

# Least Squares Revisit

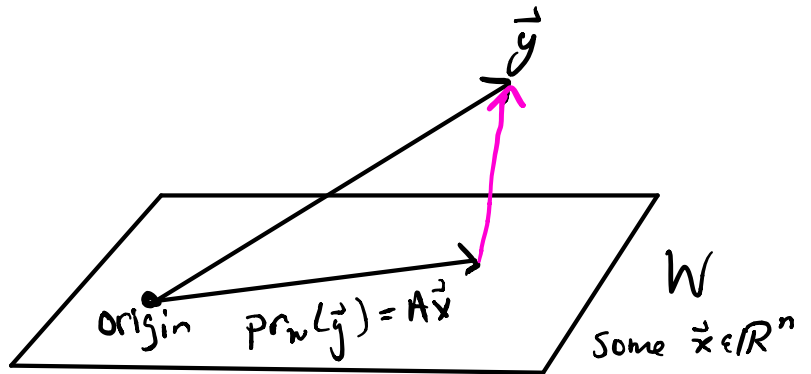
Situation:  $A$  is  $m \times n$  matrix over  $\mathbb{R}$  of rank  $n$

$$A: \mathbb{R}^n \rightarrow \mathbb{R}^m \quad (\text{w/ standard inner product})$$

Let  $W = \text{im } A \subseteq V = \mathbb{R}^m$   
then  $\dim W = n$

Have  $W^\perp$  s.t.  $V = W \oplus W^\perp$   
and  $\dim W^\perp = m - n$

Have  $\text{pr}_W: V \rightarrow V$



① Find formula for  $\text{pr}_W$

② In least squares: find  $\vec{x}$  s.t.  $\|A\vec{x} - \vec{y}\|$  is minimal  
over all  $\vec{x} \in \mathbb{R}^n$

In both cases, use  $(A\vec{x} - \vec{y}) \perp W$  i.e.  $A\vec{x} - \vec{y} \in W^\perp$

$$A\vec{x} - \vec{y} \perp W \iff A\vec{x} - \vec{y} \in W^\perp = (\text{im } A)^\perp$$

$$\iff A\vec{x} - \vec{y} \in \ker(A^T)$$

$$\iff A^T(A\vec{x}) = A^T\vec{y} \quad (\text{"normal equations"})$$

$$\iff \begin{matrix} (A^T A) & \vec{x} & = & A^T \vec{y} \\ n \times n & n \times 1 & & n \times m \quad m \times 1 \end{matrix}$$

Note:  $A^T A$  is an invertible  $n \times n$  matrix

$$\iff \vec{x} = (A^T A)^{-1} A^T \vec{y} \quad \text{wowza!}$$

So then

$$\text{pr}_w(\vec{y}) = A \vec{x} \quad \text{for this } \vec{x}$$

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

$$\therefore \text{pr}_w = \underset{m \times n}{A} \underset{n \times n}{(A^T A)^{-1}} \underset{n \times m}{A^T}$$

Question: Can't we simplify this?

$$\begin{aligned} \text{i.e. } A(A^T A)^{-1} A^T &= A A^{-1} A^{-T} A^T \\ &= I I = I \quad ? \end{aligned}$$

PROBLEM!  $A$  is NOT square

If  $A$  square,  $w$  is the same space as what you're projecting from, and that IS the identity map.

Example: Line of Best Fit

$$\begin{aligned} \text{data: } & (\hat{x}_1, \hat{y}_1) \\ & \vdots \\ & (\hat{x}_m, \hat{y}_m) \end{aligned}$$

specific example

$$\begin{bmatrix} (20, 2.7) \\ (40, 5) \\ (60, 7.5) \\ (70, 8.6) \end{bmatrix}$$

$$\text{Want } \vec{x} = \begin{pmatrix} a \\ b \end{pmatrix} \text{ s.t. } \sum_{i=1}^m ((a + b\hat{x}_i) - \hat{y}_i)^2 \text{ is minimal}$$

i.e

$$A \hat{\vec{x}} = \vec{y}$$
$$\begin{matrix} m \times 2 & \begin{bmatrix} 1 & \hat{x}_1 \\ 1 & \hat{x}_2 \\ \vdots & \vdots \\ 1 & \hat{x}_m \end{bmatrix} & \begin{bmatrix} a \\ b \end{bmatrix} & = & \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_m \end{bmatrix} & = & \vec{\hat{y}} \end{matrix}$$

Solve

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

$(2 \times 2) \quad 2 \times 1 \quad 2 \times m \quad m \times 1$

$$A^T A = \begin{bmatrix} m & \sum \hat{x}_i \\ \sum \hat{x}_i & \sum \hat{x}_i^2 \end{bmatrix}$$

$$A^T \vec{\hat{y}} = \begin{bmatrix} \sum \hat{y}_i \\ \sum \hat{x}_i \hat{y}_i \end{bmatrix}$$

In our original example

$$A = \begin{bmatrix} 1 & 20 \\ 1 & 40 \\ 1 & 60 \\ 1 & 70 \end{bmatrix}, \quad \vec{\hat{y}} = \begin{bmatrix} 2.7 \\ 5 \\ 7.5 \\ 8.6 \end{bmatrix}$$

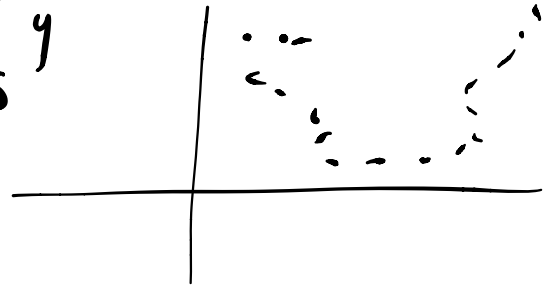
$$A^T A \hat{\vec{x}} = \begin{bmatrix} 4 & 190 \\ 190 & 10500 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 238 \\ 1306 \end{bmatrix} = A^T \vec{\hat{y}}$$

$$\text{get } \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.12 \end{bmatrix} \therefore y = 0.3 + 0.12x$$

Example: Suppose your data is  
fit the best function

$$y = a + bx + cx^2$$

choice of  $y$   
is YOURS



$$A = \begin{bmatrix} 1 & \hat{x}_1 & \hat{x}_1^2 \\ \vdots & \vdots & \vdots \\ 1 & \hat{x}_m & \hat{x}_m^2 \end{bmatrix}$$

→ solve exactly to get  
 $\vec{x} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$