## Least Squares Approximation

Starting point : $A\vec{x} = \vec{b}$ may be inconsistent!

*Idea*: try to get as close as possible to a solution.

- minimize the size of the difference $A\vec{x} - \vec{b}$

ie minimize $\|A\vec{x} - \vec{b}\|$

Beautiful Theorem: Starting with a linear system of equations in matrix form

(1) $$A\vec{x} = \vec{b},$$

the symmetrized system

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

always has a solution. Moreover the solution of (2) is a vector $\vec{x}$ which minimizes $\|A\vec{x} - \vec{b}\|$

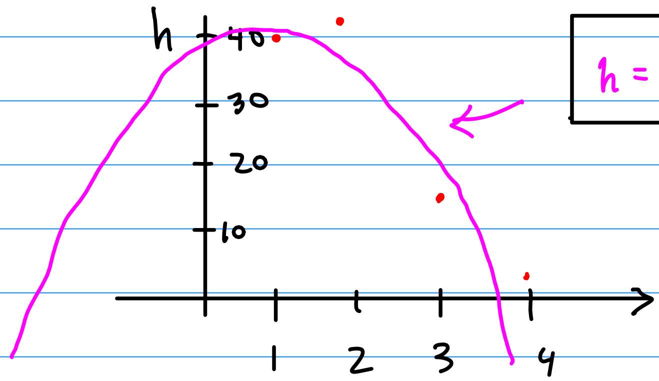
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

This has TONS of applications!

- polynomial fitting
- exponential fitting
- linear regression
- ... and lots more!

Ex: Suppose we throw a ball in the air and its height as a function of time is given by the following chart.

| $t$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $h$ | 40 | 42 | 13 | 3 |




$$\boxed{h = at^2 + bt + c}$$

I won't be able to choose $a, b, c$ to perfectly fit the data!

$$\left\{\begin{array}{l} a(1)^2 + b(1) + c = 40 \\ a(2)^2 + b(2) + c = 42 \\ a(3)^2 + b(3) + c = 13 \\ a(4)^2 + b(4) + c = 3 \end{array}\right. \rightsquigarrow \left\{\begin{array}{l} a + b + c = 40 \\ 4a + 2b + c = 42 \\ 9a + 3b + c = 13 \\ 16a + 4b + c = 3 \end{array}\right.$$

overdetermined system

$$\longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 40 \\ 42 \\ 13 \\ 3 \end{bmatrix}$$

$\leftarrow$ no solution!

Get as close as possible to a solution!

$$\begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{bmatrix}^T \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{bmatrix}^T \begin{bmatrix} 40 \\ 42 \\ 13 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 & 9 & 16 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 4 & 9 & 16 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 40 \\ 42 \\ 13 \\ 3 \end{bmatrix}$$

MATLAB: A'*A A'*b

$$\begin{bmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 373 \\ 175 \\ 98 \end{bmatrix}$$

Solve this system!

$$\begin{bmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

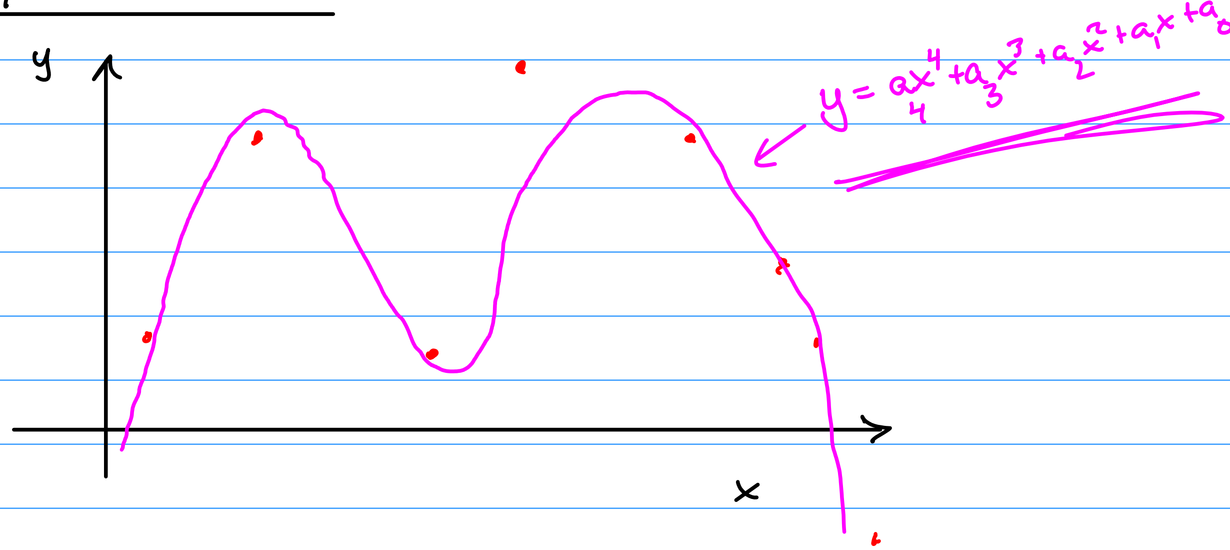
$$\boxed{\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 354 & 100 & 30 \\ 100 & 30 & 10 \\ 30 & 10 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 373 \\ 175 \\ 98 \end{bmatrix}}$$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 44.5 \end{bmatrix}$$

$$h(t) = -3t^2 + t + 44.5$$

This is **NOT** perfect! Its the polynomial which most closely approximates the data points.

## General Picture

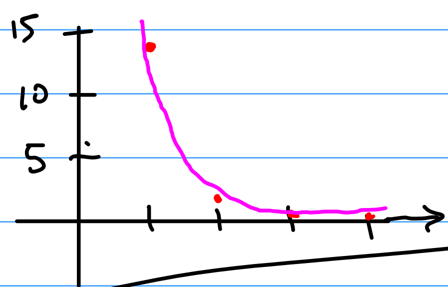



## Exponential fitting:

$$\boxed{y = ae^{bx}}$$

Ex:

| x | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| y | 13 | 2 | 0.3 | 0.05 |




$$\begin{cases} ae^{b} = 13 \\ ae^{2b} = 2 \\ ae^{3b} = 0.3 \\ ae^{4b} = 0.05 \end{cases}$$

$$\begin{cases} \ln(a) + b = \ln(13) \\ \ln(a) + 2b = \ln(2) \\ \ln(a) + 3b = \ln(0.3) \\ \ln(a) + 4b = \ln(0.05) \end{cases}$$

$c = \ln a$

$a = e^{c}$

$$\begin{cases} c + b = \ln(13) \\ c + 2b = \ln(2) \\ c + 3b = \ln(0.3) \\ c + 4b = \ln(0.05) \end{cases}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} c \\ b \end{bmatrix} = \begin{bmatrix} \ln(13) \\ \ln(2) \\ \ln(0.3) \\ \ln(0.05) \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} c \\ b \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} \ln(13) \\ \ln(2) \\ \ln(0.3) \\ \ln(0.05) \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix} \begin{bmatrix} c \\ b \end{bmatrix} = \begin{bmatrix} \ln(13) + \ln(2) + \ln(0.3) + \ln(0.05) \\ \ln(13) + 2\ln(2) + 3\ln(0.3) + 4\ln(0.05) \end{bmatrix}$$

$$\begin{bmatrix} c \\ b \end{bmatrix} = \frac{1}{120 - 100} \begin{bmatrix} 30 & -10 \\ -10 & 4 \end{bmatrix} \begin{bmatrix} \ln(13) + \ln(2) + \ln(0.3) + \ln(0.05) \\ \ln(13) + 2\ln(2) + 3\ln(0.3) + 4\ln(0.05) \end{bmatrix}$$

finish @ home!