$

Recall $\quad (\operatorname{Colsp}(A))^{\perp} = \operatorname{Ker} A^T$

$$A^T \vec{z} = 0$$

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

A solution for $A\hat{\vec{x}} = \hat{\vec{b}}$ satisfies

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

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

$$A^T\vec{b} - A^T A\hat{\vec{x}} = 0$$

$$\Rightarrow \quad \underbrace{A^T A}_{\sim}\hat{\vec{x}} = \underbrace{A^T\vec{b}}_{\sim}$$

$$\tilde{A}\hat{\vec{x}} = \tilde{b}$$

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

We want to find an $\hat{x}$ s.t.

Assume A^TA invertible

$$(A^TA)\hat{\vec{x}} = A^T\vec{b}$$

$$\underbrace{(A^TA)^{-1}(A^TA)}_{I}\hat{\vec{x}} = (A^TA)^{-1}A^T\vec{b}$$

$$I \quad \hat{\vec{x}} = \text{''}$$

$$\Rightarrow \quad \hat{\vec{x}} = \underbrace{(A^TA)^{-1}A^T}\vec{b}$$

For invertible A^TA this is A^+ pseudo-inverse of A

$$\hat{\vec{x}} = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

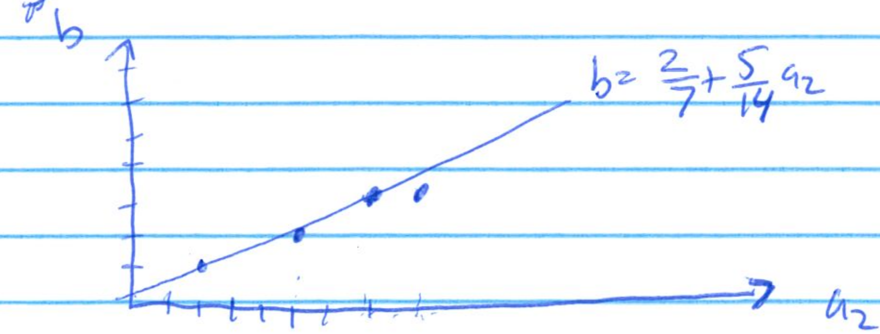
$$A\hat{x} = \begin{array}{cc} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \\ \vec{a}_1 \quad \vec{a}_2 \end{array} \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix} = \hat{b}$$

closest vector in colsp(A) to $\vec{b}$

$\equiv \operatorname{proj}_W \vec{b}$

Since a_1 is all the same let's just look at a_2 vs. b

a_2	b
2	1
5	2
7	3
8	3



In the case (A^TA) is not invertible, we still want to solve

$$(A^TA)\hat{\vec{x}} = A^T\vec{b}$$

(1) $A^TA \in \mathbb{R}^{n\times n}$

(2) $A^T\vec{b} \in \mathbb{R}^n$

(3) Put A^TA and A^Tb in an augmented matrix and solve

$$\left[\begin{array}{c|c} A^TA & A^Tb \end{array}\right]$$

Find echelon or canonical form and solve for $\hat{\vec{x}}$

Theorem: The set of least-squares solutions of $A\vec{x}=\vec{b}$ = the solution set of $(A^TA)\hat{\vec{x}} = A^T\vec{b}$

Theorem Given A is an $m\times n$ matrix the following are equivalent

a) $A\vec{x}=\vec{b}$ has a unique least square solution

b) The columns of A are linearly independent

c) A^TA is invertible