EECS 16A

Module 3, Lecture 4.

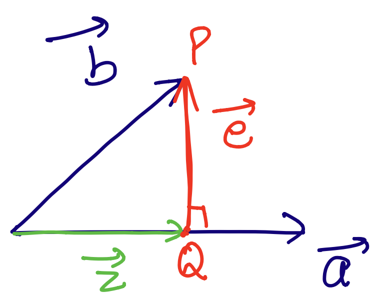
Today: • Least Squares / Projection

Last time : ① 1D Projection.

: ② Shortest distance from a point to a line is given by the perpendicular from the point to the line.

↳ Generalizes:

Shortest distance from a point to a hyperplane / vector space is given by the "orthogonal" projection of the point onto the hyperplane.

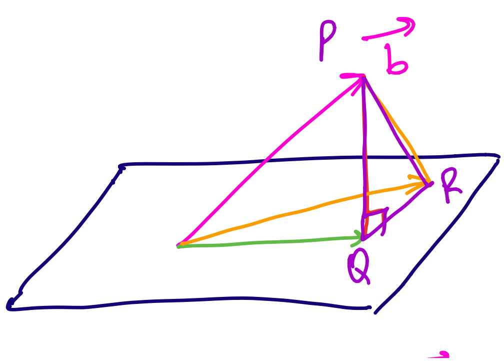


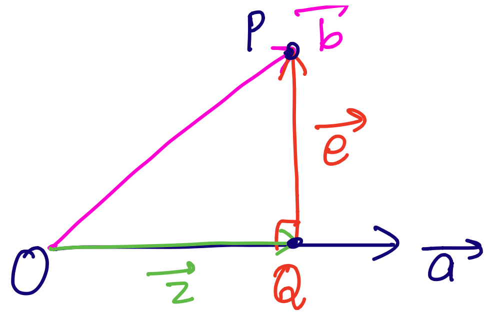
$\vec{e}$ = error

$\vec{z} + \vec{e} = \vec{b}$

$\vec{e} = \vec{b} - \vec{z}$

$\vec{z} = \alpha \cdot \vec{a}$





Want to find $\vec{z}$ in terms of $\vec{a}, \vec{b}$.

$$\boxed{\vec{z} = \alpha \cdot \vec{a}}$$

$$\boxed{\vec{z} + \vec{e} = \vec{b}}$$

$$\vec{z} = \vec{b} - \vec{e}$$
$$\vec{e} = \vec{b} - \vec{z}$$

$\langle \vec{e}, \vec{a} \rangle = 0$: Must be true since we want $\vec{e}$ orthogonal to $\vec{a}$.

$$\Rightarrow \langle \vec{b} - \vec{z}, \vec{a} \rangle = 0$$
$$\Rightarrow \langle \vec{a}, \vec{b} - \vec{z} \rangle = 0$$

inner products don't care about order.

$$\Rightarrow \vec{a}^T (\vec{b} - \vec{z}) = 0$$
$$\Rightarrow \vec{a}^T (\vec{b} - \alpha \cdot \vec{a}) = 0$$
$$\Rightarrow \vec{a}^T \vec{b} - \alpha \cdot \vec{a}^T \vec{a} = 0$$
$$\Rightarrow \langle \vec{a}, \vec{b} \rangle - \alpha \|\vec{a}\|^2 = 0$$

$$\boxed{\alpha = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\|^2}}$$

SCALAR

$$\boxed{\vec{z} = \alpha \cdot \vec{a}} = \vec{z} = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\|^2} \cdot \vec{a}$$

$$\|\vec{z}\| = \|\alpha \vec{a}\| = |\alpha| \cdot \|\vec{a}\|$$

$$= \frac{|\langle \vec{a}, \vec{b} \rangle|}{\|\vec{a}\|^{\cancel{2}}} \cdot \cancel{\|\vec{a}\|}$$

$$= \frac{|\langle \vec{a}, \vec{b} \rangle|}{\|\vec{a}\|}$$

1D Projection or Least Squares.

Systems: $A\vec{x} \approx \vec{b}$

In GPS: have multiple satellites.

A is no longer just a vector, but is an actual matrix.

$$\vec{a} x = \vec{b}$$

Moving beyond 1D.

Thm: Consider matrix A, $\vec{y} \in \text{Colspace}(A)$.

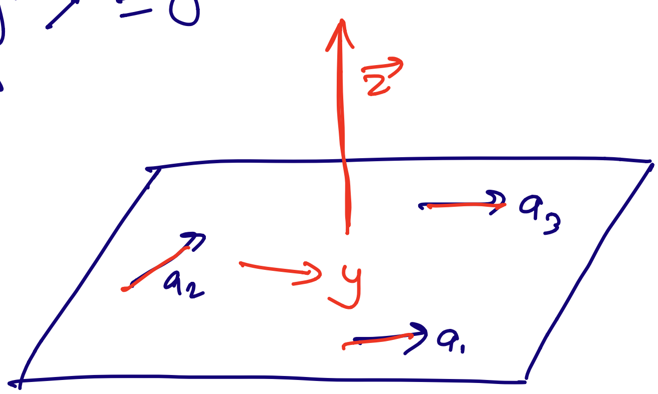
$$A = \begin{bmatrix} \vec{a}_1, & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$$

Then, consider $\vec{z}$

$\langle \vec{z}, \vec{a}_1 \rangle = 0$ $\quad \cdots \quad$ $\langle \vec{z}, \vec{a}_n \rangle = 0$

$\langle z, \vec{a}_2 \rangle = 0$

Then: $\langle \vec{z}, \vec{y} \rangle = 0$



Proof:

$\boxed{\text{eg. } A \cdot \vec{b}}$

Known:

$\vec{y} \in \text{Col}(A)$

$A \cdot (2\vec{b})$

$\vec{y}$ is a lin. comb of $\vec{a}_1, \vec{a}_2 \cdots \vec{a}_n$ $\quad = 2 \cdot A \cdot \vec{b}$

$$\vec{y} = c_1 \cdot \vec{a_1} + c_2 \cdot \vec{a_2} + \cdots + c_n \cdot \vec{a_n}$$

$$\langle \vec{z}, \vec{a_1} \rangle = 0 \; \ldots \; \langle \vec{z}, \vec{a_n} \rangle = 0$$

<u>Want:</u>

$$\langle \vec{z}, \vec{y} \rangle = 0$$

$$\langle \vec{z}, \vec{y} \rangle$$

Inner products are linear operation.

$$= \langle \vec{z}, c_1\vec{a_1} + c_2\vec{a_2} + \cdots + c_n\vec{a_n} \rangle$$

$$= \langle \vec{z}, c_1\vec{a_1} \rangle + \langle z, c_2\vec{a_2} \rangle + \cdots + \langle \vec{z}, c_n\vec{a_n} \rangle$$

$$= c_1 \langle \vec{z}, \vec{a_1} \rangle + c_2 \langle \vec{z}, \vec{a_2} \rangle + \cdots + c_n \langle \vec{z}, \vec{a_n} \rangle$$

$$= c_1 \cdot 0 + c_2 \cdot 0 + \cdots + c_n \cdot 0$$

$$= 0$$

QED.

<u>Least Squares- Algorithm</u>

<u>Trilateration</u> $\Rightarrow$ $A\vec{x} = \vec{b} + \vec{e}$

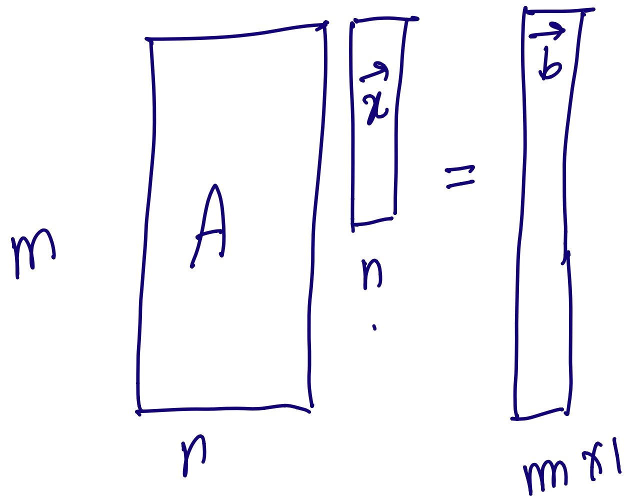
<u>Noisy equations</u>

Many equations. More equations than unknowns

$\boxed{A : m \times n}$ $\quad m > n$

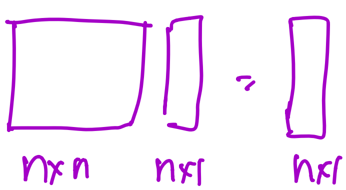
m equations

n unknowns



Overdetermined system.

Square:



Least-squares: Find $\vec{x}$, such that

$A\vec{x}$ is close to $\vec{b}$

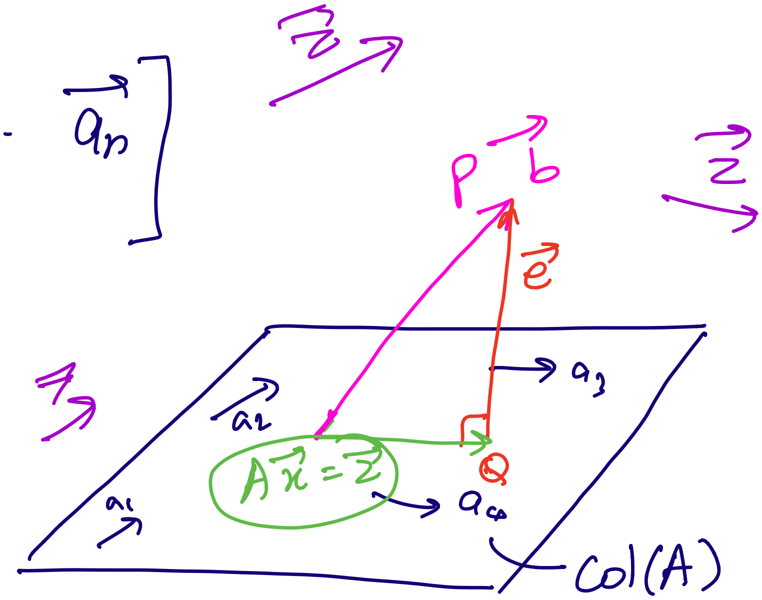
$$\vec{b} - A\vec{x} = \vec{e}$$

$$A\vec{x} = \vec{z}$$

$$\|A\vec{x} - \vec{b}\| = \|\vec{e}\| \text{ to be minimized.}$$

$$A = \begin{bmatrix} \vec{a}_1, \vec{a}_2 \cdots \vec{a}_n \end{bmatrix}$$

What is $A\vec{x} = \vec{z}$



$$\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} \begin{bmatrix} \vec{x} \end{bmatrix} = Col(A)$$

We are searching over $\vec{z} \in Col(A)$ because $A\vec{x}$ represents all vectors in the col. (A).

$$\vec{z} + \vec{e} = \vec{b}$$

$$\vec{e} = \vec{b} - \vec{z}$$

$$\langle \vec{a}_1, \vec{e} \rangle = 0 \Rightarrow \langle \vec{a}_1, \vec{b} - \vec{z} \rangle = 0$$

$$\langle \vec{a}_2, \vec{e} \rangle = 0$$

$$\vdots \qquad\qquad \vdots$$

$$\langle \vec{a}_n, \vec{e} \rangle = 0 \Rightarrow \langle \vec{a}_n, \vec{b} - \vec{z} \rangle = 0$$

$$\boxed{\begin{aligned} \vec{a}_1^T(\vec{b}-\vec{z}) &= 0 \\ \vec{a}_2^T(\vec{b}-\vec{z}) &= 0 \\ &\vdots \\ \vec{a}_n^T(\vec{b}-\vec{z}) &= 0 \end{aligned}}$$

$$A = [\vec{a}_1 \ldots \vec{a}_n]$$

$$A^T = \begin{bmatrix} \text{---}\ \vec{a}_1^T\ \text{---} \\ \text{---}\ \vec{a}_2^T\ \text{---} \\ \vdots \\ \text{---}\ \vec{a}_n^T\ \text{---} \end{bmatrix}$$

$$A = m \times n$$

$$A^T : n \times m$$

$$\begin{bmatrix} \text{---}\ \vec{a}_1^T\ \text{---} \\ \text{---}\ \vec{a}_2^T\ \text{---} \\ \vdots \\ \text{---}\ \vec{a}_n^T\ \text{---} \end{bmatrix} \begin{bmatrix} \vec{b} - \vec{z} \\ \\ \\ \end{bmatrix} = 0$$

$$\underbrace{A^T}_{n\times m}\,(\overbrace{\vec{b}}^{m\times 1}-\overbrace{\underbrace{\vec{z}}}^{m\times 1})=0$$

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

$$A^T\vec{b}-A^TA\cdot\underbrace{\vec{x}}=0$$

$$\boxed{A^TA\vec{x}=A^T\vec{b}} \qquad A\vec{x}=\vec{b}$$

$$\underset{n\times m}{A^T}\cdot\underset{m\times n}{A}\ :\ n\times n$$

Use inversion! if A^TA is invertible:

$$\boxed{\hat{\vec{x}}=(A^TA)^{-1}A^T\cdot\vec{b}\quad !!!}$$

"General least squares algorithm":

$$\boxed{\boxed{\vec{z}=A\vec{x}}=A(A^TA)^{-1}A^T\cdot\vec{b}}$$

Example:

$$A = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

2×1

$m = 2$

$n = 1$

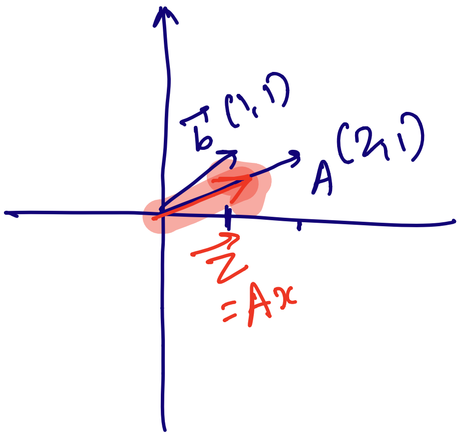
" $A\vec{x} = \vec{b}$ "

$$\boxed{\begin{bmatrix} 2 \\ 1 \end{bmatrix} [x] = \begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

Gaussian Elimination

$$\left[\begin{array}{c|c} 2 & 1 \\ 1 & 1 \end{array}\right] \to \left[\begin{array}{c|c} 1 & 1/2 \\ 1 & 1 \end{array}\right] \to \left[\begin{array}{c|c} 1 & 1/2 \\ 0 & 1/2 \end{array}\right]$$

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



$$A^T = [2 \quad 1]$$

$$A^T A = [2 \quad 1]\begin{bmatrix}2\\1\end{bmatrix}$$

$$= 5$$

$$(A^T A)^{-1} = \frac{1}{5}$$

$$A^T \cdot \vec{b} = [2 \quad 1]\begin{bmatrix}1\\1\end{bmatrix} = 3$$

$$\hat{x} = \frac{1}{5} \cdot 3 = 3/5$$

$$\vec{z} = A\hat{x} = \begin{bmatrix}2\\1\end{bmatrix} \cdot \frac{3}{5}$$

$$\begin{bmatrix}1 & 2\\2 & 3\\1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}$$
