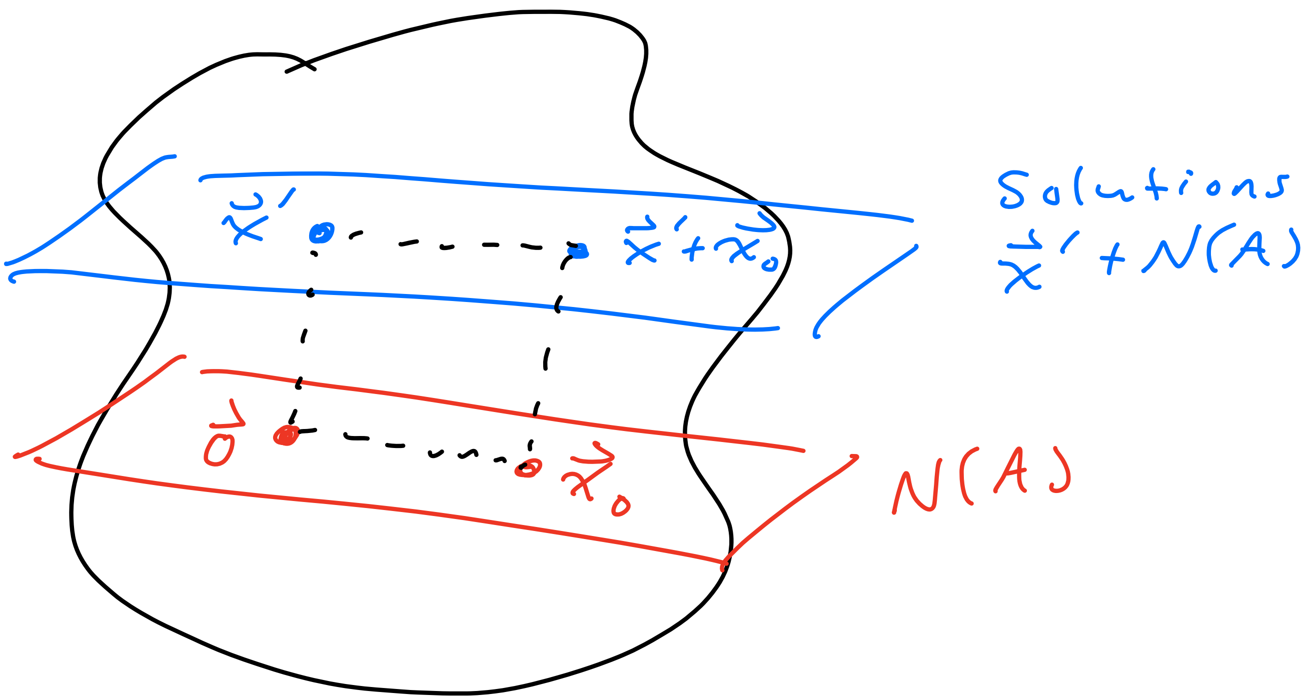
To solve $A\vec{x} = \vec{b}$.

Given one particular solution $\vec{x}'$, the general solution is

$$\vec{x} = \vec{x}' + N(A)$$

$$= \{\text{all vectors of the form } \vec{x}' + \vec{x}_0 \text{ for } \vec{x}_0 \in N(A)\}$$

Picture:



Check: If $\vec{x}_0 \in N(A)$ so $A\vec{x}_0 = \vec{0}$

then $A(\vec{x}' + \vec{x}_0) = A\vec{x}' + A\vec{x}_0$

$$= \vec{b} + \vec{0} = \vec{b}.$$

Example.

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

$$\begin{pmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

$$\xrightarrow[\text{RREF}]{} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{cases} x + 2z = -1 \\ y + 2z = 1 \\ 0 = 0 \end{cases}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1-2z \\ 1-2z \\ z \end{pmatrix}$$

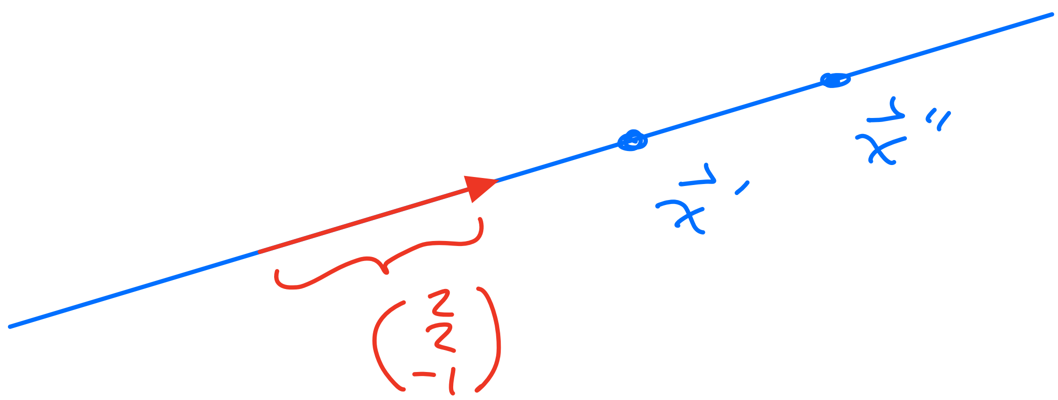
$$= \underbrace{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}}_{\text{some particular } \vec{x}'} + \underbrace{t\begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}}_{N(A)}$$

Another solution:

$$\vec{x}'' = \begin{pmatrix}-1\\1\\0\end{pmatrix} + \begin{pmatrix}2\\2\\-1\end{pmatrix} = \begin{pmatrix}1\\3\\-1\end{pmatrix}$$

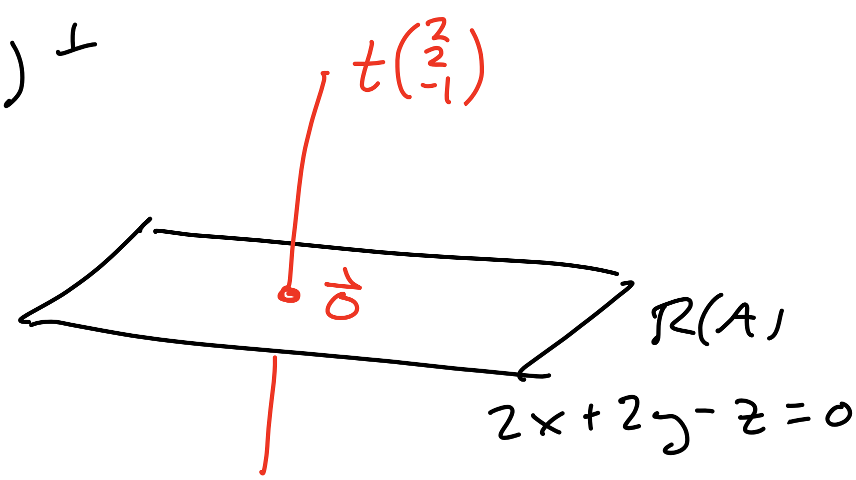
We can also express the solution

$$\begin{pmatrix}x\\y\\z\end{pmatrix} = \underbrace{\begin{pmatrix}1\\3\\-1\end{pmatrix}}_{\vec{x}''} + \underbrace{t\begin{pmatrix}2\\2\\-1\end{pmatrix}}_{N(A)}$$



Where is the row space?

$R(A) = N(A)^{\perp}$



Recall $C(A) = \{ A\vec{x} : \vec{x} \in \mathbb{R}^n \}$

$= \{ \text{linear combos of cols of } A \}$.

So $A\vec{x} = \vec{b}$ has a solution $\vec{x}$ (possibly non-unique)

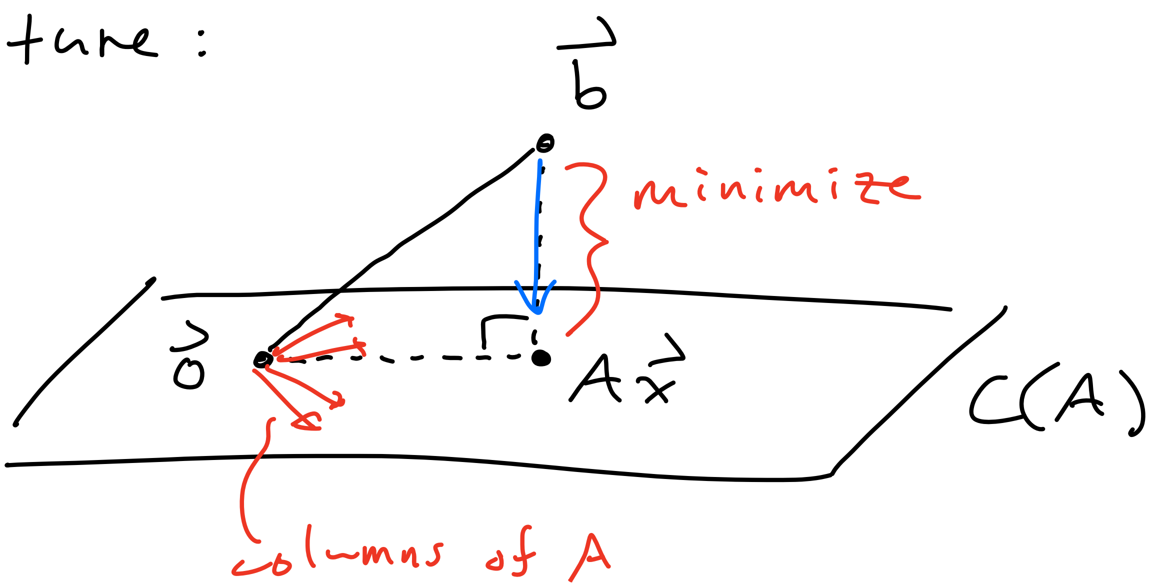
$\iff \vec{b} \in C(A)$.

What if $\vec{b} \notin C(A)$?

WANT : $\vec{x}$ such that

$\| A\vec{x} - \vec{b} \|$ is minimized.

Picture :



Geometry: Distance minimized when vector $A\vec{x} - \vec{b}$ is $\perp$ to $C(A)$, i.e., $A\vec{x} - \vec{b}$ is $\perp$ to every col of A.

TRICK (Memorize!):

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

$\perp$ to every row of A^T
i.e. to every col of A.

[Details: $A = \begin{pmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{pmatrix}$

$$A^T \vec{v} = \begin{pmatrix} - & \vec{a}_1^T & - \\ & \vdots & \\ - & \vec{a}_m^T & - \end{pmatrix} \vec{v}$$

$$= \begin{pmatrix} \vec{a}_1^T \vec{v} \\ \vdots \\ \vec{a}_m^T \vec{v} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Means $\vec{a}_i^T \vec{v} = 0 \quad \forall i$

i.e. $\vec{v} \perp \vec{a}_i$ for all i.

Summary:

$$A^T \vec{v} = \vec{0} \iff \vec{v} \perp \text{ to every col of } A.$$

Memorize

$A\vec{x} = \vec{b}$ NO SOLUTION

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

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

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

The normal equation.

This always has a solution!

The matrix $A^T A$.

- Symmetric & Square
- $N(\underbrace{A^TA}_{n\times n}) = N(\underset{\substack{\uparrow \\ m\times n}}{A})$ [HW 3.1a]
- If A has independent columns, so $\dim N(A) = 0$, then $\dim N(A^TA) = 0$, hence A^TA has ind. columns. Since A^TA is SQUARE

$\rightsquigarrow$ $(A^TA)^{-1}$ exists.

Summary:

$$\underset{r=n}{A \text{ ind. cols}} \Longrightarrow (A^TA)^{-1} \text{ exists.}$$

If A has ind cols, then

there always exists a UNIQUE least squares solution:

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

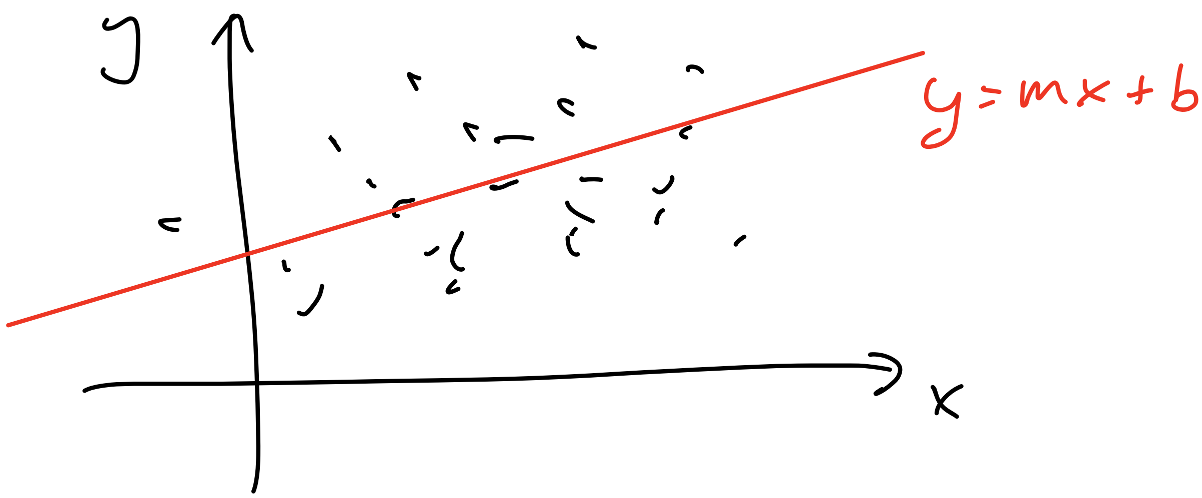
$$\rightsquigarrow \quad A^T A \vec{x} = A^T \vec{b}$$

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

[If A has non-ind cols, the least squares sol is not unique.]

Most common example:

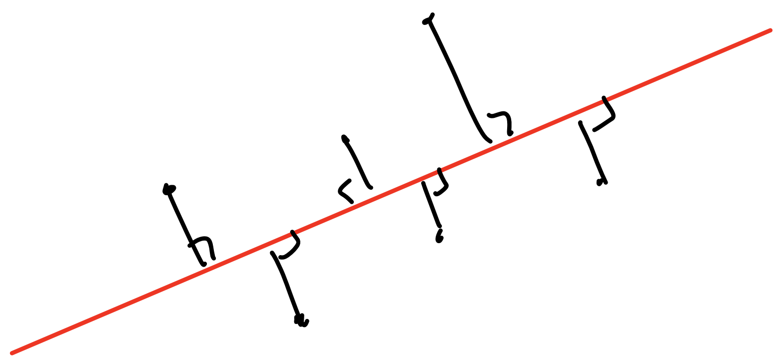
Data points $(x_1, y_1), \ldots, (x_n, y_n)$.



Find the best fit line $y=mx+b$.
i.e. Find m & b.

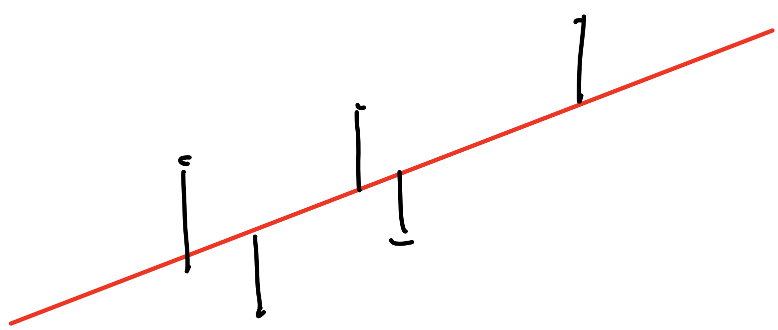
Solution: Depends what you mean by "best".

Most obvious: Minimize distances:



Hard non-linear problem. Later we'll solve it with SVD.

Easier: Minimize vert. distances:



This is what $A^TA\vec{x} = A^T\vec{b}$ does.

Method: Pretend all of the data points are on the line, so

$$\begin{cases} y_1 = mx_1 + b \\ y_2 = mx_2 + b \\ \vdots \\ y_n = mx_n + b. \end{cases}$$

system of n linear eqns in 2 unknowns m & b.

Matrices:

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$A \begin{pmatrix} m \\ b \end{pmatrix} = \vec{y}$$

No solution, so look at normal eq.

$$A^T A \begin{pmatrix} m \\ b \end{pmatrix} = A^T \vec{y}$$

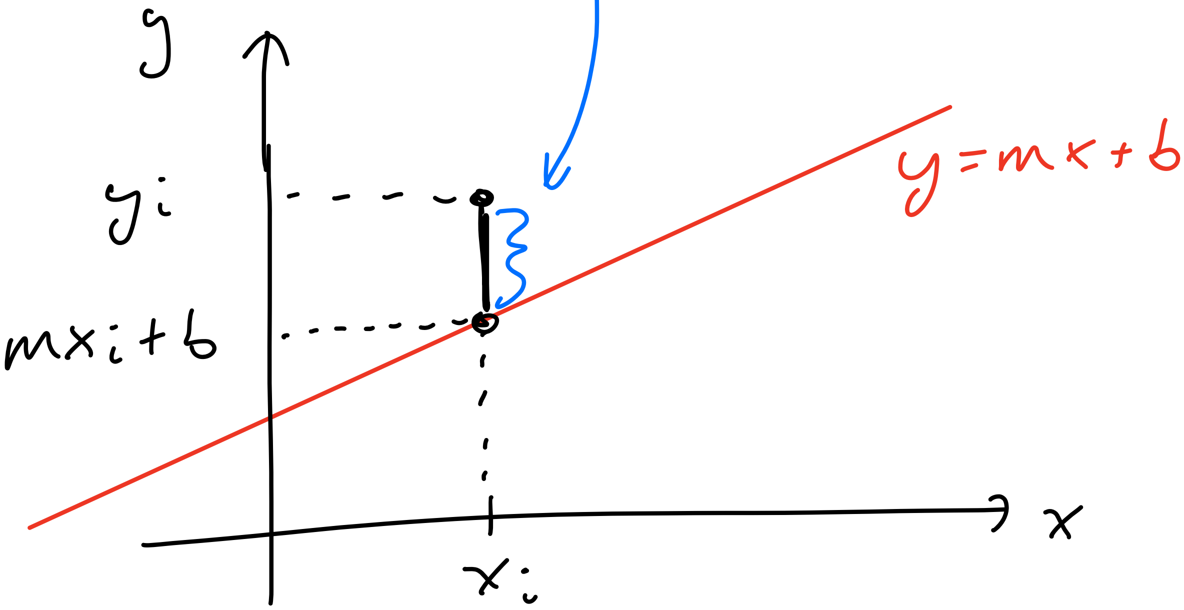
Then solve.

$$\begin{pmatrix} m \\ b \end{pmatrix} = (A^T A)^{-1} A^T \vec{y}.$$

What is minimized?

$$\left\| A \begin{pmatrix} m \\ b \end{pmatrix} - \vec{y} \right\|^2$$

$$= \sum \Big(\underbrace{(mx_i + b) - y_i} \Big)^2$$



Explicit:

$$A^T A \begin{pmatrix} m \\ b \end{pmatrix} = A^T \vec{y}$$

$$\begin{pmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$

2 linear eqs in 2 unknows.
Usually derived from Calculus in a statistics course.

General least squares has many more applications.

HW 3.4: Find distance between skew lines in $\mathbb{R}^3$