

Welcome to EECS 16A!

Designing Information Devices and Systems I



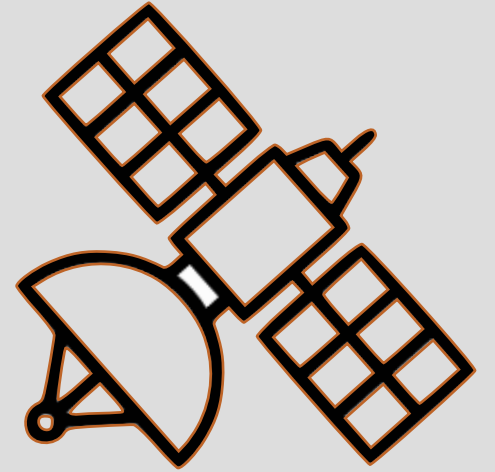
Ana Arias and Miki Lustig
Fall 2021

Lecture 13A
More Least squares



GPS

- 24 satellites
 - Known position
 - Time synchronized
 - 8 usually visible
- Problem:
 - Classify which satellite is transmitting
 - Estimate distance to GPS
 - Estimate position from noisy data
- Tools:
 - Inner product
 - Cross correlation
 - Least Squares



Least Squares

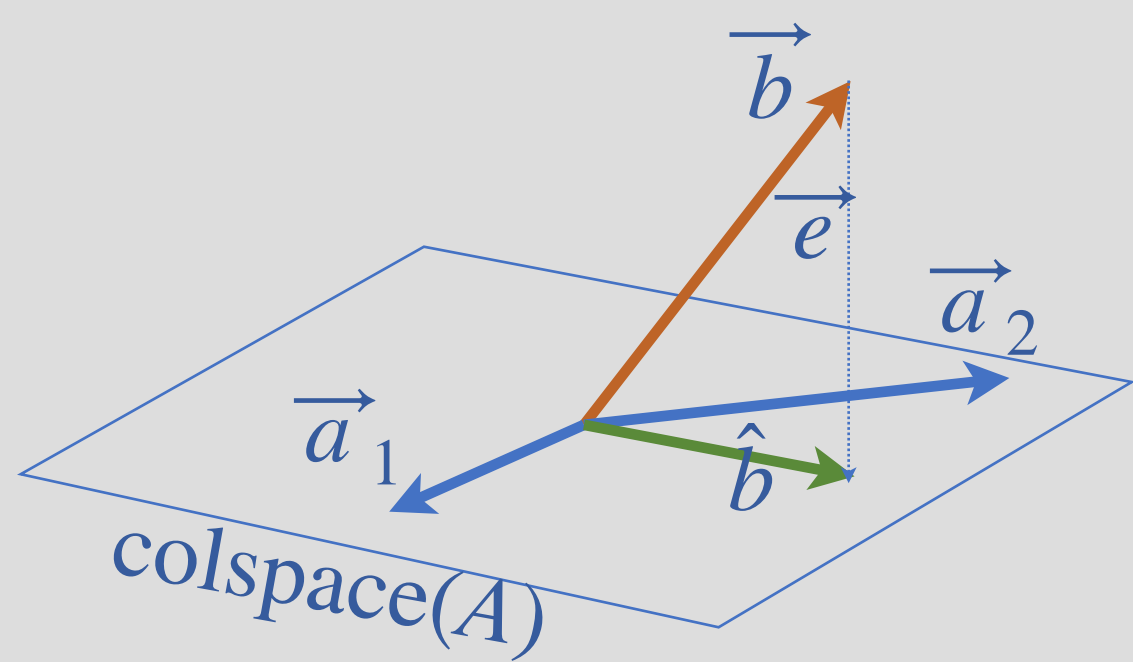
$$\operatorname{argmin}_{\vec{x}} \|\vec{e}\| = \|A\vec{x} - \vec{b}\|$$

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

$$\text{Since } \vec{e} \perp \operatorname{col}(A), \langle \vec{a}_i, \vec{e} \rangle = 0$$

$$\langle \vec{a}_i, \vec{b} - \hat{b} \rangle = 0$$

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



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

$$\begin{cases} A\vec{x} \in \operatorname{colspace}(A) \\ \text{Find } \hat{b} = A\hat{x} \end{cases}$$

Least Squares

$$\operatorname{argmin}_{\vec{x}} \|\vec{e}\| = \|A\vec{x} - \vec{b}\|$$

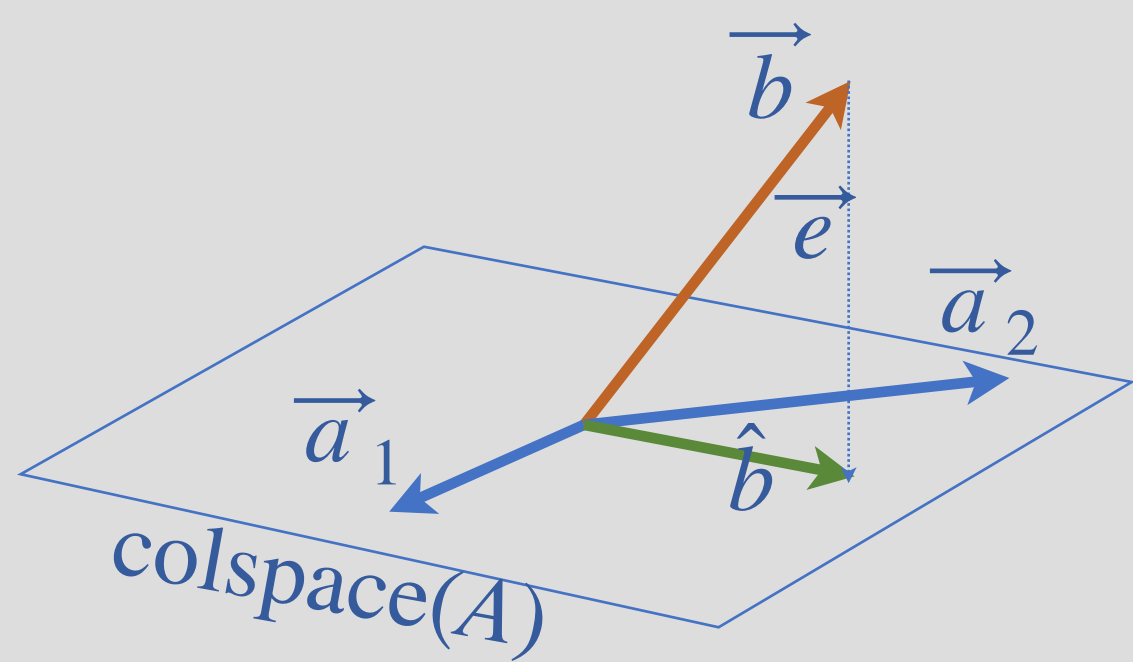
$$\vec{e} = \vec{b} - \hat{b}$$

$$\text{Since } \vec{e} \perp \operatorname{col}(A), \langle \vec{a}_i, \vec{e} \rangle = 0$$

$$\langle \vec{a}_i, \vec{b} - \hat{b} \rangle = 0$$

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

$$\begin{bmatrix} - & \vec{a}_1^T & - \\ - & \vec{a}_2^T & - \\ & \vdots & \\ - & \vec{a}_N^T & - \end{bmatrix} \begin{bmatrix} | \\ \vec{b} - \hat{b} \\ | \end{bmatrix} = 0$$



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

$A\vec{x} \in \operatorname{colspace}(A)$
Find $\hat{b} = A\hat{x}$

Least Squares

$$\begin{bmatrix} - & \vec{a}_1^T & - \\ - & \vec{a}_2^T & - \\ & \vdots & \\ - & \vec{a}_N^T & - \end{bmatrix} \begin{bmatrix} | \\ \vec{b} - \hat{\vec{b}} \\ | \end{bmatrix} = 0$$

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

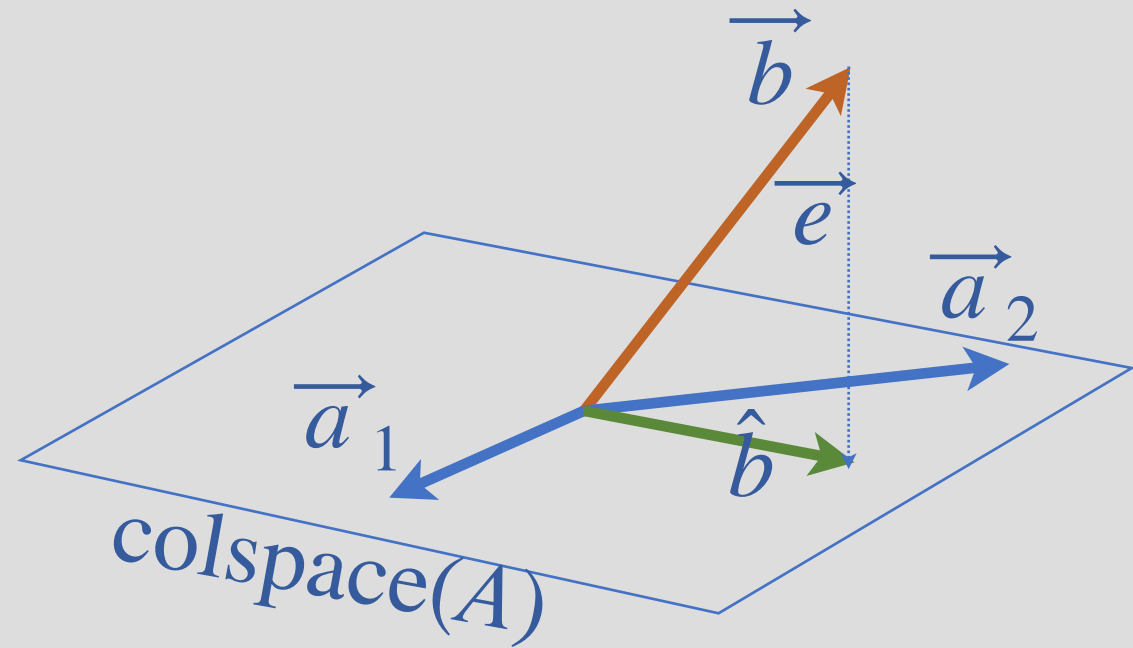
$$A^T \vec{b} - A^T A \hat{x} = 0$$

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

If A is full Rank, then $A^T A$ is invertible

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

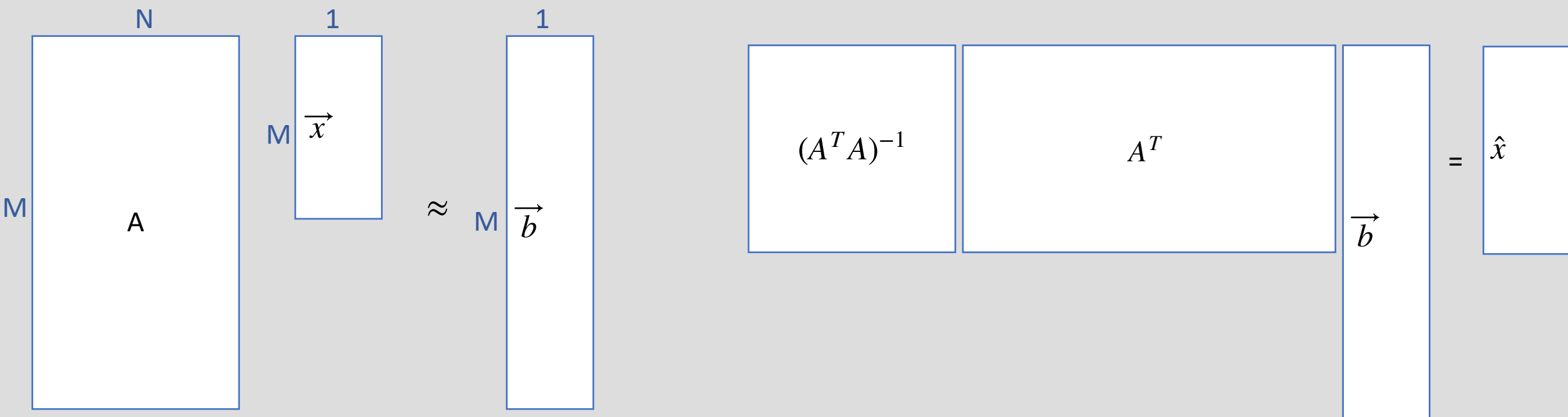
$$\hat{\vec{b}} = A(A^T A)^{-1} A^T \vec{b}$$



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

$A\vec{x} \in \text{colspace}(A)$
Find $\hat{\vec{b}} = A\hat{x}$

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





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



Example 1

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

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

$$\left. \begin{array}{l} 2x = 1 \\ 1 \cdot x = 1 \end{array} \right\}$$

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

Inconsistent! No solution

Least Squares:

$$\begin{aligned} \hat{x} &= (A^T A)^{-1} A^T \vec{b} \\ &= \begin{pmatrix} 1 \\ 5 \end{pmatrix} [2 \ 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \begin{pmatrix} 1 \\ 5 \end{pmatrix} \cdot 3 = \boxed{\frac{3}{5}} \end{aligned}$$

$$\begin{aligned} A^T A &= [2 \ 1] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5 \\ (A^T A)^{-1} &= \frac{1}{5} \end{aligned}$$

Example 2

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

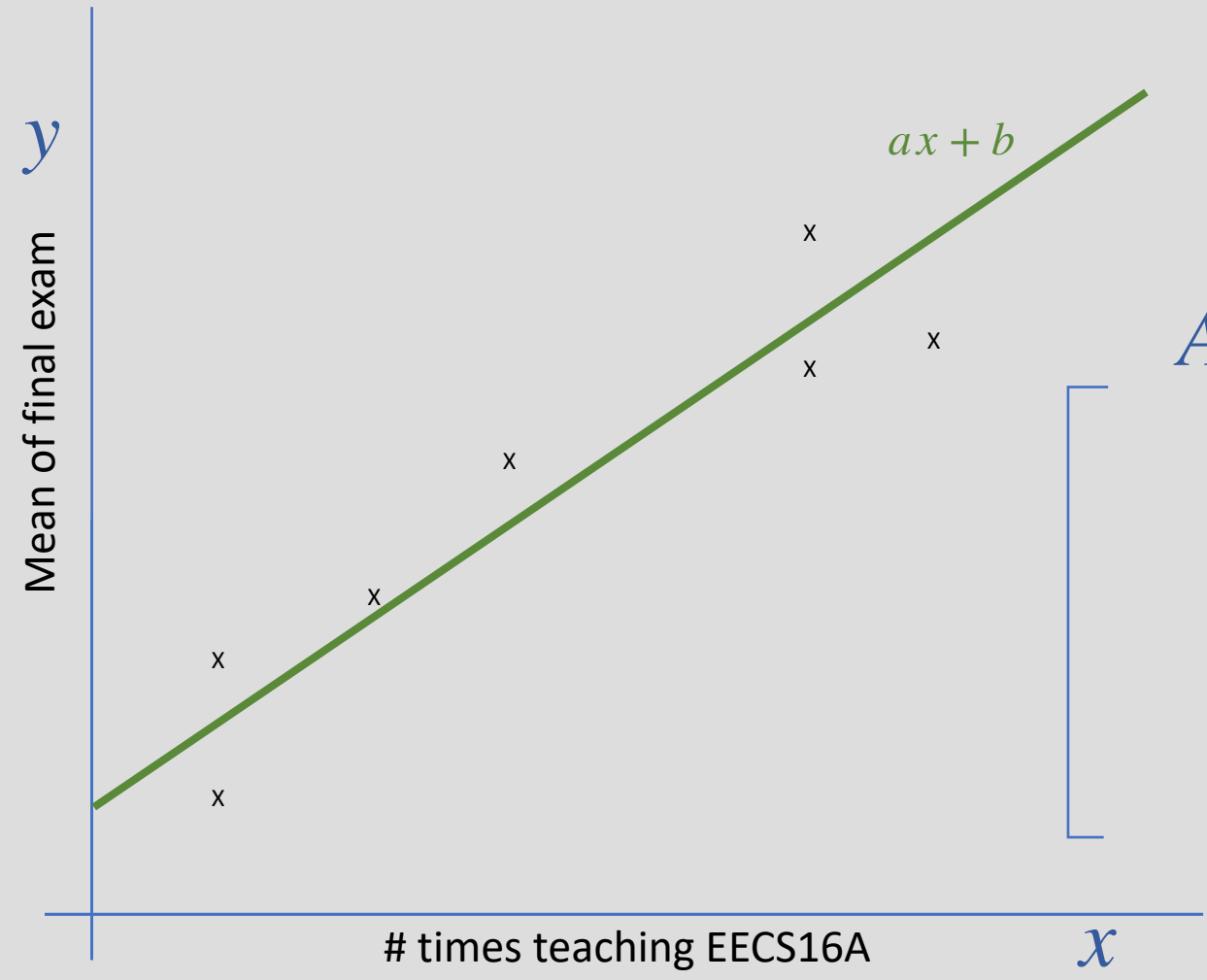
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \left. \begin{array}{l} x_1 = 1 \\ x_2 = 2 \\ x_2 = 3 \end{array} \right\} 2.5!$$

Least Squares:

$$\begin{aligned} \hat{x} &= (A^T A)^{-1} A^T \vec{b} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2.5 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} (A^T A)^{-1} &= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \right)^{-1} \\ &= \left(\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \end{aligned}$$

Example 3: Linear Regression



Model: $y = ax + b$

Known:

- Waller: (x_1, y_1)
- Sahai: (x_2, y_2)
- Alon: (x_4, y_4)
- Stojanovic: (x_5, y_5)
- Ranade: (x_6, y_6)
- Courtade: (x_7, y_7)
- Liu: (x_8, y_8)

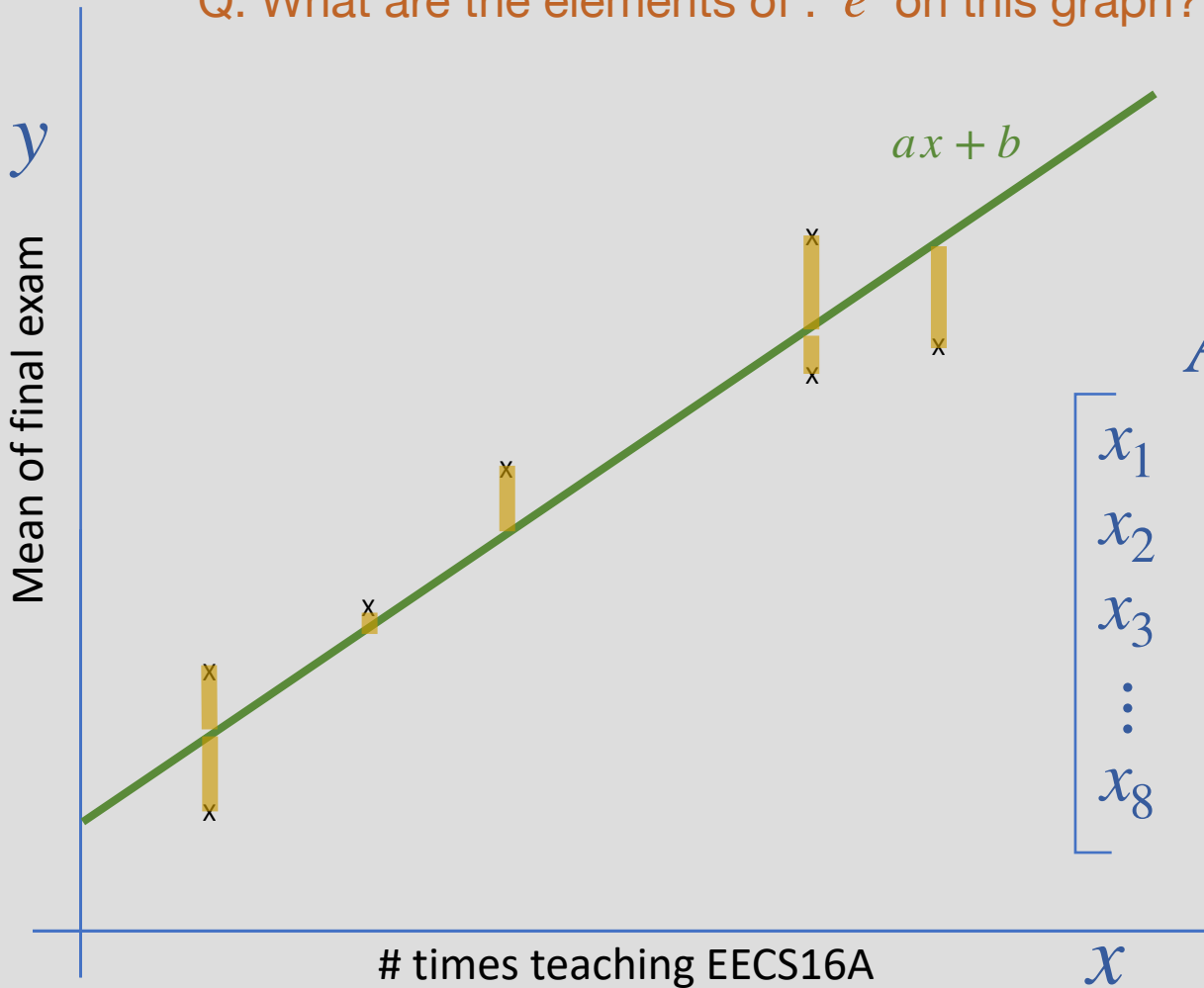
$$\begin{bmatrix} \\ \\ \\ \\ \\ \\ \\ \end{bmatrix} A \begin{bmatrix} \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \\ \\ \\ \\ \end{bmatrix}$$

$$\hat{p} = (A^T A)^{-1} A^T \vec{y}$$

Example 3: Linear Regression

$$\vec{e} = A\hat{p} - \vec{y}$$

Q: What are the elements of \vec{e} on this graph?



Model: $y = ax + b$

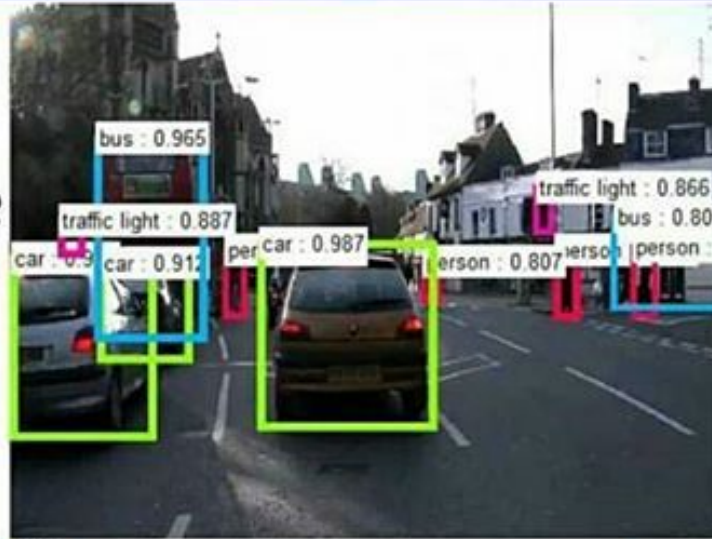
- Known:
- Waller: (x_1, y_1)
 - Sahai: (x_2, y_2)
 - Alon: (x_4, y_4)
 - Stojanovic: (x_5, y_5)
 - Ranade: (x_6, y_6)
 - Courtade: (x_7, y_7)
 - Liu: (x_8, y_8)

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ \vdots & \vdots \\ x_8 & 1 \end{bmatrix} \vec{p} = \vec{y}$$

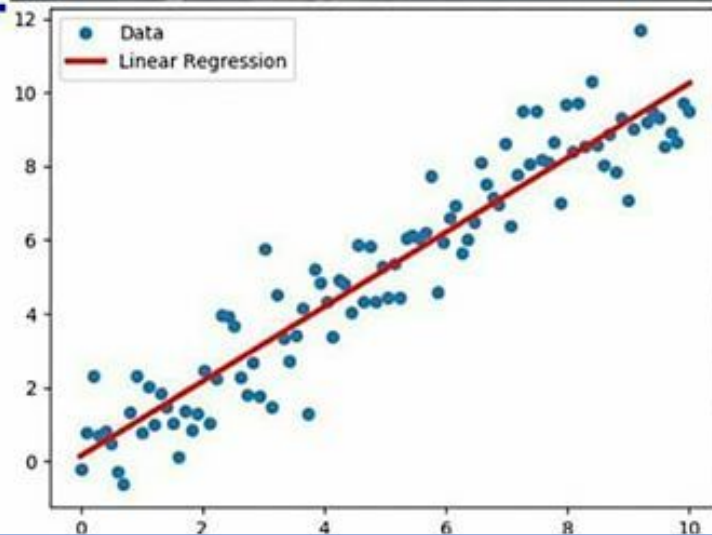
$$\hat{p} = (A^T A)^{-1} A^T \vec{y}$$

Online Courses

What they promise
you will learn



What you actually
learn



BUT, not
everything fits to a
line!?!

Example 4: Regression

Gauss found Ceres by using Kepler's laws:

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

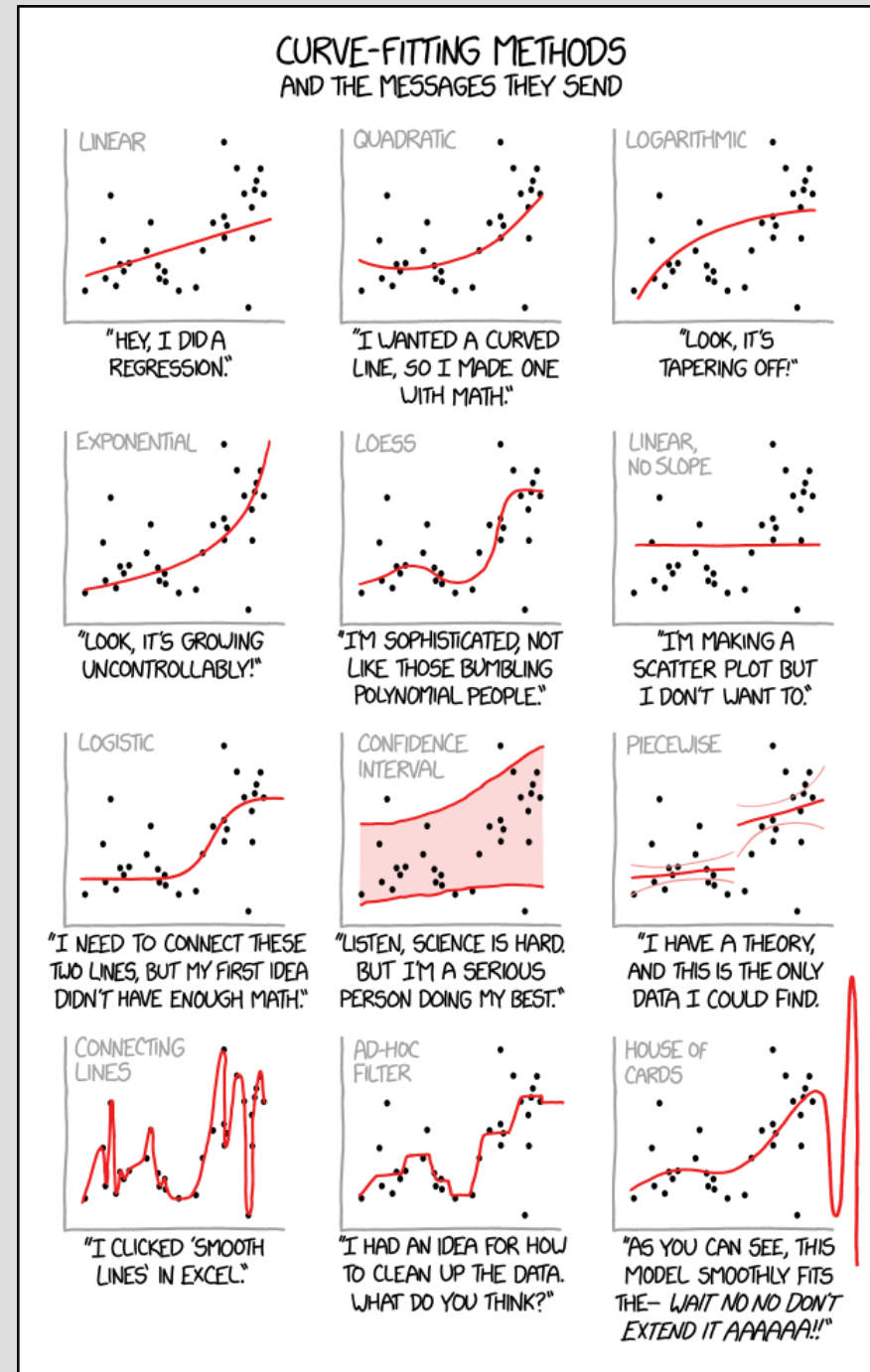
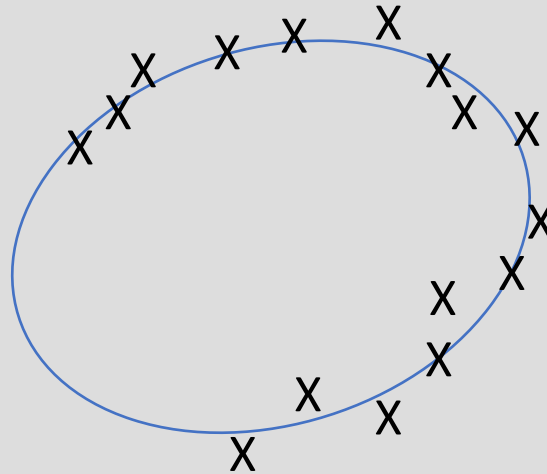
Q: Is this a linear fit?

A: Yes!

Knowns: $(x_1, y_1) (x_2, y_2) \dots (x_N, y_N)$

Unknowns: $\vec{p} = [a \ b \ c \ d \ e]^T$

$$\begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}
 \begin{matrix} A \\ \\ \\ \\ \end{matrix}
 \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}
 \vec{p}
 =
 \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}
 \vec{y}$$



Example 4: Regression

Gauss found Ceres by using Kepler's laws:

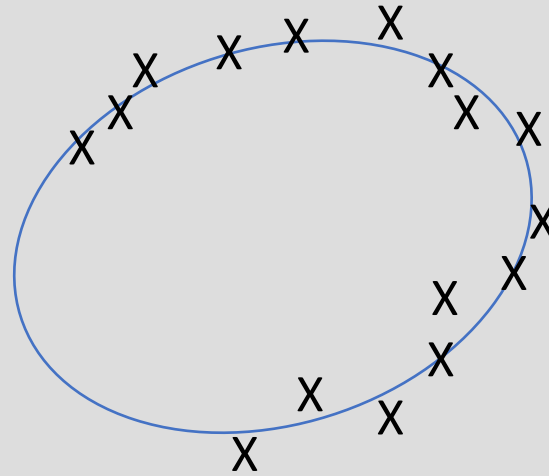
Model: $ax^2 + by^2 + cxy + dx + ey = 1$

Q: Is this a linear fit?

A: Yes!

Knowns: $(x_1, y_1) (x_2, y_2) \dots (x_N, y_N)$

Unknowns: $\vec{p} = [a \ b \ c \ d \ e]^T$



$$\begin{bmatrix} x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 \\ x_2^2 & y_2^2 & x_2 y_2 & x_2 & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_N^2 & y_N^2 & x_N y_N & x_N & y_N \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\hat{p} = (A^T A)^{-1} A^T \vec{y}$$

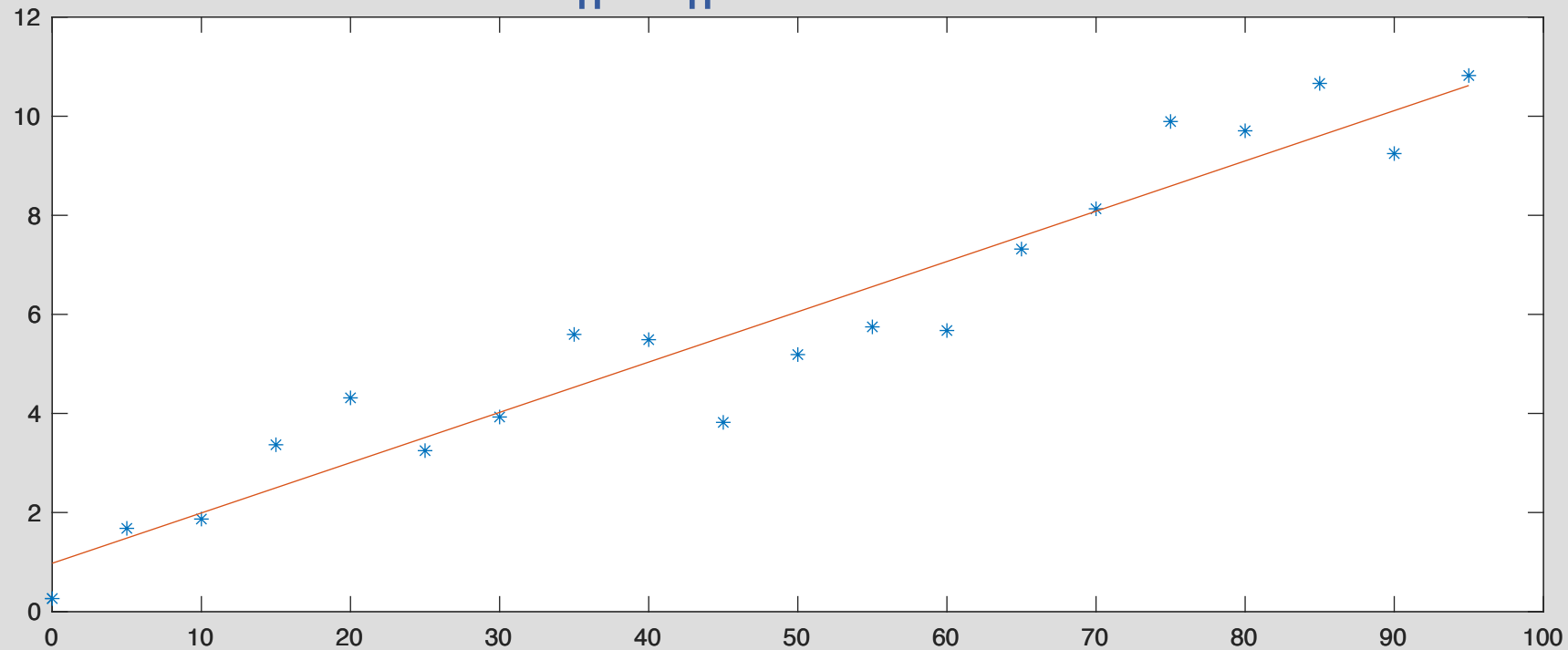
Example 5: Over Fitting

- Consider noisy measurements of $y = 0.1x + 1$:

Model: $y = ax + b$

$$\vec{p} = [0.1015 \quad 0.9757]^T$$

$$\|\vec{e}\| = 3.85$$

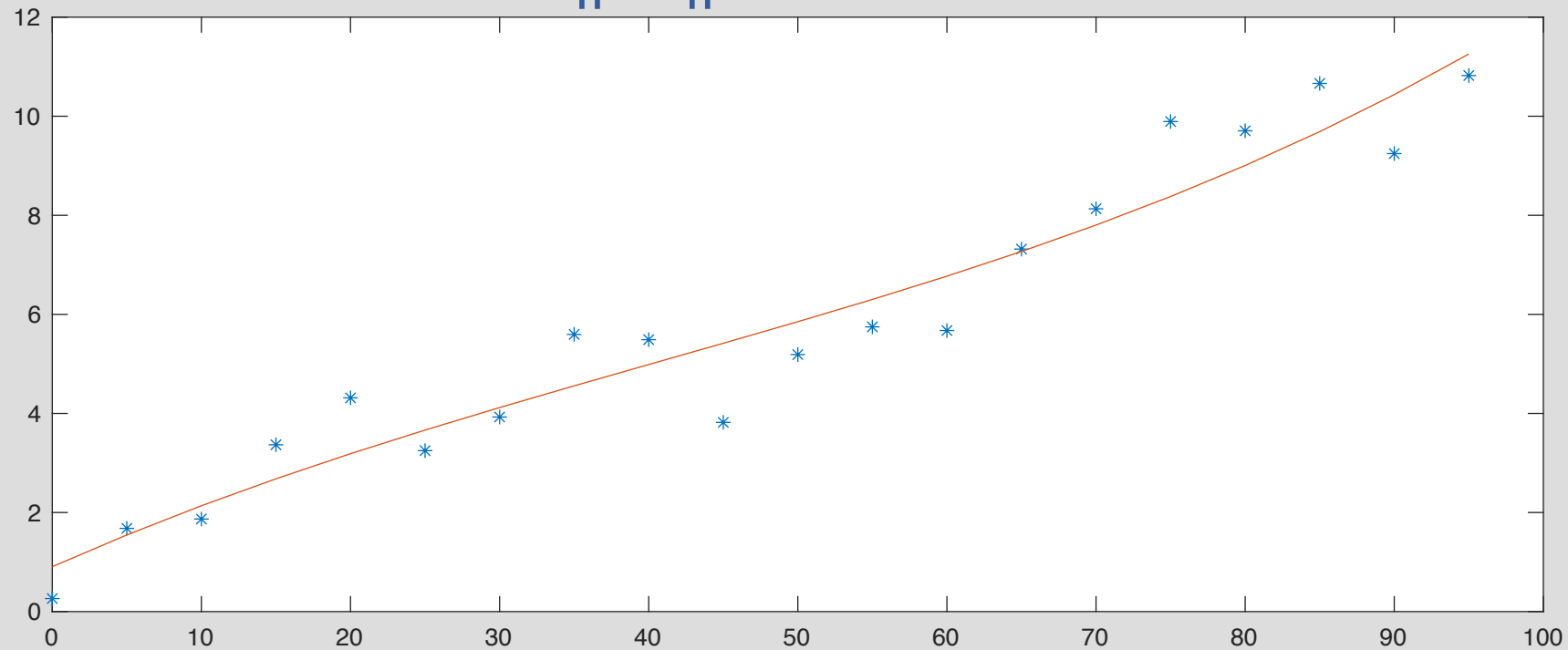


Example 5: Over Fitting

- Consider noisy measurements of $y = 0.1x + 1$:

Model: $y = ax^3 + bx^2 + cx + d$

$$\|\vec{e}\| = 3.71$$

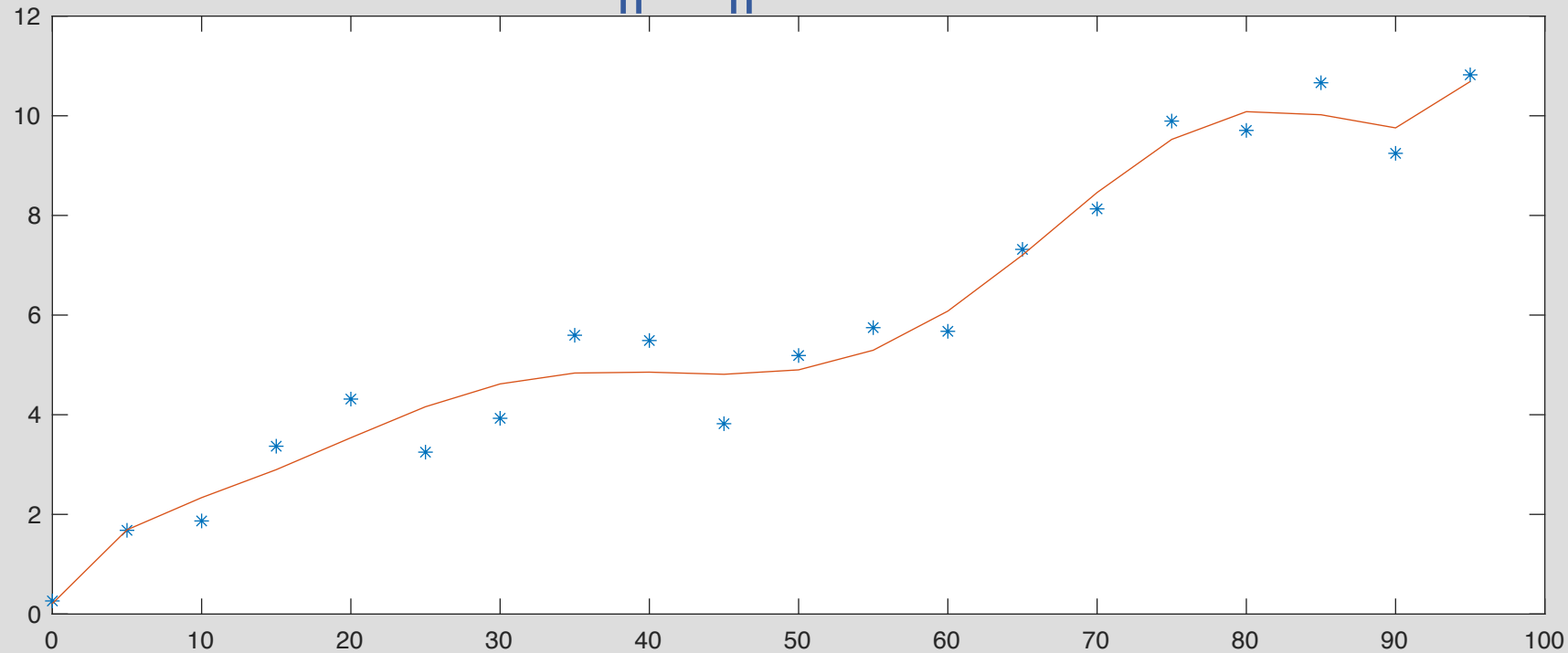


Example 5: Over Fitting

- Consider noisy measurements of $y = 0.1x + 1$:

Model: $y = ax^7 + bx^6 + cx^5 + dx^4 + ex^3 + fx^2 + gx + h$

$$\|\vec{e}\| = 2.42$$



Example 5: Exponential Regression

Model: $y = ce^{ax}$

Q: Is this a linear fit?

A: No! But, can be made linear.....

New Model: $\log(y) = \log c + ax = b + ax$

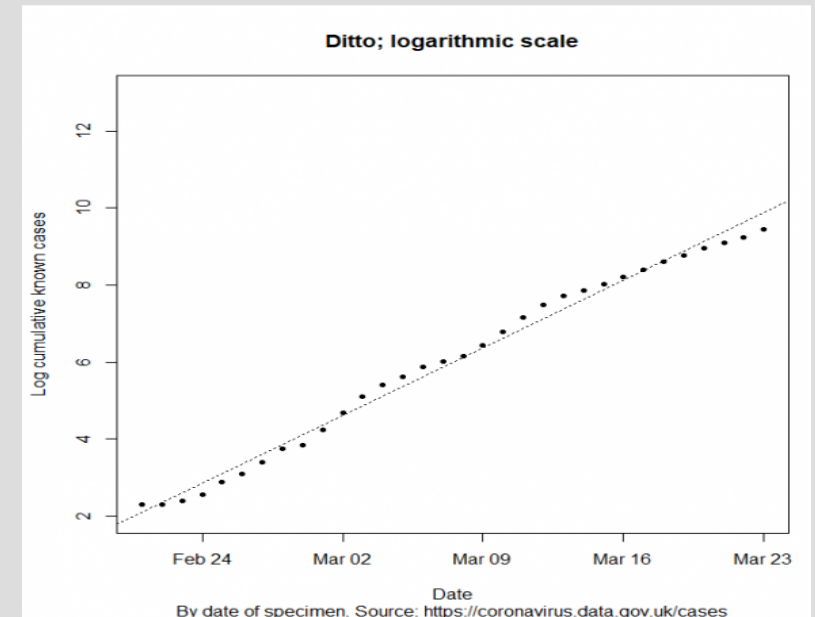
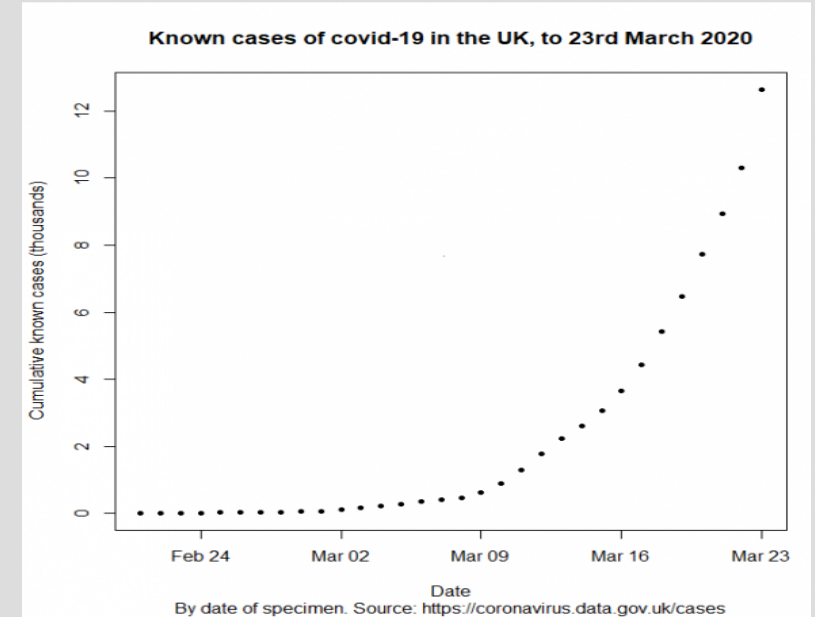
Knowns: $(x_1, \log(y_1)) (x_2, \log(y_2)) \dots (x_N, \log(y_N))$

Unknowns: $\vec{p} = [a \ b]^T$

$$\begin{bmatrix} A & \vec{p} \\ \begin{matrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{matrix} & \begin{bmatrix} a \\ b \end{bmatrix} \\ = & \begin{bmatrix} \vec{y} \\ \log y_1 \\ \log y_2 \\ \vdots \\ \log y_N \end{bmatrix} \end{bmatrix}$$

$$\hat{p} = (A^T A)^{-1} A^T \vec{y}$$

$$\hat{c} = e^{\hat{b}}$$



Multi-Lateration

$$2(\vec{a}_1 - \vec{a}_2)^T \vec{x} - 2C^2 \Delta\tau_2 \tau_1 = \|\vec{a}_1\|^2 - \|\vec{a}_2\|^2 + C^2(\Delta\tau_2)^2$$

$$2(\vec{a}_1 - \vec{a}_3)^T \vec{x} - 2C^2 \Delta\tau_3 \tau_1 = \|\vec{a}_1\|^2 - \|\vec{a}_3\|^2 + C^2(\Delta\tau_3)^2$$

$$2(\vec{a}_1 - \vec{a}_4)^T \vec{x} - 2C^2 \Delta\tau_4 \tau_1 = \|\vec{a}_1\|^2 - \|\vec{a}_4\|^2 + C^2(\Delta\tau_4)^2$$

$$2(\vec{a}_1 - \vec{a}_5)^T \vec{x} - 2C^2 \Delta\tau_5 \tau_1 = \|\vec{a}_1\|^2 - \|\vec{a}_5\|^2 + C^2(\Delta\tau_5)^2$$

More equations than unknowns

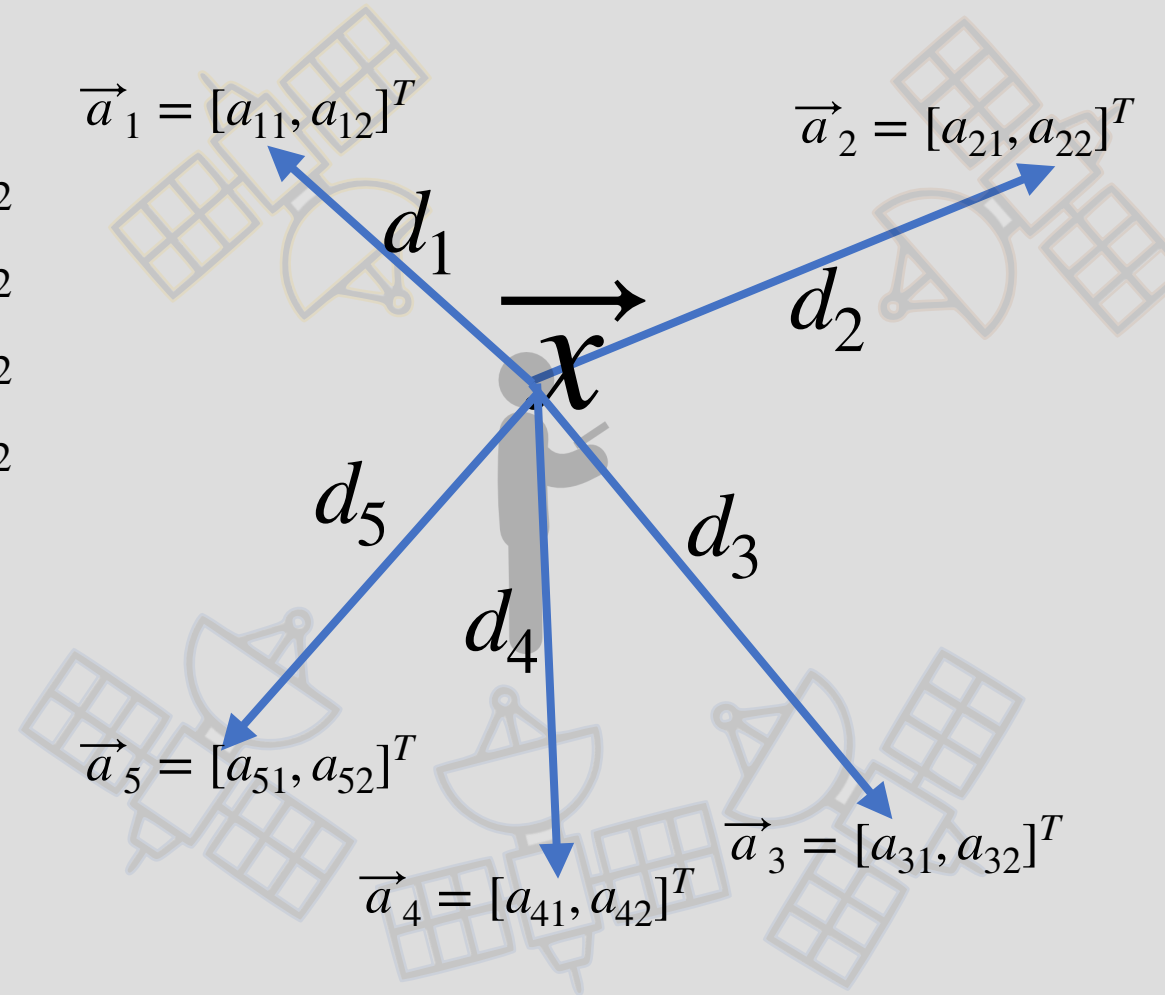
$$\begin{matrix} \boxed{A} \\ \vec{p} \end{matrix} = \begin{matrix} \vec{b} \end{matrix}$$

Over-determined — Solve via Least-Squares

Q: How do we know if $A^T A$ is invertible?

A: if A is full rank!?!?

$$\hat{p} = (A^T A)^{-1} A^T \vec{y}$$



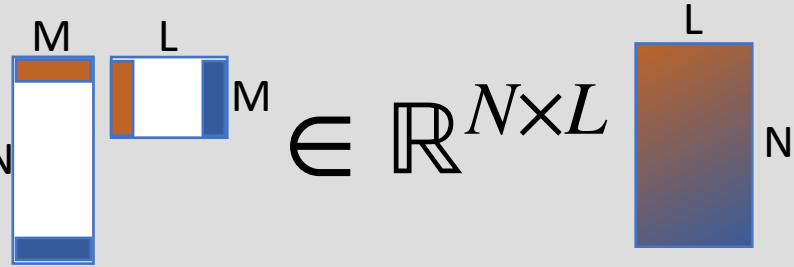
Matrix Transposes

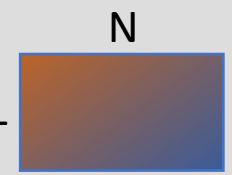
$$A \in \mathbb{R}^{N \times M}$$


$$B \in \mathbb{R}^{M \times L}$$


$$A^T \in \mathbb{R}^{M \times N}$$


$$B^T \in \mathbb{R}^{L \times M}$$


$$AB \in \mathbb{R}^{N \times L}$$


$$(AB)^T = \left(\begin{matrix} \text{Diagram of } A \\ \text{Diagram of } B \end{matrix} \right)^T = \left(\begin{matrix} \text{Diagram of } AB \end{matrix} \right)^T \in \mathbb{R}^{L \times N}$$


Matrix Transposes

$$(AB)^T \left(\begin{array}{|c|c|} \hline \color{red}{M} & \color{red}{L} \\ \hline \color{blue}{N} & \color{blue}{M} \\ \hline \end{array} \right)^T = \left(\begin{array}{|c|} \hline \color{red}{L} \\ \hline \color{blue}{N} \\ \hline \end{array} \right)^T \in \mathbb{R}^{L \times N} \quad \begin{array}{|c|} \hline \color{red}{N} \\ \hline \color{blue}{L} \\ \hline \end{array}$$

$$B^T A^T \begin{array}{|c|c|} \hline \color{red}{L} & \color{red}{M} \\ \hline \color{blue}{M} & \color{blue}{N} \\ \hline \end{array} \in \mathbb{R}^{L \times N} \quad \begin{array}{|c|} \hline \color{red}{N} \\ \hline \color{blue}{L} \\ \hline \end{array}$$

$$(AB)^T = B^T A^T$$

Invertibility of $A^T A$

- Invertible \Rightarrow Trivial null space \Rightarrow Linear independent cols/rows....

The matrix $A^T A$ is invertible iff $\text{Null}(A^T A) = \vec{0}$

Theorem: $\text{Null}(A^T A) = \text{Null}(A)$

Proof: (1) show that if $\vec{w} \in \text{Null}(A)$, then $\vec{w} \in \text{Null}(A^T A)$
(2) show that if $\vec{v} \in \text{Null}(A^T A)$, then $\vec{v} \in \text{Null}(A)$

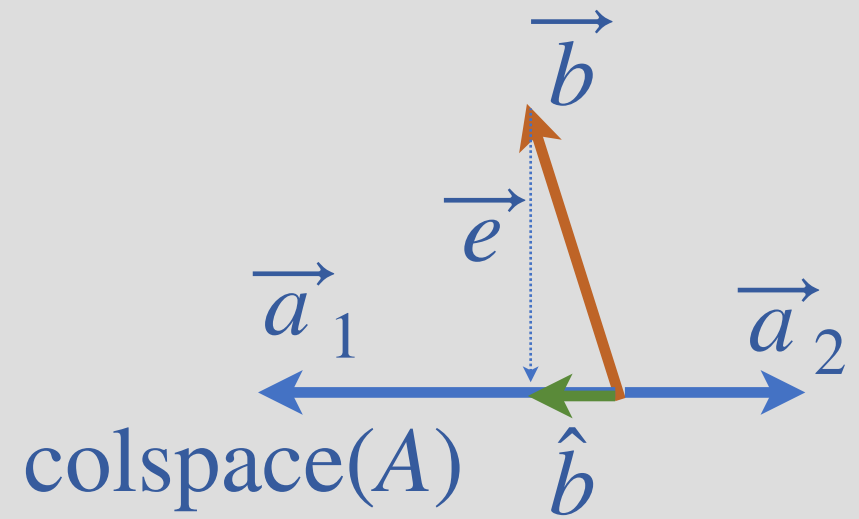
$$\begin{aligned} (1). \quad & \vec{w} \in \text{Null}(A) \\ & A\vec{w} = \vec{0} \\ & A^T A\vec{w} = A^T \vec{0} \\ & A^T A\vec{w} = \vec{0} \quad \checkmark \end{aligned}$$

$$\begin{aligned} (2). \quad & \vec{v} \in \text{Null}(A^T A) \\ & A^T A\vec{v} = \vec{0} \quad \text{Need to show } A\vec{v} = \vec{0} \\ & \quad \text{Or... } \|A\vec{v}\| = 0 \\ & \|A\vec{v}\|^2 = (A\vec{v})^T (A\vec{v}) \\ & \quad = \vec{v}^T A^T (A\vec{v}) \\ & \quad = \vec{v}^T (A^T A\vec{v}) = 0 \quad \checkmark \end{aligned}$$

Invertibility of $A^T A$

- What if $A^T A$ is not invertible

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



A: \hat{x} will have infinite solutions with the same $\vec{e} = A\hat{x} - \vec{b}$