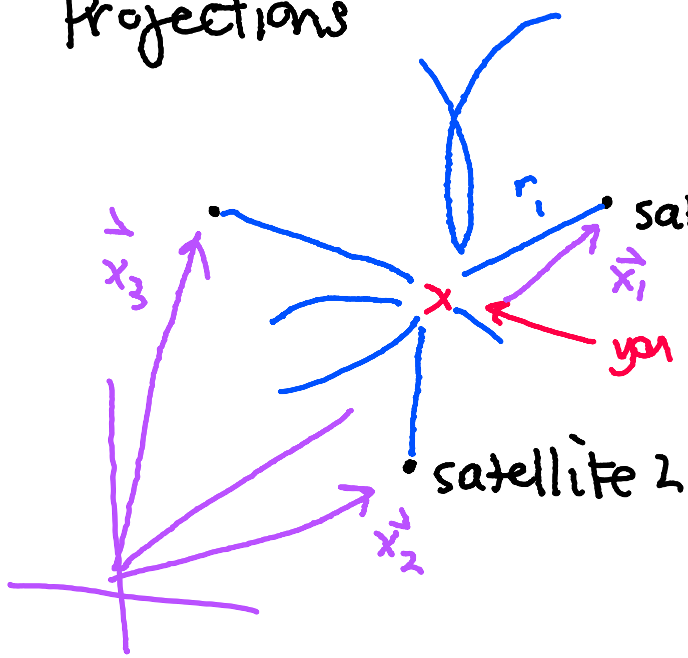
EECS16A Su24 Lecture 25

8/6/24

"Projections"



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

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

↑ noise

A is a tall matrix

$m \times n$

$m > n$

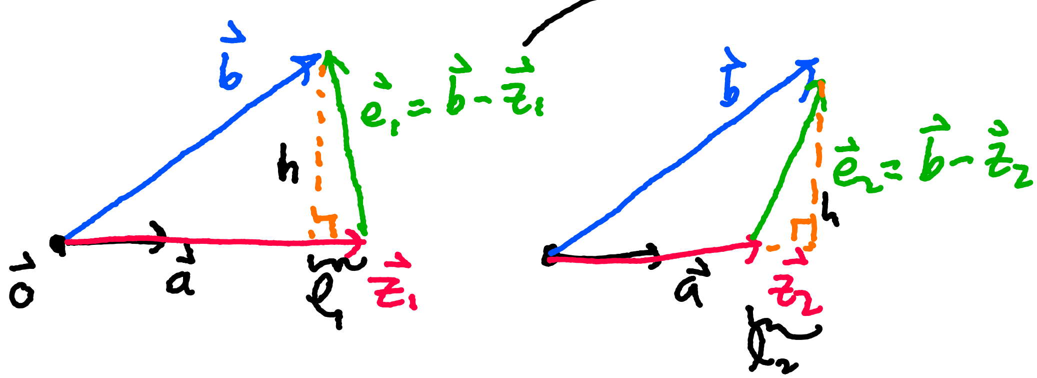
$A \rightarrow$ [ ]

$A = \underset{n\times 1}{\vec{a}}$

$\rightarrow$ $[\vec{a}]\, x \approx [\vec{b}]$

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

Geometrically



$||\vec{e}_1||^2 = h^2 + \ell_1^2$

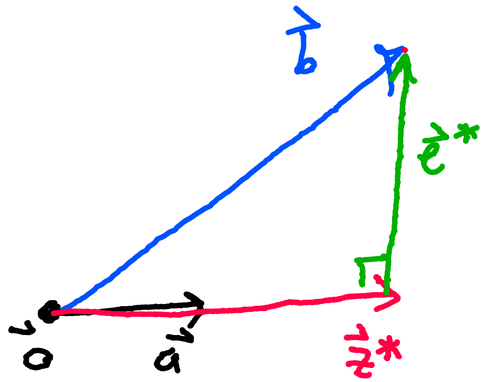
$||\vec{e}_2||^2 = h^2 + \ell_2^2$

Best we can do:

make $\ell = 0$

$||\vec{e}||^2 = h^2$

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 $\vec{e}^* \perp \vec{z}^*$)

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

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

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

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

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

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

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

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

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

Q: Why $\vec{a}$ not $A_{m \times n}$?

A: Solve simpler problem first

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

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

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

$\vec{z}^*$ is called the <u>vector projection</u> of $\vec{b}$ onto $\vec{a}$

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

$\vec{b}$ & x^* are <u>not</u> in the same space.

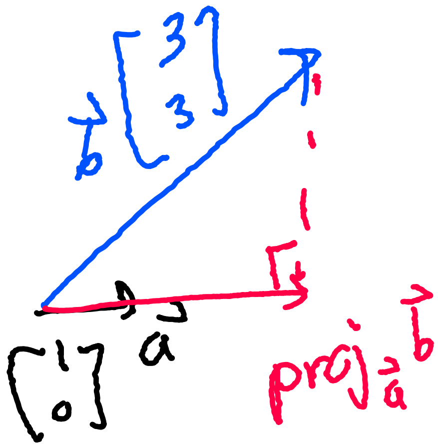
think of x^* as a coordinate (it multiplies the $\vec{a}$)

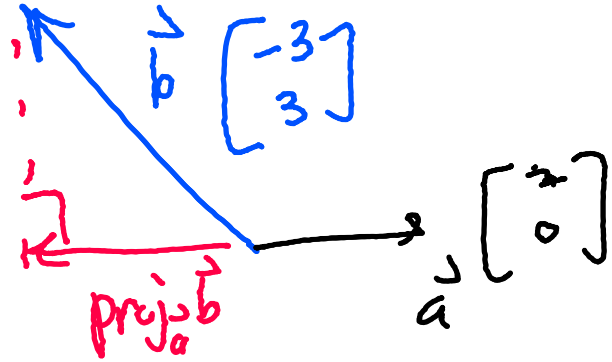
x^* is called the <u>scalar projection</u> of $\vec{b}$ onto $\vec{a}$

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

we're looking for $\vec{x}$ such that $\vec{e} = A\vec{x} - \vec{b} \in \mathbb{R}^n$ is of the smallest length.

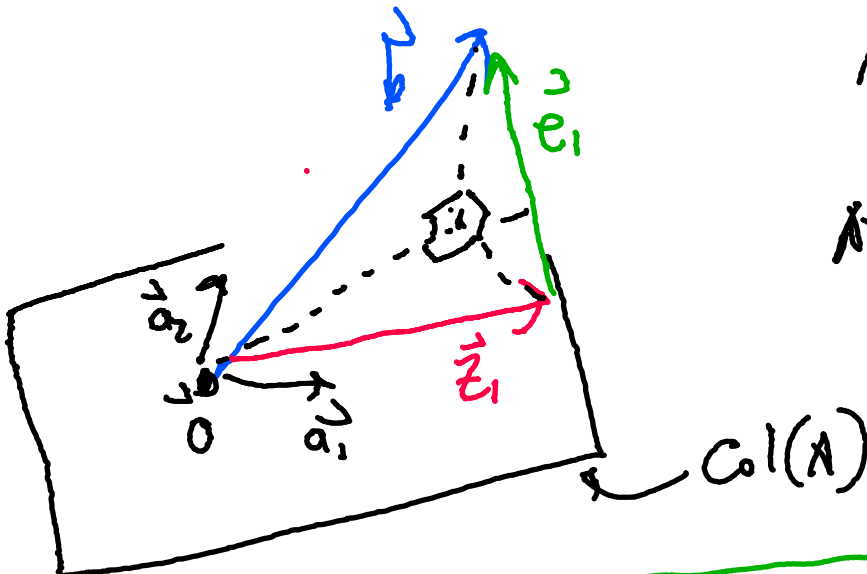
$$\min_{\vec{x}} \|A\vec{x} - \vec{b}\| = \|\vec{e}^*\|$$

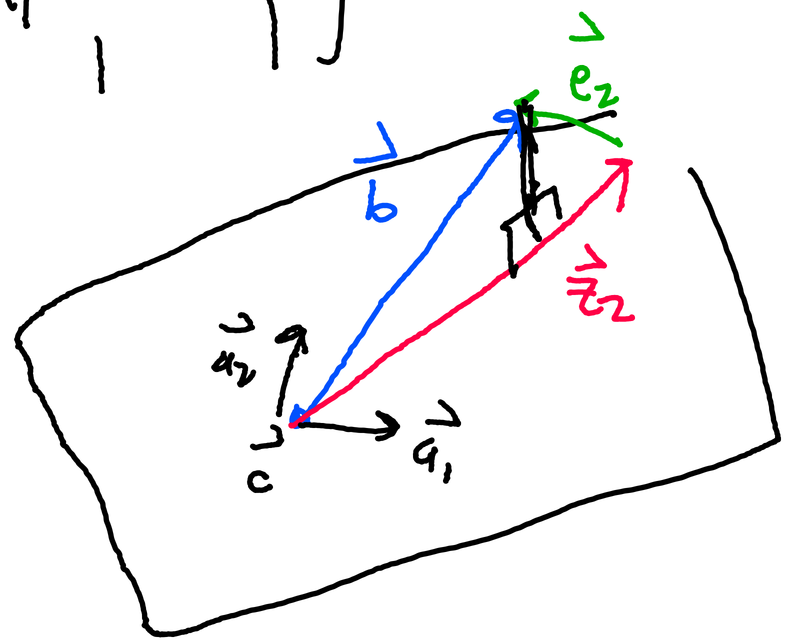




$$A\vec{x} \approx \vec{b}$$

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





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

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

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

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

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

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

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

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

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

x^*

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

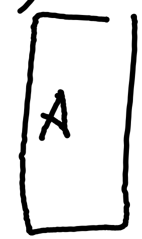
$$= \begin{bmatrix} -3 \\ 0 \end{bmatrix}$$

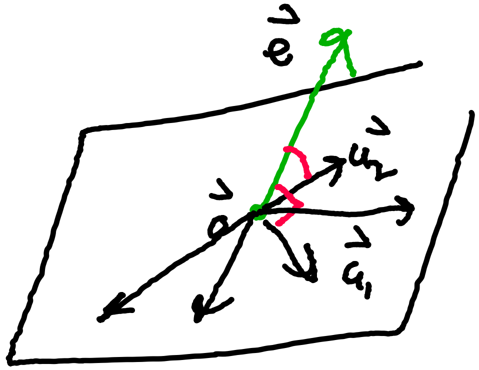
Want $\langle \vec{z}, \vec{e} \rangle = 0$ to ensure $\vec{z} \perp \vec{e}$, $\vec{z} \in \mathrm{Col}(A)$

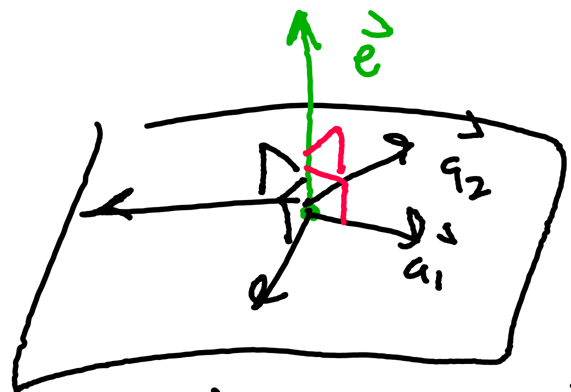
$$\vec{z} = A\vec{x}$$
$$= \vec{a}_1 x_1 + \vec{a}_2 x_2 + \vec{a}_3 x_3 + \cdots + \vec{a}_n x_n \quad (A \in \mathbb{R}^{m \times n})$$

m rows

n cols







Claim: if $\vec{e} \perp \vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n \rightarrow \left(\begin{matrix} \vec{e} \perp \vec{a}_1 \\ \vec{e} \perp \vec{a}_2 \\ \vdots \\ \vec{e} \perp \vec{a}_n \end{matrix} \right)$

then $\vec{e} \perp \vec{z}$, $\vec{z} \in \mathrm{span}\{\vec{a}_1, \ldots, \vec{a}_n\}$

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

$\vec{e}^{\,T} \vec{a}_i = 0$ for $i = 1, \ldots, n$

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

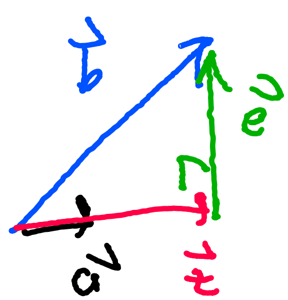
$$\vec{z} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \cdots + \alpha_n \vec{a}_n$$

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

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

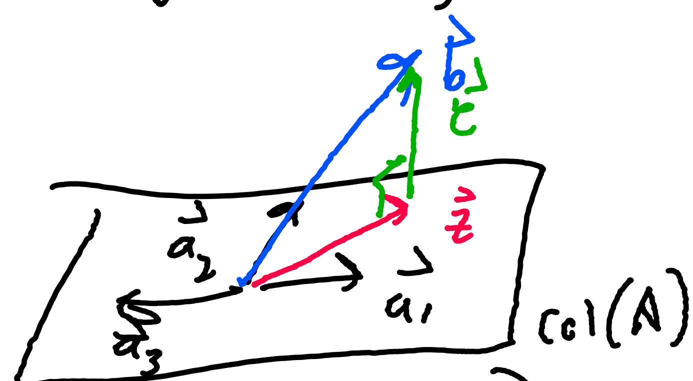
Yes, $\vec{z} \perp \vec{e}$.

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



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

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



$$\left.\begin{array}{l} \langle \vec{a}_1, \vec{e} \rangle = 0 \\ \langle \vec{a}_2, \vec{e} \rangle = 0 \\ \quad \vdots \\ \langle \vec{a}_n, \vec{e} \rangle = 0 \end{array}\right\} \Rightarrow \langle \vec{z}, \vec{e} \rangle$$

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

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

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

$$\vdots$$

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

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

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

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

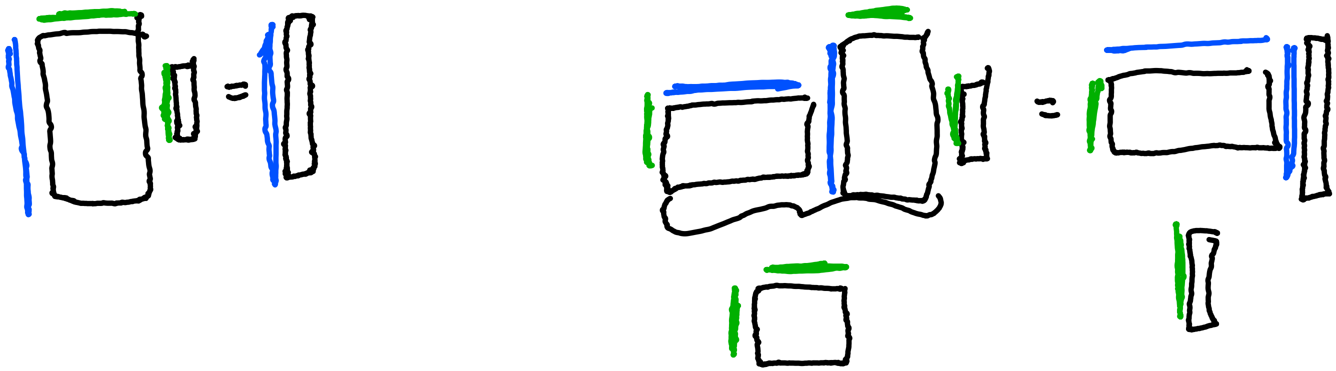
$$A^T(\vec{b} - \vec{z}) = \vec{0}$$

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

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

$$\underbrace{A^T\vec{b} = A^T A\vec{x}}_{\text{a system of equations}}$$

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



If A^TA is invertible (it's a square matrix)

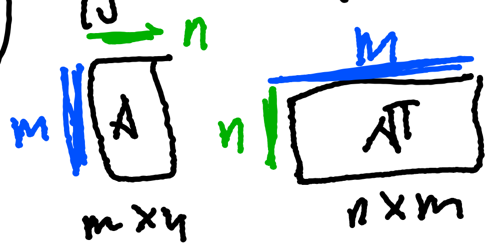
The solution $\vec{x}$ is: $\underbrace{(A^TA)^{-1}A^T\vec{b}}_{\text{not } \vec{x}}$

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

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

Is there ever a case A^TA is not square?

(A is tall) $m > n$



$\underset{n\times m \;\; m\times n}{A^TA} \Rightarrow \underset{n\times n}{\text{square}}$

A: A^TA is always square.

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

$A^TA =$

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

$A^TA =$

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

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

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

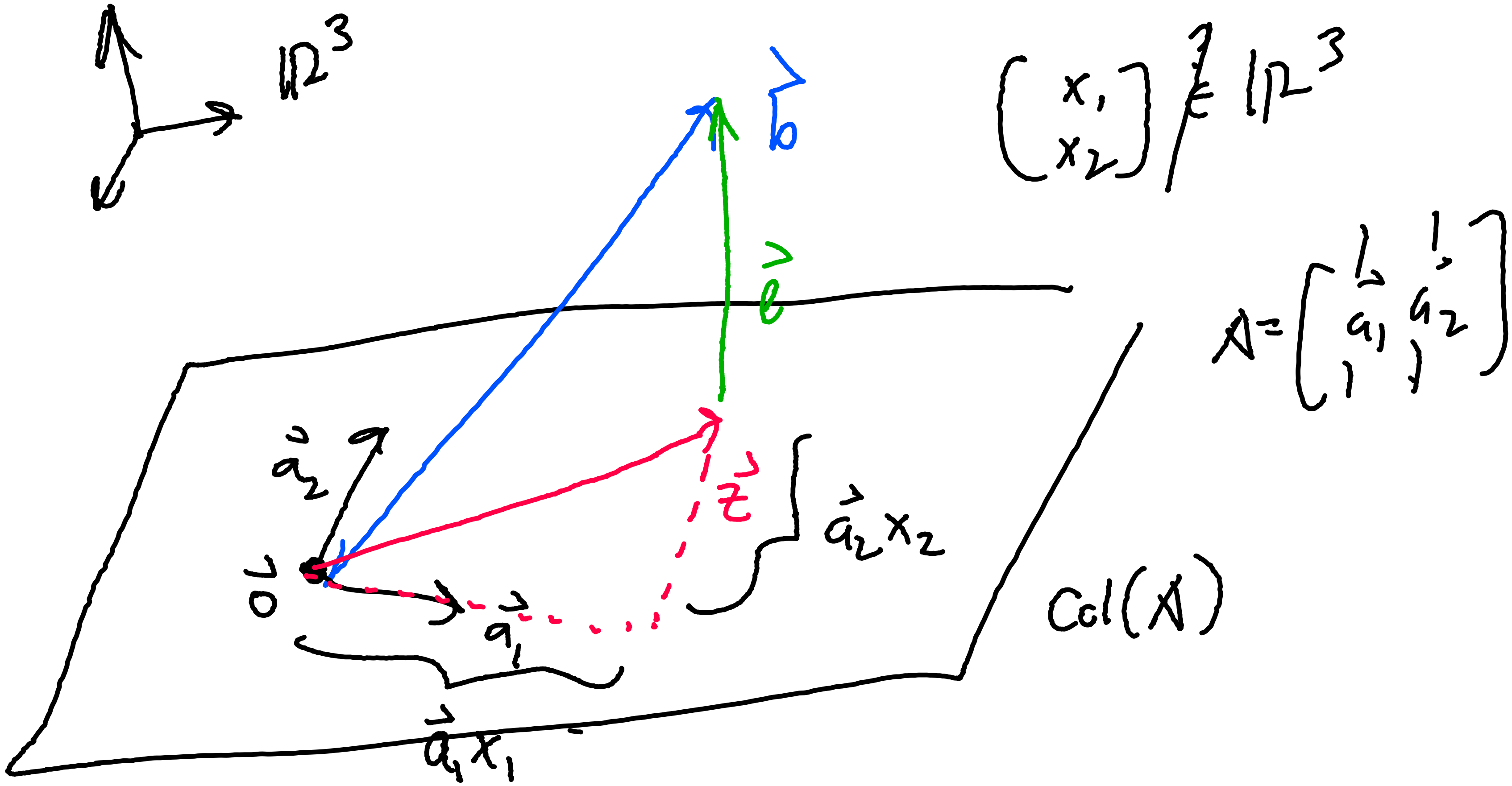
$\vec{z} = A(A^TA)^{-1}A^T\vec{b}$

vector projection

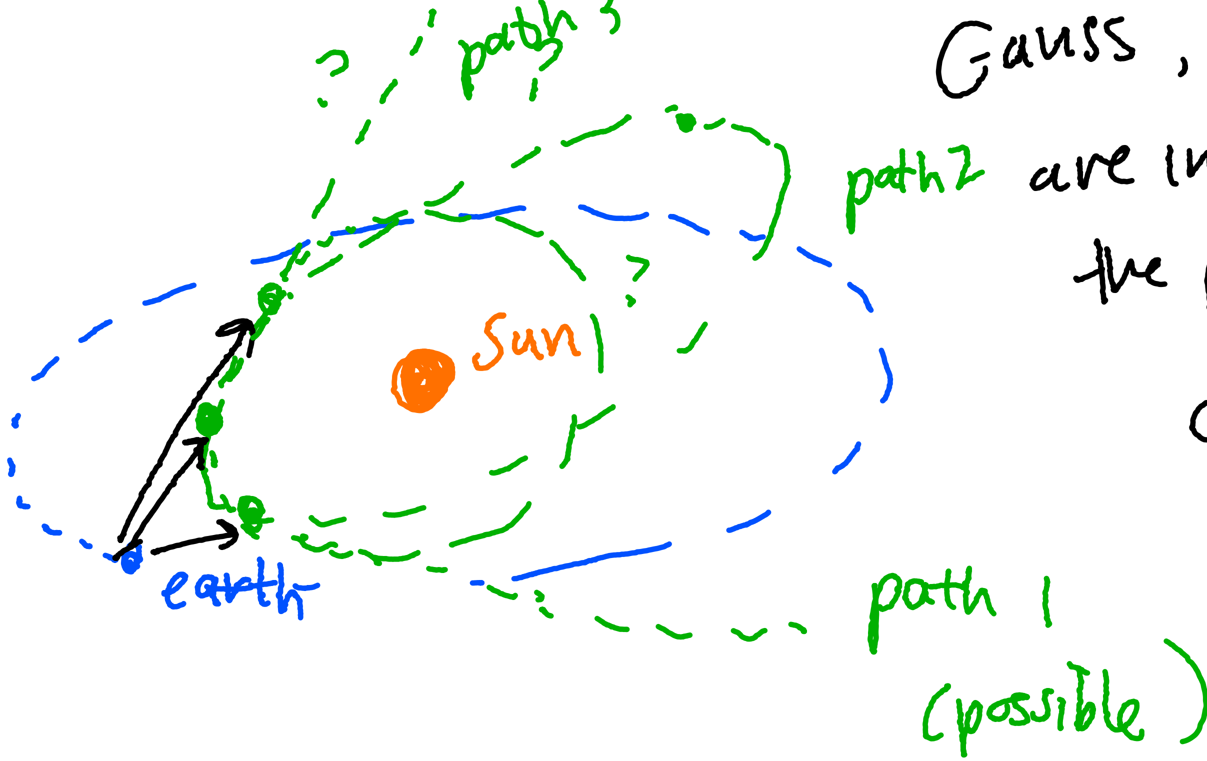
$\vec{z} = \text{proj}_{col(A)}\vec{b}$.

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

"coordinates of projection"

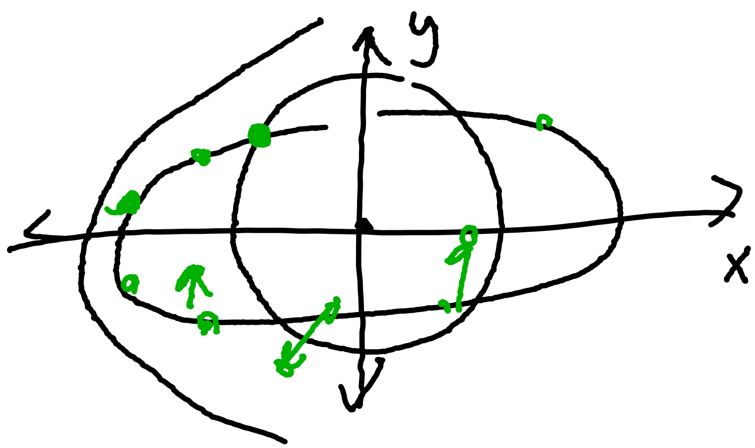


$\mathbb{R}^3$
$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \notin \mathbb{R}^3$
$\vec{b}$
$\vec{e}$
$A = \begin{bmatrix} | & | \\ a_1 & a_2 \\ | & | \end{bmatrix}$
$\vec{a}_2$
$\vec{z}$
$\vec{a}_2 x_2$
$\vec{0}$
$\vec{a}_1$
$\mathrm{Col}(A)$
$\vec{a}_1 x_1$



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

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



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

$$\underline{a}\,x_1^2 + \underline{b}\,y_1^2 + \underline{c}\,x_1 y_1 + \underline{d}\,x_1 + \underline{e}\,y_1 = 1$$

$$\vdots$$

$$\begin{bmatrix} x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & y_n^2 & x_n y_n & x_n & y_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \left.\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}\right\} n$$