Least Squares Revisit

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

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

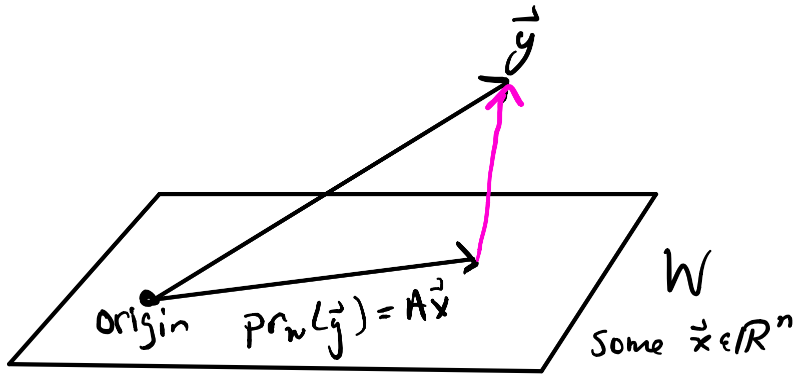
Let $W = \text{im } A \subseteq V = \mathbb{R}^m$

then $\dim W = n$

Have $W^\perp$ s.t. $V = W \oplus W^\perp$

and $\dim W^\perp = m - n$

Have $pr_W : V \to V$



① Find formula for 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 $(Ax - y) \perp W$ i.e. $Ax - y \in W^\perp$

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

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

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

$$\iff \underset{n\times n}{(A^TA)}\ \underset{n\times 1}{\vec{x}} = \underset{n\times m}{A^T}\ \underset{m\times 1}{\vec{y}}$$

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

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

wowza!

So then

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

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

$$\therefore \quad \mathrm{pr}_W = \underset{m\times n}{A}\,\underset{n\times n}{(A^TA)^{-1}}\,\underset{n\times m}{A^T}$$

Question: Can't we simplify this?

i.e $A(A^TA)^{-1}A^T = AA^{-1}A^{-T}A^T$

$= II = I$?

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

data: $(\hat{x}_1, \hat{y}_1)$

$\vdots$

$(\hat{x}_m, \hat{y}_m)$

specific example

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

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

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

Solve

$$\underset{(2\times 2)}{A^TA}\, \underset{2\times 1}{\vec{x}} = \underset{2\times m}{A^T}\, \underset{m\times 1}{\vec{y}}$$

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

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

In our original example

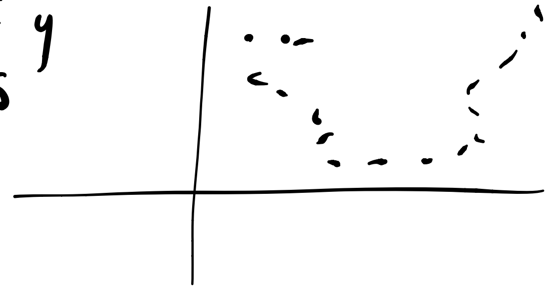
$$A = \begin{bmatrix} 1 & 20 \\ 1 & 40 \\ 1 & 60 \\ 1 & 70 \end{bmatrix}, \quad \hat{y} = \begin{bmatrix} 2.7 \\ 5 \\ 7.5 \\ 8.6 \end{bmatrix}$$

$$A^TA\vec{x} = \begin{bmatrix} 4 & 190 \\ 140 & 10500 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 238 \\ 1306 \end{bmatrix} = A^T\vec{y}$$

get

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

Example: Suppose your data is



fit the best function

$y = a + bx + cx^2$ choice of y is YOURS

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

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