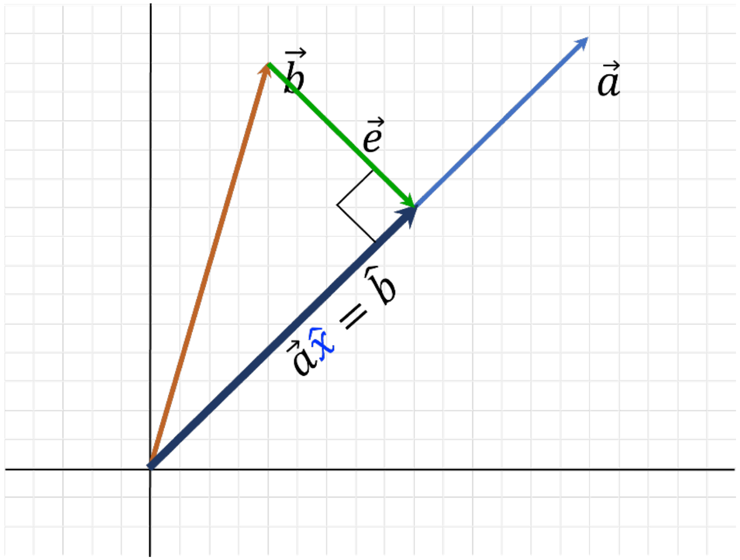
$y = mx + b$

Is My Favorite One-Liner

EECS 16A
More Least Squares

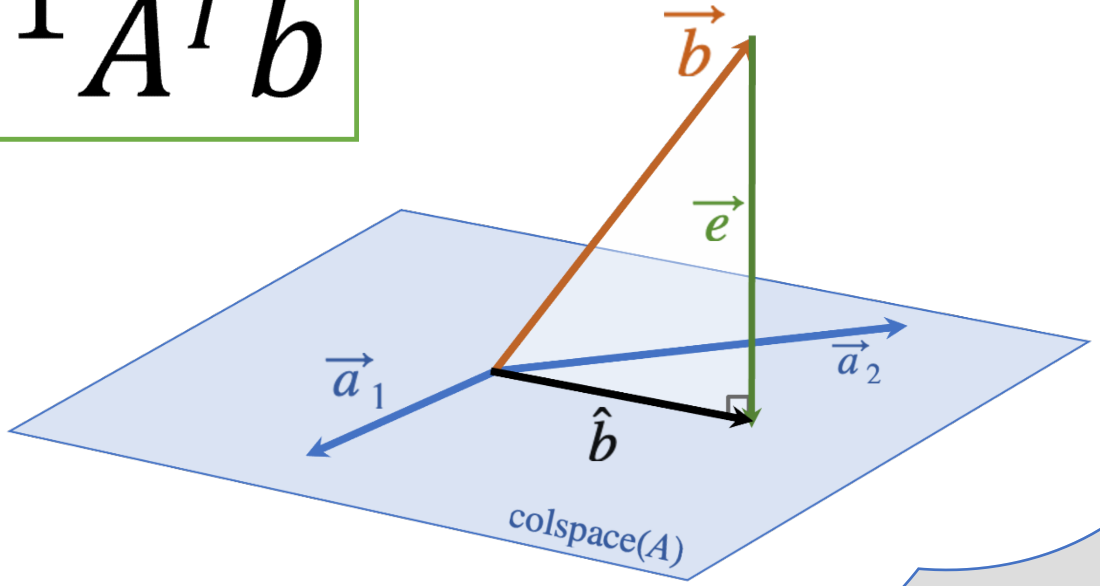
Given vectors $\vec{a}, \vec{b}$,we say that the **orthogonal projection** of $\vec{b}$ onto $\vec{a}$ is:

$$\mathrm{Proj}_{\vec{a}}(\vec{b}) = \frac{\vec{a}^T \vec{b}}{\| \vec{a} \|^2} \vec{a}$$



Least squares solution to an overdetermined problem:

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



Demo: fitting Facebook stock data to a line

Linear regression:
(fit to a line)

"known" measured data

$$y = mx + c$$

slope

intercept

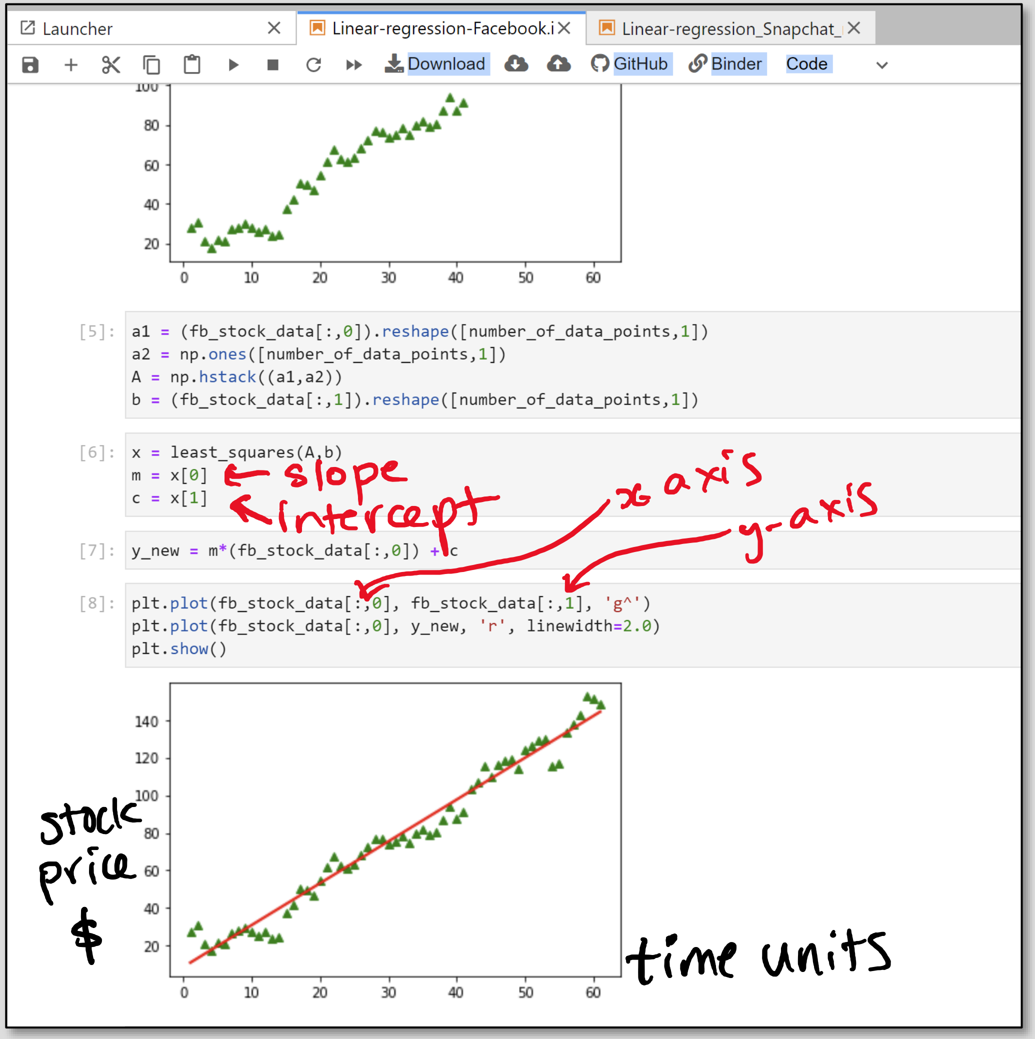
"unknowns" to solve for

write some eqns:

$$y_1 = m x_1 + c$$
$$y_2 = m x_2 + c$$
$$\vdots$$
$$y_N = m x_N + c$$

solve for

$$\underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_N & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m \\ c \end{bmatrix}}_{\vec{w}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}}_{\vec{y}} \rightarrow \text{LS} \quad \hat{w} = (A^T A)^{-1} A^T \vec{y}$$



```
[5]: a1 = (fb_stock_data[:,0]).reshape([number_of_data_points,1])
     a2 = np.ones([number_of_data_points,1])
     A = np.hstack((a1,a2))
     b = (fb_stock_data[:,1]).reshape([number_of_data_points,1])

[6]: x = least_squares(A,b)
     m = x[0]
     c = x[1]

[7]: y_new = m*(fb_stock_data[:,0]) + c

[8]: plt.plot(fb_stock_data[:,0], fb_stock_data[:,1], 'g^')
     plt.plot(fb_stock_data[:,0], y_new, 'r', linewidth=2.0)
     plt.show()
```

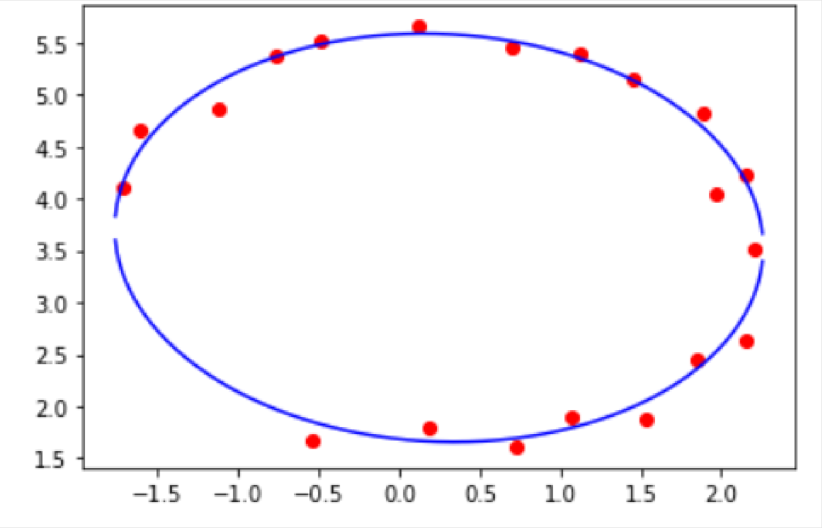
Let's be lazy and do it with Python:

Fitting to non-linear curves can be a linear problem!

unknown

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

Known
(measurements)



Example: Over Fitting

every 10 jokes improves score 1 pt.!

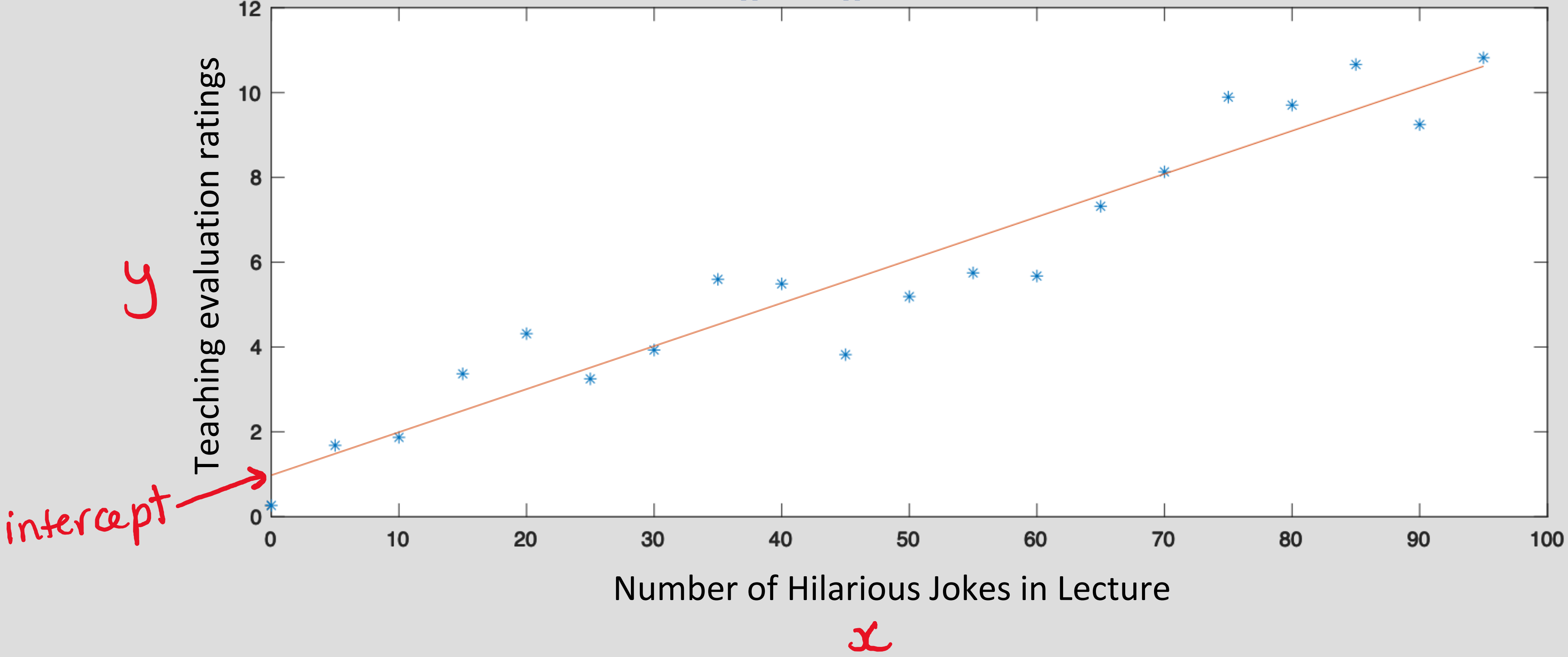
1 point is minimum

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

Model: $y = ax + b$

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

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

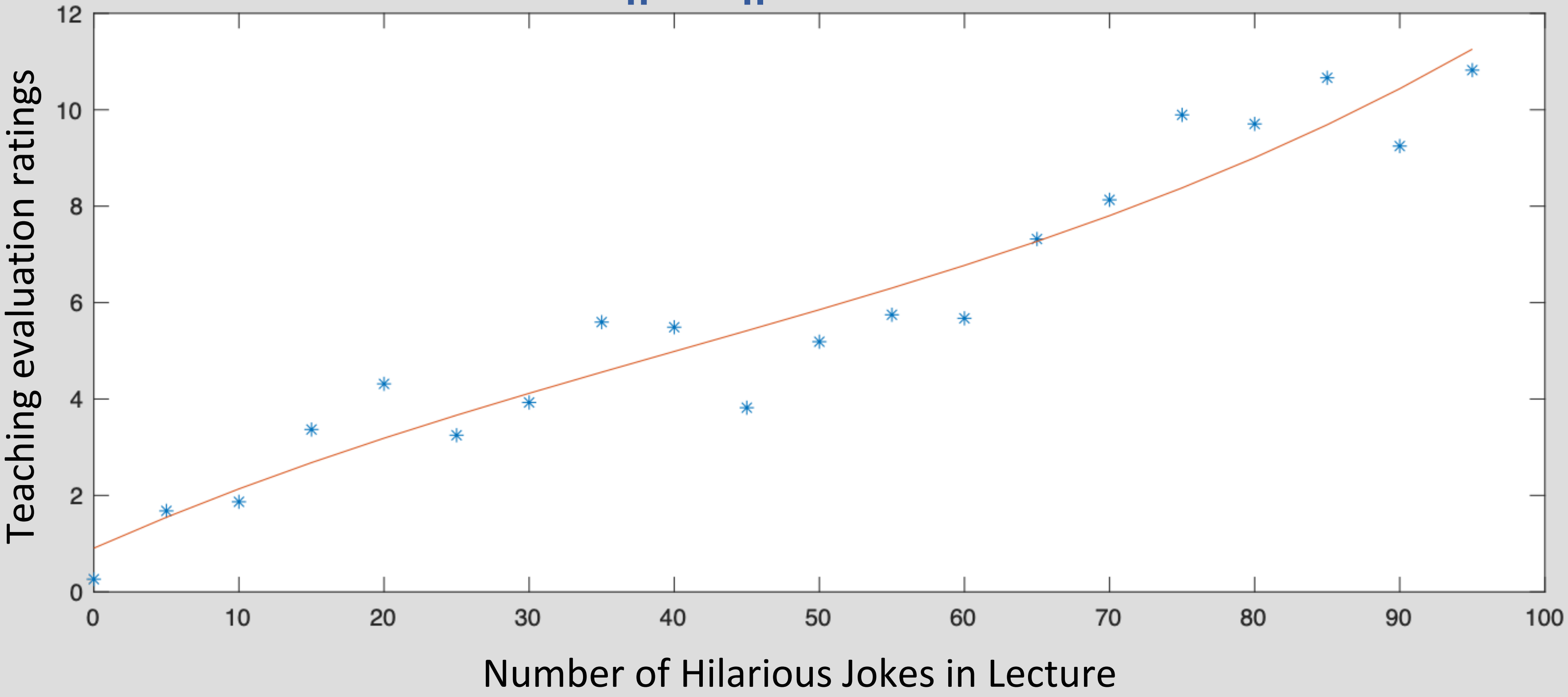


Example: Over Fitting

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

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

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

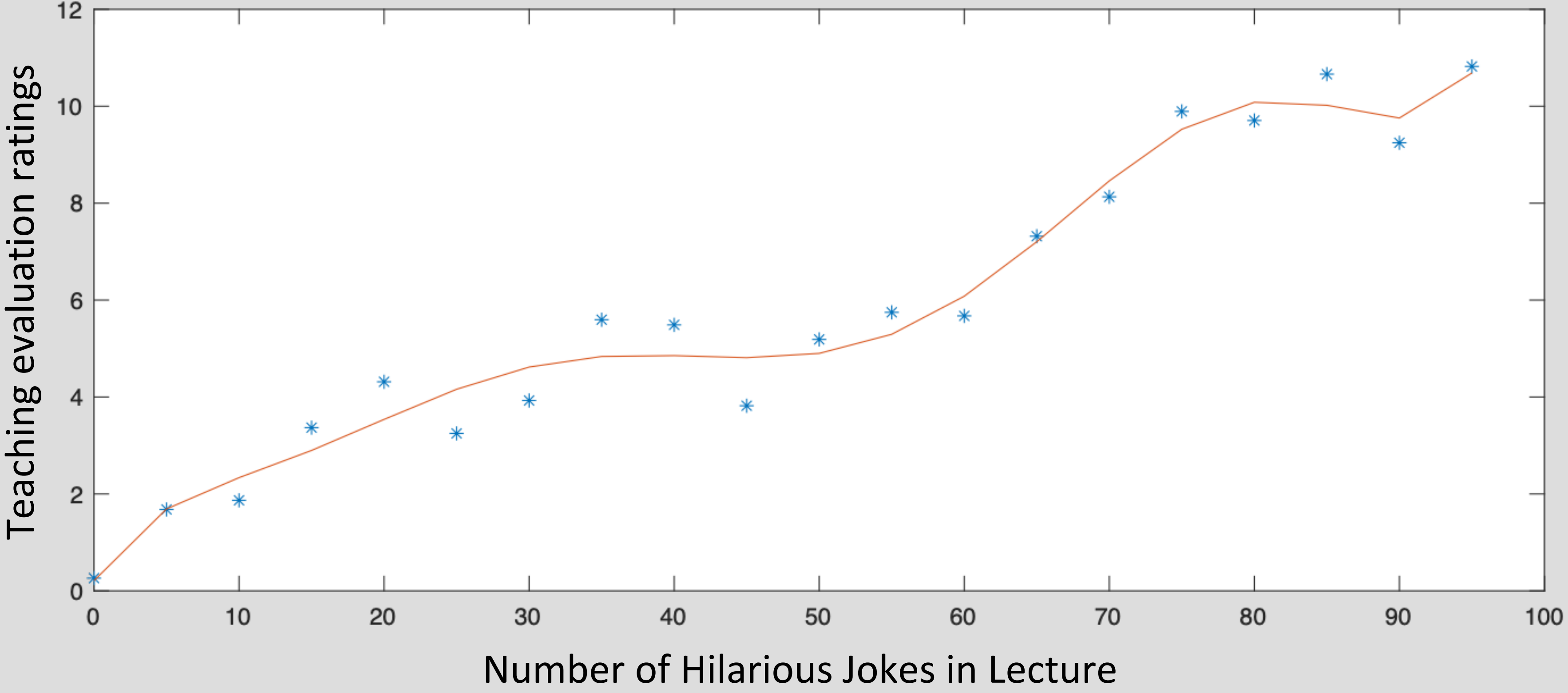


Example: Over Fitting

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

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

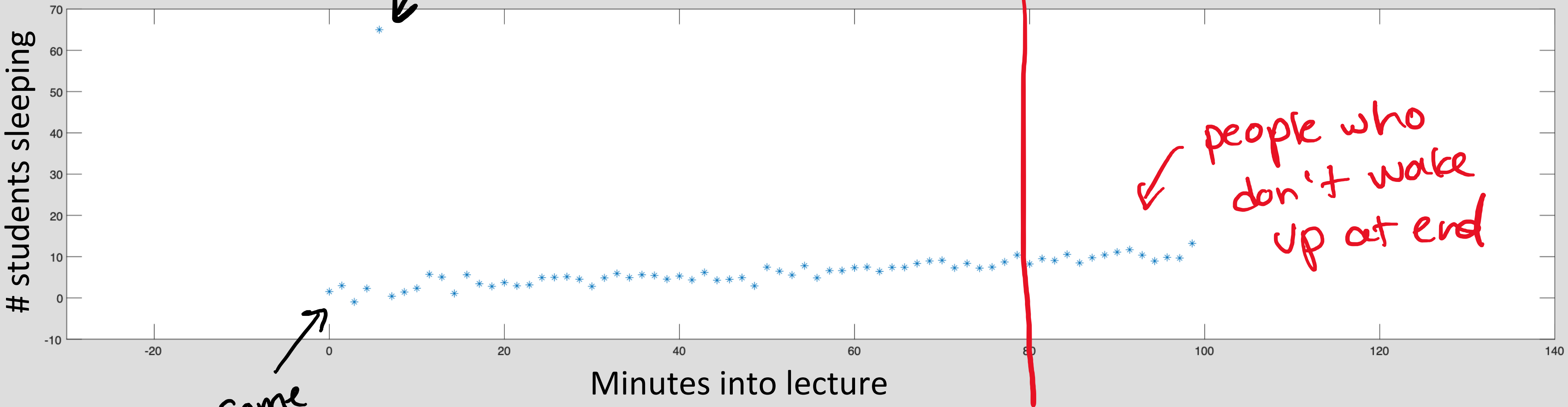
$$\| \vec{e} \| = 2.42$$



Example: Outlier

Model: $y = ax + b$

OUTLIER (bad measurement?)



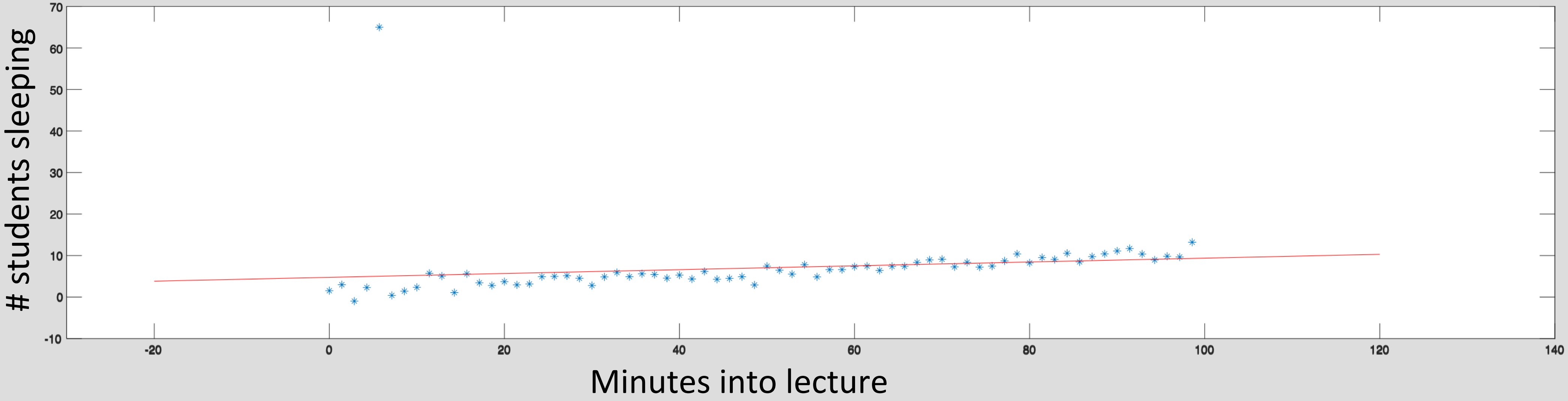
people who don't wake up at end

Some fall asleep right away (not my fault)

Lec is only 80 min

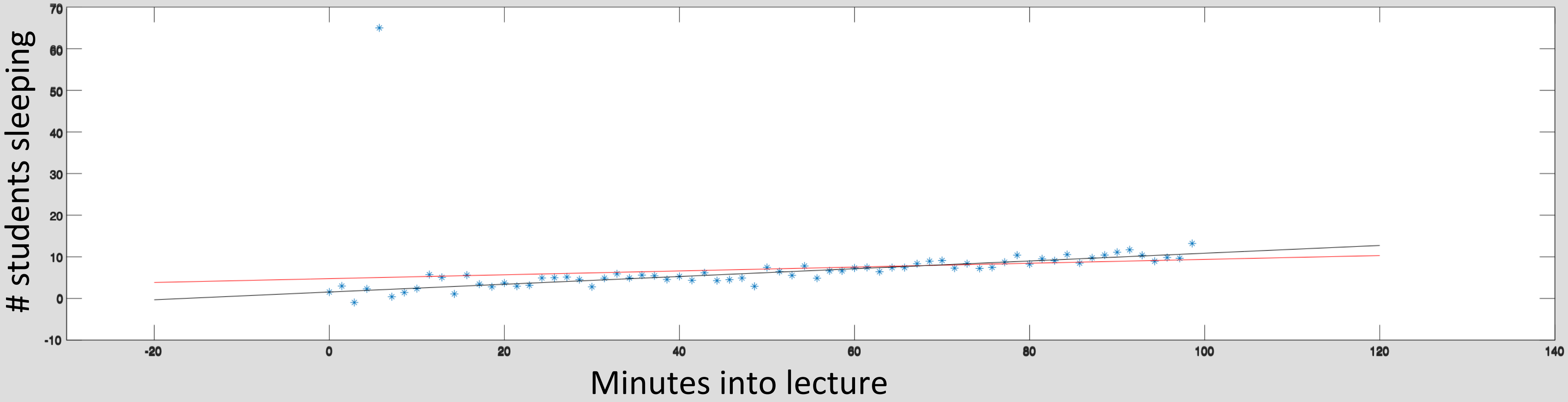
Example: Outlier

Model: $y = ax + b$



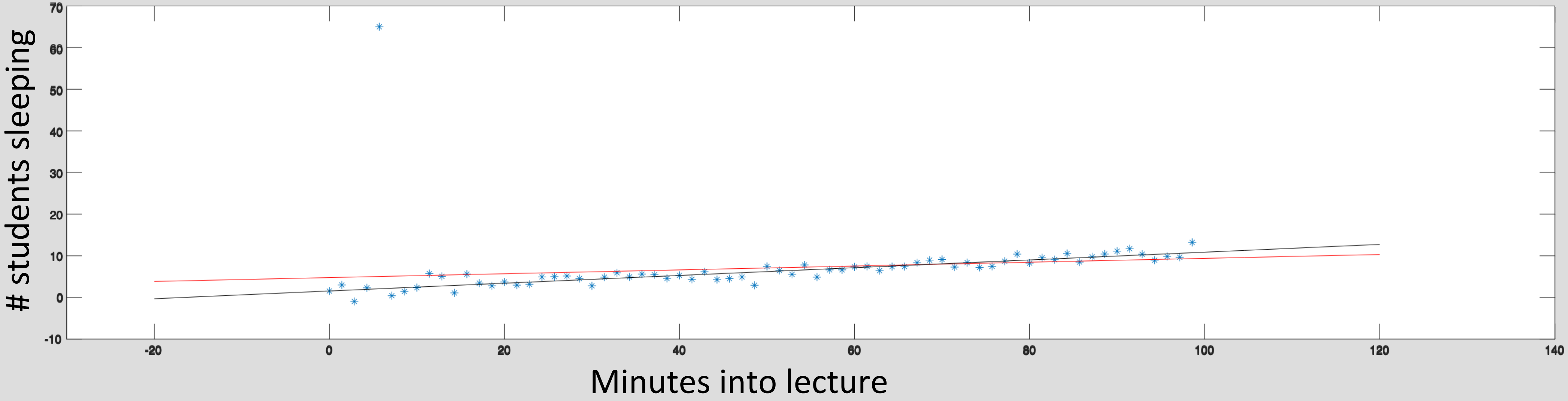
Example: Outlier

Model: $y = ax + b$



Example: Outlier

Model: $y = ax + b$

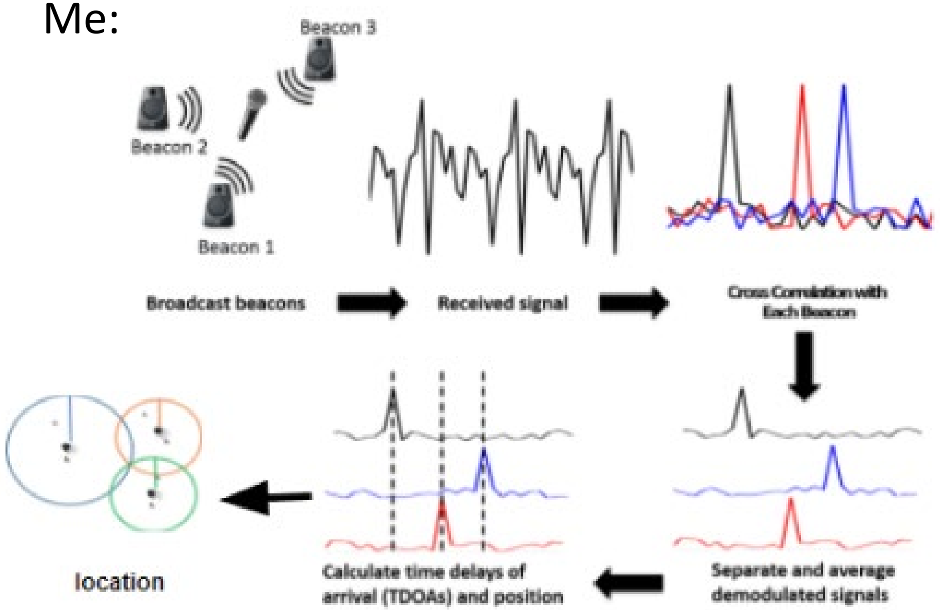


How to solve for GPS coordinates:

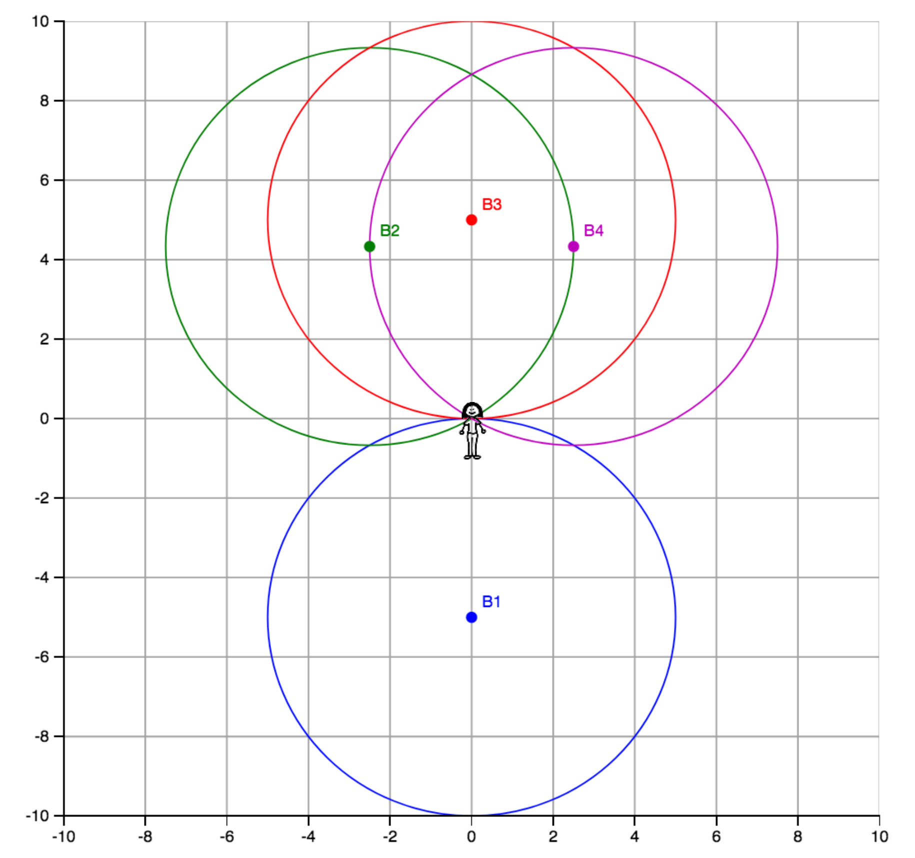
1. ***Identify*** which satellites are ‘on’
2. Find the ***delay/shift*** for each satellite
3. Use shifts to find ***distances*** to each satellite
4. ***~~Tril~~ateration*** to find my coordinates

   Multilateration

Friend: Come over!
Me: I have no idea where i am and all I have is this recording that sounds like trash
Friend: I have chocolate ☺
Me:



Example: Multi-lateration with no noise → unique solution

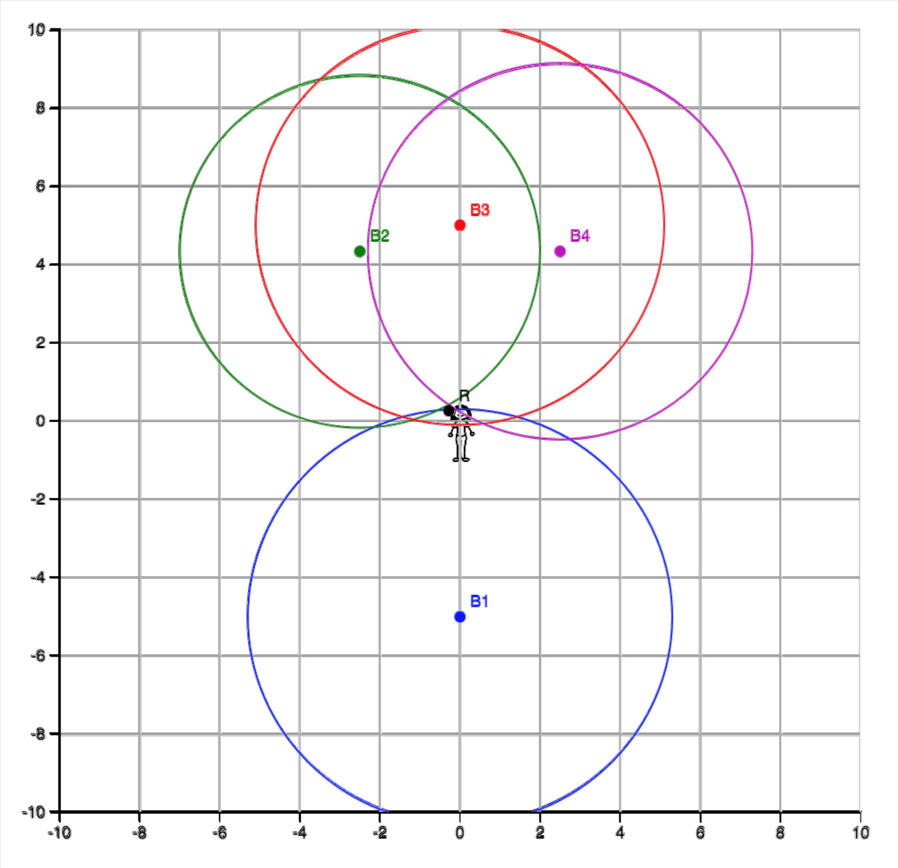


Correct measurements:

B1: 5m
B2: 5m
B3: 5m
B4: 5m

Least squares estimate:
(0,0) ✓

Example: Multi-lateration with random noise



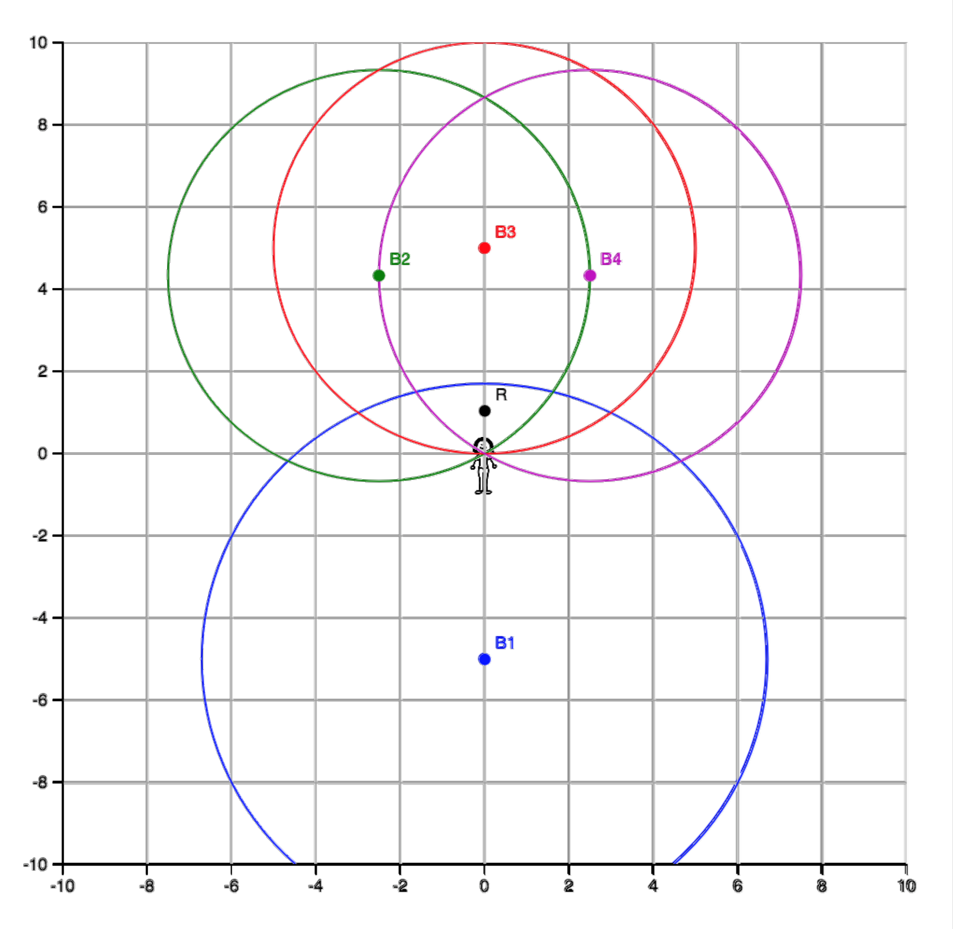
All measurements have error:

B1: 5.3m
B2: 4.5m
B3: 5.1m
B4: 4.8m

Least squares estimate:
(-0.28,0.26)

Estimate has some error, but will get smaller with more measurements (if error is random)

Example: Multi-lateration with ***not*** random noise



All measurements have error:

B1: 6.8m
B2: 5m
B3: 5m
B4: 5m

Least squares estimate:
(0,1.04)

Error is not spread evenly (random), 3 measurements were correct, 1 wrong

Example: Back to GPS

$\|\vec{x} - \cancel{\vec{a}_1}\|^2 = d_1^2 \rightarrow \|\vec{x}\|^2 = d_1^2$ ①

$\|\vec{x} - \vec{a}_2\|^2 = d_2^2 \rightarrow (\vec{x} - \vec{a}_2)^T(\vec{x} - \vec{a}_2) = \|\vec{x}\|^2 - 2\vec{a}_2^T\vec{x} + \|\vec{a}_2\|^2 = d_2^2$ ②

$\|\vec{x} - \vec{a}_3\|^2 = d_3^2 \longrightarrow \|\vec{x}\|^2 - 2\vec{a}_3^T\vec{x} + \|\vec{a}_3\|^2 = d_3^2$ ③

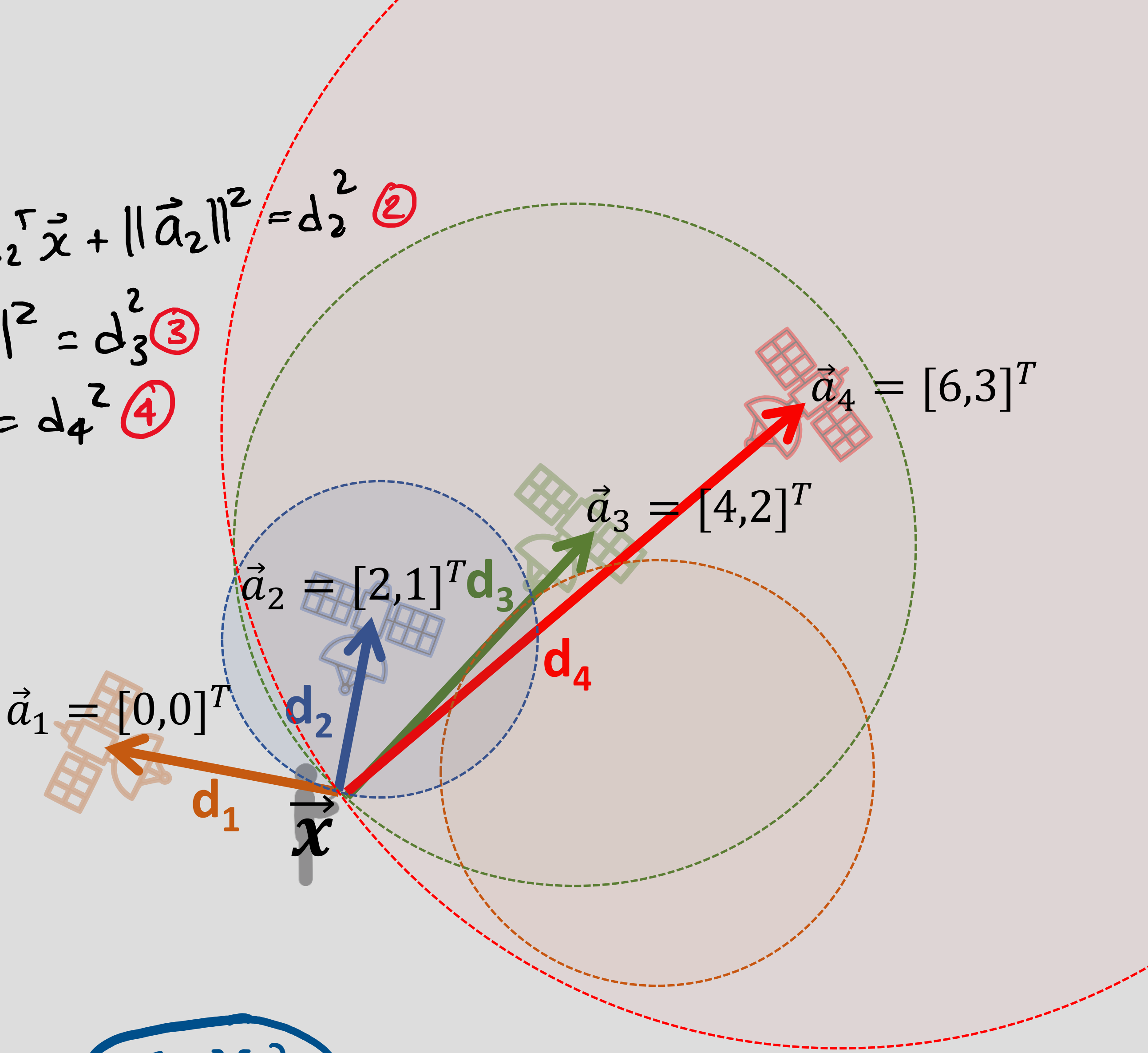
$\|\vec{x} - \vec{a}_4\|^2 = d_4^2 \longrightarrow \|\vec{x}\|^2 - 2\vec{a}_4^T\vec{x} + \|\vec{a}_4\|^2 = d_4^2$ ④

② - ① $\quad 2\vec{a}_2^T\vec{x} - \|\vec{a}_2\|^2 = d_1^2 - d_2^2$

③ - ① $\quad 2\vec{a}_3^T\vec{x} - \|\vec{a}_3\|^2 = d_1^2 - d_3^2$

④ - ① $\quad 2\vec{a}_4^T\vec{x} - \|\vec{a}_4\|^2 = d_1^2 - d_4^2$

$$\begin{bmatrix} 4 & 2 \\ 8 & 4 \\ 12 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \|\vec{a}_2\|^2 + d_1^2 - d_2^2 \\ \|\vec{a}_3\|^2 + d_1^2 - d_3^2 \\ \|\vec{a}_4\|^2 + d_1^2 - d_4^2 \end{bmatrix}$$



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

↳ A is Lin. Dependent! ☹ What does it mean for $(A^TA)^{-1}$?

$$A^TA = \begin{bmatrix}4 & 8 & 12\\ 2 & 4 & 6\end{bmatrix}\begin{bmatrix}4 & 2\\ 8 & 4\\ 12 & 6\end{bmatrix} = \begin{bmatrix}224 & 112\\ 112 & 56\end{bmatrix} \rightarrow \frac{1}{56}\begin{bmatrix}4 & 2\\ 2 & 1\end{bmatrix}$$

col 1 = 2 col 2

Lin. Dependent!

No inverse exists

$$\det\begin{bmatrix}4 & 2\\ 2 & 1\end{bmatrix} = 4-4=0$$

☹

$$\begin{bmatrix}4 & 2\\ 8 & 4\\ 12 & 6\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} = \begin{bmatrix}\|\vec{a}_2\|^2+d_1^2-d_2^2\\ \|\vec{a}_3\|^2+d_1^2-d_3^2\\ \|\vec{a}_4\|^2+d_1^2-d_4^2\end{bmatrix}$$

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

↰ A is Lin. Dependent! ☹ What does it mean for $(A^TA)^{-1}$?

Invertibility of $A^T A$

Recall: matrix is invertible if → Nullspace is 'trivial'
→ Lin. ind. cols/rows

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

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

if true, these fully overlap

Proof: Need to show two things:

① some vector $\vec{w} \in \text{Null}(A)$ is also in $\text{Null}(A^TA)$

known: $\vec{w} \in \text{Null}(A)$

$$A\vec{w} = \vec{0}$$

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

$$(A^TA)\vec{w} = \vec{0} \checkmark$$

$\vec{w}$ is also in Nullspace of A^TA

② some vector $\vec{v} \in \text{Null}(A^TA)$ is also in $\text{Null}(A)$

known: $A^TA\vec{v} = \vec{0}$

can't just divide by this... hmm

Recall: Properties of transposes

$$\underbrace{(AB)^T}_{\substack{(n\times m\ m\times k)^T \\ (n\times k)^T}} = \underbrace{B^T A^T}_{\substack{k\times m\ m\times n \\ k\times n}} \checkmark ☺$$

Trick! Consider:

$$\|A\vec{v}\|^2 = \langle A\vec{v}, A\vec{v}\rangle = (A\vec{v})^T A\vec{v}$$
$$= \vec{v}^T A^T (A\vec{v})$$
$$= \vec{v}^T \underbrace{(A^TA\vec{v})}$$
$$= \vec{v}^T(\vec{0})$$
$$= \vec{0}$$

if mag. is zero, can only be $\vec{0}$ vector

Recall: If $\|\vec{x}\| = 0$, then $\vec{x} = \vec{0}$

Proof: $\|\vec{x}\| = 0$

$\|\vec{x}\|^2 = 0$

$x_1^2 + x_2^2 + \ldots x_n^2 = \langle \vec{x}, \vec{x}\rangle = 0$

all terms $\geqslant 0$

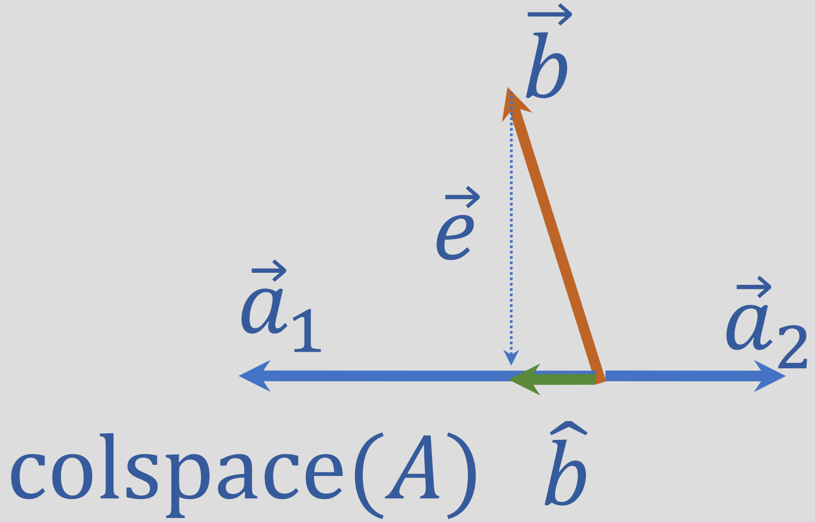
so all terms $= 0$ → so $\vec{x} = \vec{0}$! ✓☺

$$\|A\vec{v}\| = \vec{0} \rightarrow A\vec{v} = 0 \checkmark$$

in nullspace of A!

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

EECS16A: we’re do course evaluations now

If over 80% of the class fill outs the official course evaluation for this class (deadline of May 7), we will award 2 bonus points to everyone in the course (2 bonus points out of a total of 300 course points).

course-evaluations.berkeley.edu