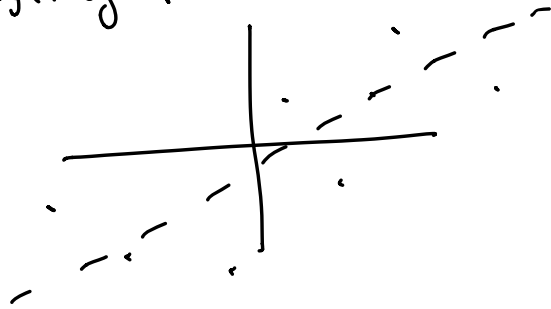


Given a set  $\{(x_i, y_i)\}_{i=1}^N$  of points.

Before: thought of this as  $N$  pts in  $\mathbb{R}^2$



Now:  $\vec{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$   $\vec{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$   $\vec{E} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$\vec{X}, \vec{Y}, \vec{E} \in \mathbb{R}^N$

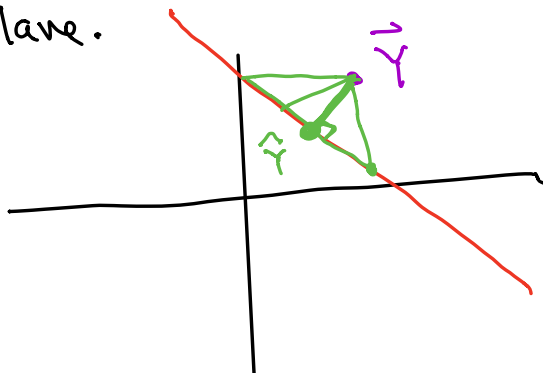
NOTICE: if  $y_i = mx_i + b$  for all  $i$  then

$$\vec{Y} = m\vec{X} + b\vec{E}.$$

in other words the 3 vectors are linearly dependent.

But that's NOT TRUE.

$\vec{X}, \vec{E}$  span a plane (a 2-dim'l space) in  $\mathbb{R}^N$ . and  $\vec{Y}$  doesn't belong to that plane.



$\vec{Y} = m\vec{X} + b\vec{E}$   
 for some  $m, b$   
 $\vec{Y} = \begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_N \end{pmatrix}$  are predicted  $y$ -values  
 plane spanned by  
 $\vec{X}, \vec{E} = \{m\vec{X} + b\vec{E}\}$

$$\hat{Y}(m,b) = m\vec{X} + b\vec{E}$$

$$\|\hat{Y} - \vec{Y}\|^2 = \|m\vec{X} + b\vec{E} - \vec{Y}\|^2 \quad \leftarrow \text{minimize this}$$

$$\|A\|^2 = A \cdot A$$

$$D = \|\hat{Y} - \vec{Y}\|^2 = (m\vec{X} + b\vec{E} - \vec{Y}) \cdot (m\vec{X} + b\vec{E} - \vec{Y})$$

$$\frac{\partial D}{\partial m} = 2\vec{X} \cdot (m\vec{X} + b\vec{E} - \vec{Y})$$

$$\frac{\partial D}{\partial b} = 2\vec{E} \cdot (m\vec{X} + b\vec{E} - \vec{Y})$$

$$\text{Set } \frac{\partial D}{\partial m} = 0 \quad \text{and} \quad \frac{\partial D}{\partial b} = 0$$

$$\vec{X} \cdot (m\vec{X} + b\vec{E} - \vec{Y}) = 0$$

$$\vec{E} \cdot (m\vec{X} + b\vec{E} - \vec{Y}) = 0$$

$A(m,b)$  is  $\perp$  to both  $\vec{X}$  and  $\vec{E}$   
and so is  $\perp$  to the plane they span.

$$m(\vec{X} \cdot \vec{X}) + b(\vec{X} \cdot \vec{E}) = \vec{X} \cdot \vec{Y}$$

$$m(\vec{E} \cdot \vec{X}) + b(\vec{E} \cdot \vec{E}) = \vec{E} \cdot \vec{Y}$$

SAME 2 equations we obtained by minimizing the MSE.

$$\vec{X} \cdot \vec{X} = \sum x_i^2$$

$$\vec{X} \cdot \vec{E} = \sum x_i$$

$$\vec{X} \cdot \vec{Y} = \sum x_i y_i$$

$$\vec{Y} \cdot \vec{E} = \sum y_i$$

$$\vec{E} \cdot \vec{E} = N$$

