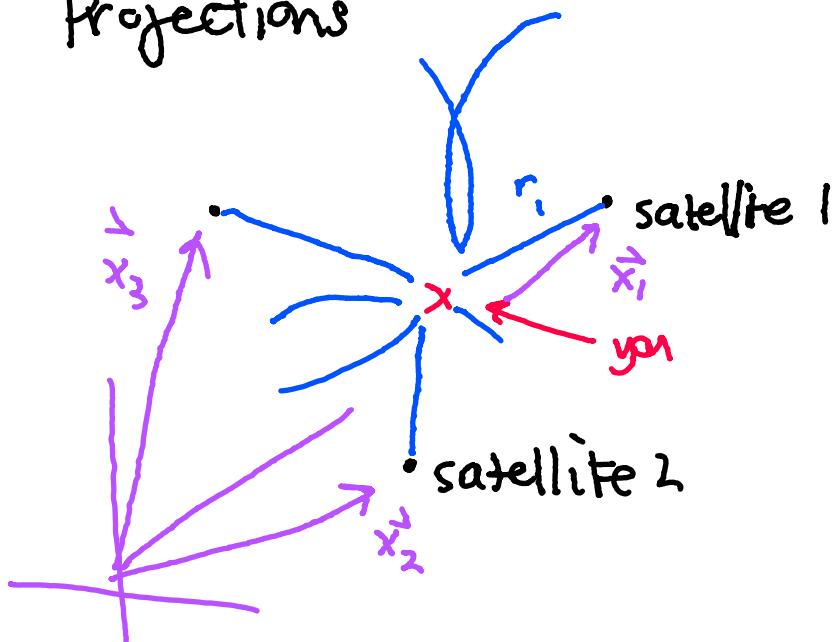


EECS16A Su24 Lecture 25

8/6/24

"Projections"



$$A \vec{x} \neq \vec{b}$$

$$A \vec{x} + \vec{n} = \vec{b} \quad \xrightarrow{\text{noise}}$$

$$A = \begin{bmatrix} \vec{a} \\ \vdots \\ \vec{a} \end{bmatrix}_{n \times 1}$$

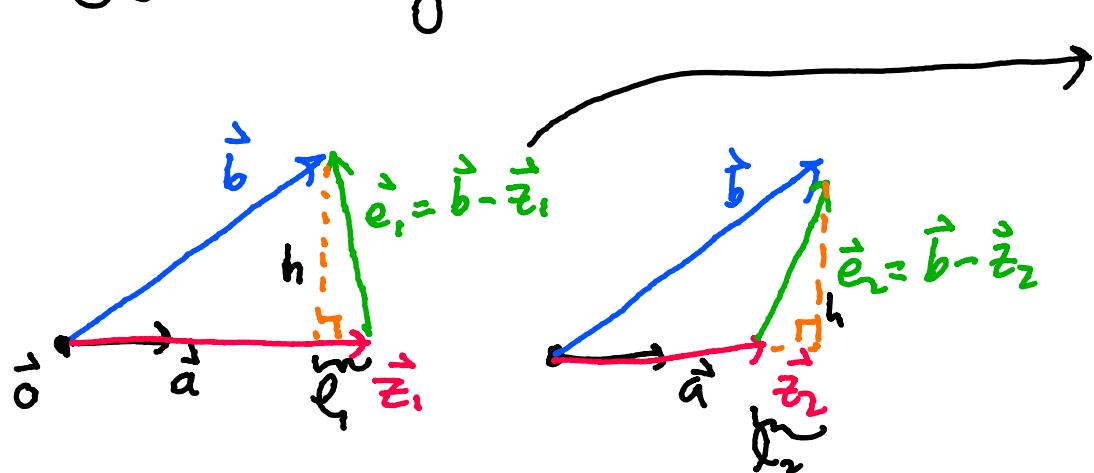
$$\begin{bmatrix} \vec{a} \\ \vdots \\ \vec{a} \end{bmatrix} \times \vec{x} \approx \begin{bmatrix} \vec{b} \\ \vdots \\ \vec{b} \end{bmatrix}$$

A is a tall matrix
 $m \times n$
 $m > n$

$$A \rightarrow \boxed{}$$

Find a
 approximate
 x that
 best satisfies
 $\vec{a} x = \vec{b}$

Geometrically

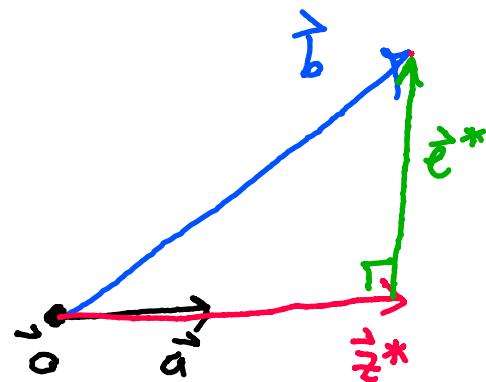


$$\|\vec{e}_1\|^2 = h^2 + l_1^2$$

$$\|\vec{e}_2\|^2 = h^2 + l_2^2$$

Best we can do:
 make $l = 0$
 $\|\vec{e}\|^2 = h^2$

Solution for best \vec{z} to get closest to \vec{b} (\vec{z} is along \vec{a})
 is the \vec{z} that makes $\vec{e} \perp \vec{z}$



How to find \vec{z}^* in terms of \vec{a}, \vec{b} ?
 (if we know $\vec{e}^* \perp \vec{z}^*$)

$$\langle \vec{e}^*, \vec{z}^* \rangle = 0$$

$$\vec{z}^* = \vec{a}x^* - \dots$$

$$\vec{e}^* = \vec{b} - \vec{z}^* \quad ; \quad \langle \vec{b} - \vec{z}^*, \vec{z}^* \rangle = 0$$

$$\dots \rightarrow \langle \vec{b} - \vec{a}x^*, \vec{a}x^* \rangle = 0$$

$$(\vec{b} - \vec{a}x^*)^T (\vec{a}x^*) = 0 \quad \square$$

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

$$\vec{b}^T \vec{a}^T x^* - \vec{a}^T \vec{a} x^{*2} = 0$$

$$\vec{b}^T \vec{a}^T x^* = \vec{a}^T \vec{a} x^{*2}$$

$$x^* = \frac{\vec{b}^T \vec{a}}{\vec{a}^T \vec{a}}$$

Q: Why \vec{a} not \vec{m}_n ?
 A: Solve simpler problem
 - first

$$\langle \vec{v}, \vec{w} \rangle$$

$$\frac{1}{2} \vec{z}^* = \vec{a} x^* = \vec{a} \cdot \frac{\vec{b}^T \vec{a}}{\vec{a}^T \vec{a}}$$

\vec{b} , \vec{z}^* are in the same vector space

\vec{z}^* is called the vector projection of \vec{b} onto G

$$\vec{z}^* = \text{proj}_{\vec{a}} \vec{b}$$

How about $A \vec{x} \approx \vec{b}$?

We're looking for \vec{x} such that $A \vec{x}$ is of the smallest length.

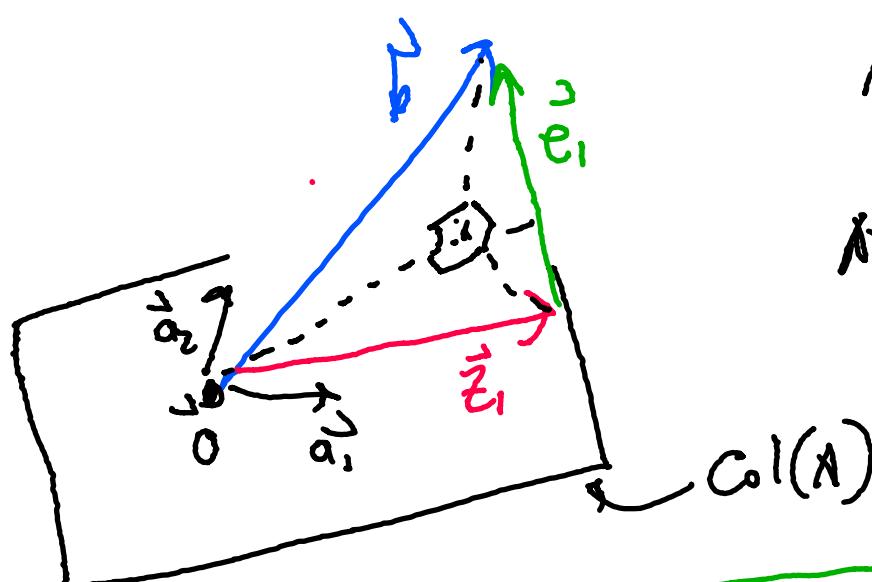
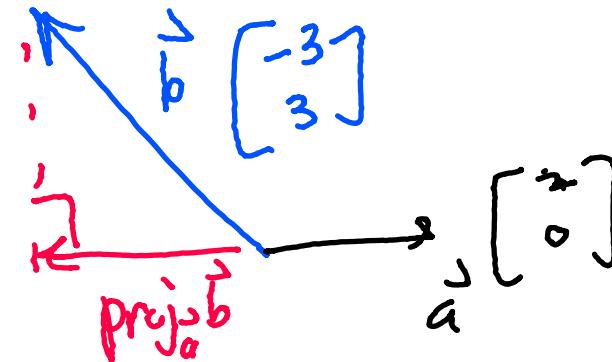
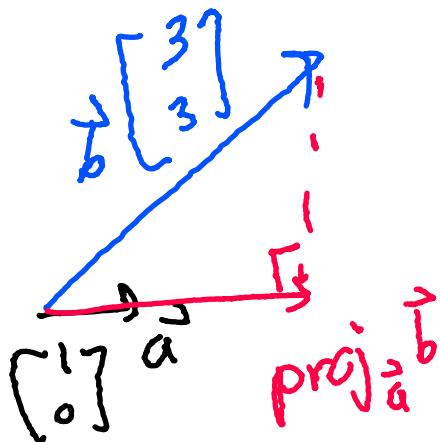
\vec{b} , x^* are not in the same space.

think of x^* as a coordinate (it multiplies the \vec{a})

x^* is called the scalar projection of \vec{b} onto \vec{a}

$$\min_{\vec{x}} \|A \vec{x} - \vec{b}\| = \|\vec{e}^*\|$$

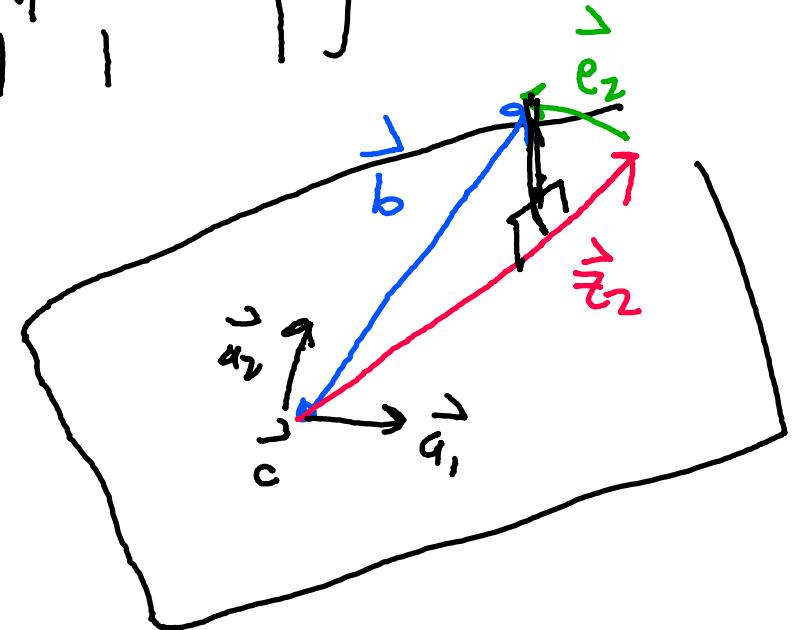
$$\vec{e} = A \vec{x} - \vec{b} \in \mathbb{R}^{2^n}$$



Want that $\vec{e} \perp \vec{z}$
 $\vec{z} \in \text{Col}(A)$

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

$$\vec{x} = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \dots & \vec{q}_n \end{bmatrix}$$



$$\vec{a} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$\text{proj}_{\vec{a}} \vec{b} = \frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}} \vec{a}$$

$\xrightarrow{x^*}$

$$= \frac{\begin{bmatrix} 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 3 \\ 3 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \frac{3}{1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

$$\vec{a} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$$

$$\text{proj}_{\vec{a}} \vec{b} = \frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}} \vec{a}$$

$$= \frac{\begin{bmatrix} 2 \\ 0 \end{bmatrix}^T \begin{bmatrix} -3 \\ 3 \end{bmatrix}}{\begin{bmatrix} 2 \\ 0 \end{bmatrix}^T \begin{bmatrix} 2 \\ 0 \end{bmatrix}} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$= \frac{-6}{4} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -3 \\ 0 \end{bmatrix}$$

Want $\langle \vec{z}, \vec{e} \rangle = 0$ to ensure $\vec{z} \perp \vec{e}$, $\vec{z} \in \text{col}(A)$

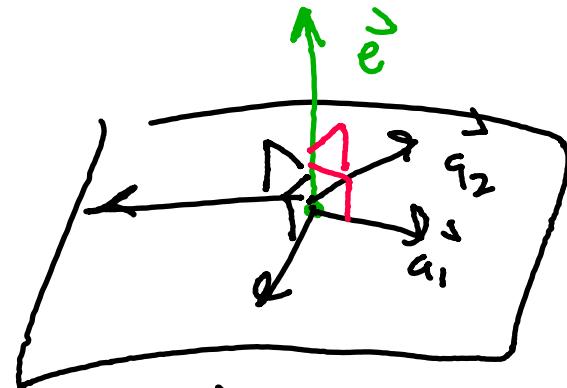
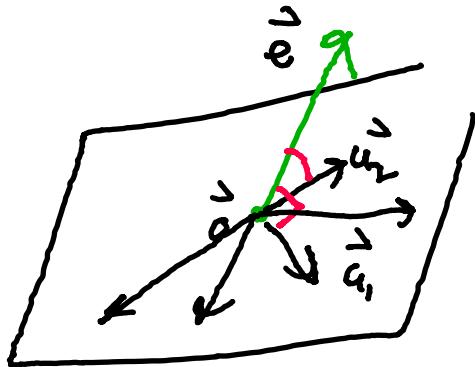
$$\vec{z} = A\vec{x}$$

$$= \vec{a}_1 x_1 + \vec{a}_2 x_2 + \vec{a}_3 x_3 + \cdots + \vec{a}_n x_n$$

$A \in \mathbb{R}^{m \times n}$

m rows
n cols

A



Claim: if $\vec{e} \perp \vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \rightarrow$
then $\vec{e} \perp \vec{z}, \vec{z} \notin \text{Span}\{\vec{a}_1, \dots, \vec{a}_n\}$

If $\vec{e} \perp \vec{a}_1, \dots, \vec{e} \perp \vec{a}_n \Rightarrow \langle \vec{e}, \vec{a}_i \rangle = 0$ for $i=1, \dots, n$
 $\vec{e} \perp \vec{a}_i = 0$ for $i=1, \dots, n$

Want as a conclusion that $\langle \vec{e}, \vec{z} \rangle = 0$

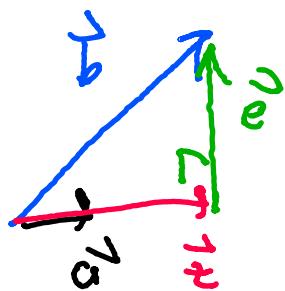
$$\vec{z} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \cdots + \alpha_n \vec{a}_n$$

$$\langle \vec{e}, \vec{z} \rangle = \langle \vec{e}, \alpha_1 \vec{a}_1 + \cdots + \alpha_n \vec{a}_n \rangle ? = 0$$

$$= \alpha_1 \langle \vec{e}, \vec{a}_1 \rangle + \alpha_2 \langle \vec{e}, \vec{a}_2 \rangle + \cdots + \alpha_n \langle \vec{e}, \vec{a}_n \rangle ? = 0$$

Yes, $\vec{z} \perp \vec{e}$.

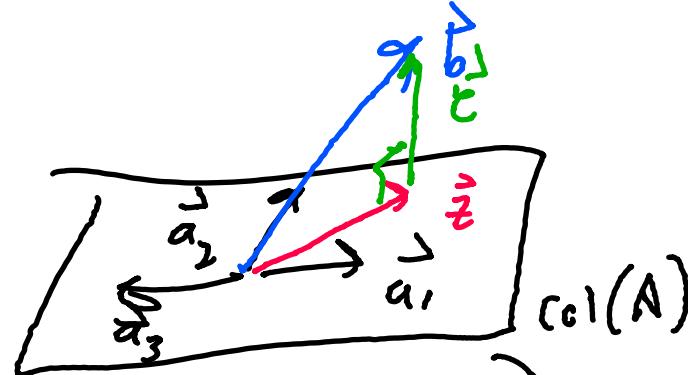
Before ($\vec{a} x = \vec{b}$)



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

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

Now ($A \vec{x} = \vec{b}$)



$$\langle \vec{a}_1, \vec{e} \rangle = 0$$

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

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

Goal: Express $\vec{z}, \vec{e}, \vec{x}$ in terms of \vec{a}, \vec{b}

$$\vec{a}_1^T \vec{e} = 0$$

$$\vec{a}_2^T \vec{e} = 0$$

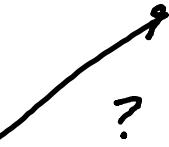
⋮

$$\vec{a}_n^T \vec{e} = 0$$



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

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \end{bmatrix}$$



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

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

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

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

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

a system of equations

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

A diagram illustrating the equation $A\vec{x} = \vec{b}$. It shows a tall, narrow blue rectangle labeled 'A' above a shorter blue rectangle labeled 'b'. A green bracket on the left indicates the height of A, and a blue bracket on the right indicates the length of b.

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

A diagram illustrating the equation $A^T A \vec{x} = A^T \vec{b}$. It shows a square blue rectangle labeled 'A^T A' above a shorter blue rectangle labeled 'A^T b'. A green bracket on the left indicates the width of A^T A, and a blue bracket on the right indicates the length of A^T b. Below the main equation, there is a smaller diagram showing a green rectangle labeled 'x' above a blue rectangle labeled 'b'.

If $A^T A$ is invertible (it's a square matrix)

The solution \vec{x} is : $(A^T A)^{-1} A^T \vec{b}$

not $\vec{0}$

coordinates of how
much to go along
each of $\vec{a}_1, \vec{a}_2, \dots$ gain
to get to the best
approximation of \vec{b}

Q: How do we know $A^T A$ is
square when A is not?

Is there ever a case $A^T A$
is not square?

(A is tall)
 $m > n$

A diagram showing a tall blue rectangle labeled 'A' above a shorter blue rectangle labeled 'AT'. A green bracket on the left indicates the height of A, and a blue bracket on the right indicates the length of AT. Below the rectangles, their dimensions are given as $m \times n$ and $n \times m$ respectively.

$$A^T A \Rightarrow \begin{matrix} \text{square} \\ n \times n \end{matrix}$$

A: $A^T A$ is always
square.

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$b = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix}$$

$$A^T A =$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 \\ 6 & 5 \end{bmatrix}$$

$$A^T b =$$

$$\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = 6$$

$$A^T A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

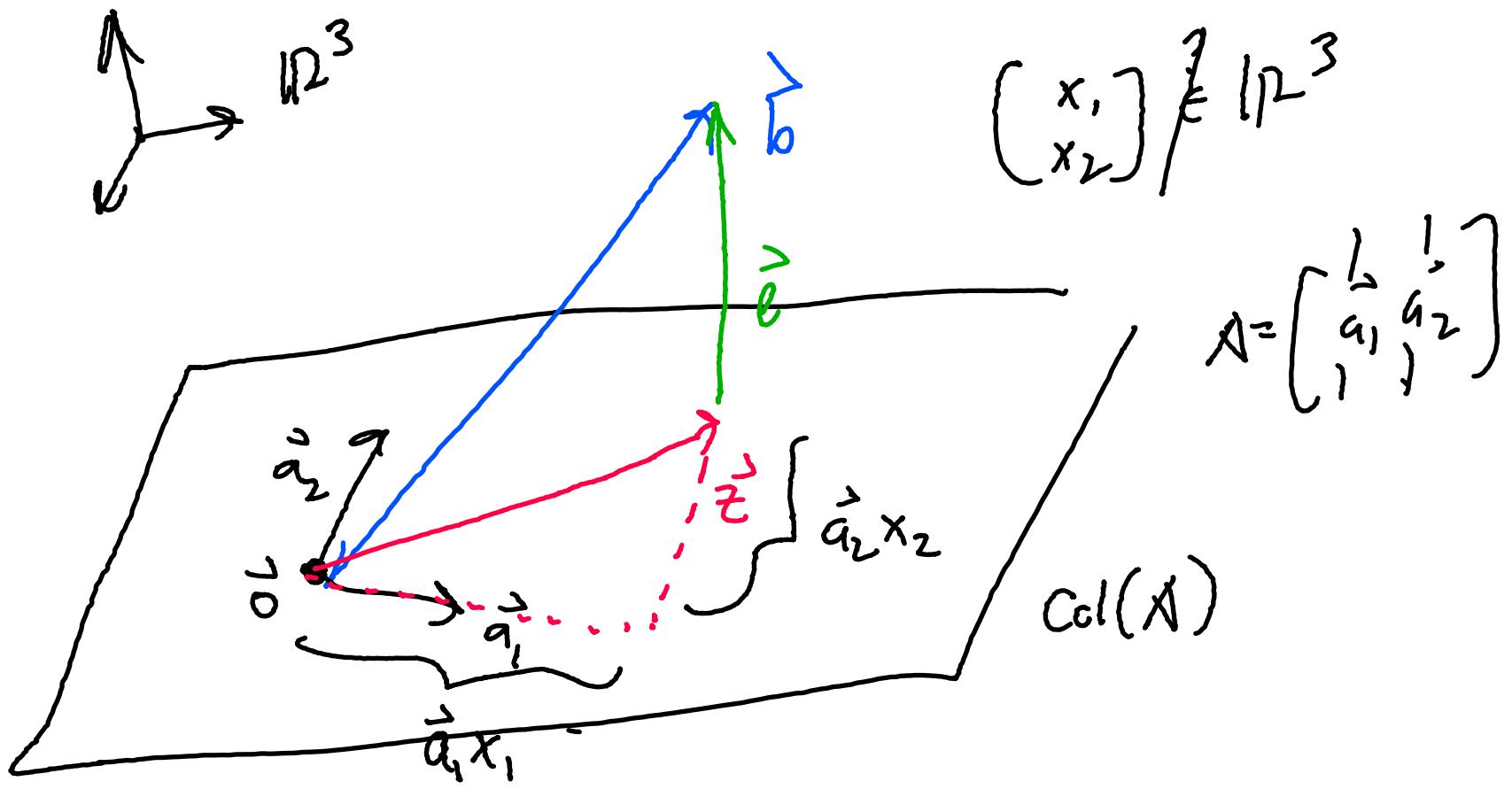
$$\vec{z} = A(A^T A)^{-1} A^T \vec{b}$$

vector projection

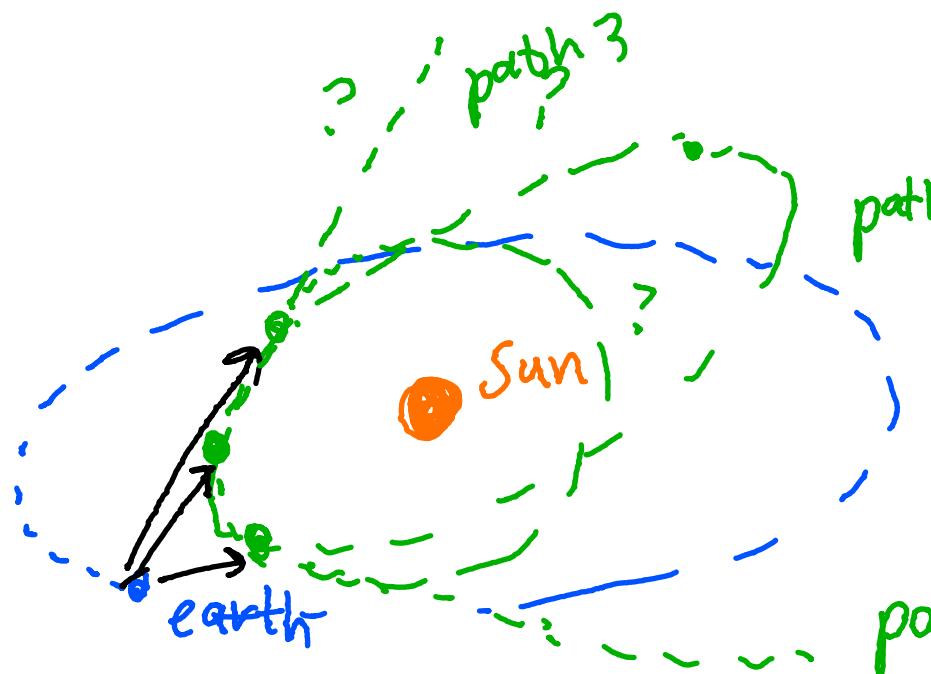
$$\vec{z} = \text{proj}_{\text{col}(A)} \vec{b}$$

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

"coordinates of projection"



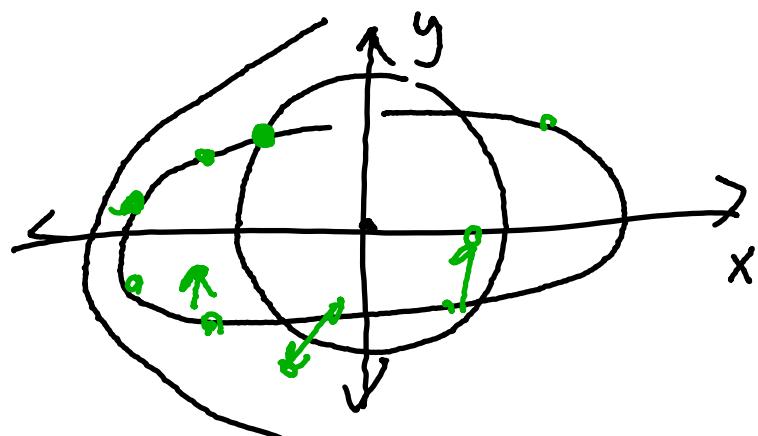
$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix}$$



Gauss, Piazzi, other folks
are interested in finding
the path of celestial object
ceres.

path 1
(possible)

$$ax^2 + bxy^2 + cxy + dx + ey = 1$$



Data: (x_1, y_1)
 (x_2, y_2)
 \vdots
 (x_3, y_3)
 \vdots
 (x_n, y_n)

$$ax_1^2 + by_1^2 + cx_1y_1 + dx_1 + ey_1 = 1$$

⋮
⋮
⋮

$$\begin{bmatrix} x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & y_n^2 & x_ny_n & x_n & y_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^n$$