(0)

Recap

Orthogonal Complement of subspace (W):

- Denoted $W^\perp$

$W^\perp = \{$ all $\vec{z}$ such that $\vec{z} \perp W \}$

or

$W^\perp = \{$ all $\vec{z}$ such that $\vec{z} \perp \vec{u}$ for any $\vec{u} \in W \}$

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

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

dim W & dim $W^\perp$

- Given above it is not ~~s~~ suprising dim W + dim $(W^\perp) = n$ where $W, W^\perp$ are in $\mathbb{R}^n$

1

Orthogonal Set

$\{\vec{u}_1, \ldots, \vec{u}_p\}$ is an orthogonal set in $\mathbb{R}^n$ if $\vec{u}_i \perp \vec{u}_j$ for all $j \neq i$

$\Rightarrow \vec{u}_i \cdot \vec{u}_j = 0$ for all $i \neq j$

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

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

$\Rightarrow \{\vec{u}_1, \ldots, \vec{u}_p\}$ is a basis for W.

Orthogonal Decomposition Theorem

Let $\{\vec{u}_1, \ldots, \vec{u}_p\}$ be a orthogonal basis for ~~[illegible]~~ $W \in \mathbb{R}^n$

For each $\vec{y}$ in $\mathbb{R}^n$ it can be decomposed as $\vec{y} = \hat{y} + \vec{z}$

where $\vec{\hat{y}} \in W$ and $z \in W^\perp$

$\vec{\hat{y}} =$ ~~proj~~ $\text{proj}_W \vec{y}$

$$\text{proj}_W \vec{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}(\vec{u}_1) + \cdots + \frac{\vec{y} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p}(\vec{u}_p)$$

2

Consider

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

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

$$\left[\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 5 & 2 \\ 1 & 7 & 3 \\ 1 & 8 & 3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 5 & 2 \\ 0 & 6 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 1/3 \\ 0 & 0 & 0 \end{array}\right]$$

- Still intested in finding a <u>best</u> solution
- why??
- We could be studying non-determinstic (correlation)
- hours of study vs. exam grade

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

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

- $\| \ \|$ is a distance metric
- In this case, this the n-dimensional version of the 1-D least squares metric $\sqrt{(b - ax)^2}$

So we want to ~~solve~~ find $\hat{\vec{x}}$

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

- For this we need one more tool and the stuff from last lecture

Best Approximation Theorem

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

$\rightarrow \quad \|\vec{y} - \text{proj}_W \vec{y}\| < \|\vec{y} - \vec{v}\|$

for any ~~$v \in$~~ $v \neq \text{proj}_W \vec{y}$

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

$$\left\{ c_1 \begin{bmatrix} \vec{a}_1 \end{bmatrix} + \dots + c_n \begin{bmatrix} \vec{a}_n \end{bmatrix} \right\}$$

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

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

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

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

$\rightarrow \quad \|\vec{b} - \hat{\vec{b}}\| \leq \|\vec{b} - \vec{v}\|$ for all $v \in W$

by best approximation theorem

⑥

Does $\vec{\hat{b}}$ exists as a ~~sol~~ solution to $A\vec{x}$?

$\vec{\hat{b}} = \text{proj}_W \vec{b}$ which by def is in W

$\Rightarrow$ $\vec{\hat{b}}$ is a linear combination of ~~[illegible]~~ the columns of A

$\Rightarrow$ $A\vec{\hat{x}} = \vec{\hat{b}}$ is consistent

$\Rightarrow$ follow ~~[illegible]~~ close to the same reasoning for $\vec{v}$ ~~[illegible]~~ for $A\vec{x} = \vec{v}$

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

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

How do we calculate $\vec{\hat{b}}$ for ~~[illegible]~~ $A\vec{\hat{x}}$

-Use orthogonal decomposition

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

$$\vec{z} = (\vec{b} - \vec{\hat{b}})$$

We know $\vec{z} \in W^\perp$ $\quad \vec{z} \in (\text{col}(A))^\perp$

Recall $(\text{Col } A)^\perp = \text{Null } A^T$

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

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

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

$\uparrow$ is the normal equation for $A\vec{x} = \vec{b}$

Conversly $A^TA\hat{\vec{x}} = A^Tb$ satisfies

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

Since $\text{Null } A^T = (\text{Col}(A))^\perp$

$$\rightarrow (\vec{b} - A\hat{\vec{x}}) \in (\text{Col } A)^\perp = W^\perp$$

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

Back to our problem

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

$$\cancel{(A^TA)\hat{\vec{x}} = A^T\hat{b}}$$
$$\cancel{(A^TA)^{-1}(A^TA)\hat{\vec{x}} = A^T\vec{b}\,(A^TA)^{-1}}$$

8

We want to find an $\vec{\hat{x}}$ such that

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

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

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

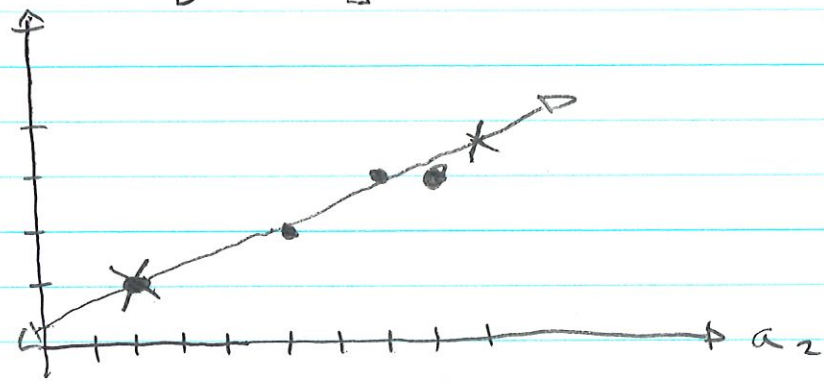
$\rightarrow$ $\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$ get as close to $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$ as possible

9

Since a_1 is all the ~~same~~ same lets just look at a_2 vs b

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \end{bmatrix}$$

a_2	b
2	1
5	2
7	3
8	3



$$b = \frac{2}{7} + \frac{5}{14} a_2$$

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

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

① Find A^TA which will be a $n \times n$ matrix

② Find $A^T\vec{b}$ which will be a $1 \times n$ vector

(10)

11

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

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

And $\tilde{\vec{x}}$ can be any solution that satisfies the above.

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

<u>Theorem</u> 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