Least Squares Approximation

Recall: suppose given vectors $\vec{v}_1, \ldots, \vec{v}_n$ in $\mathbb{R}^m$ which are independent

$$\underset{m\times n}{A} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

subspace $V \subset \mathbb{R}^m$ spanned by $\vec{v}_1, \ldots, \vec{v}_n$ is $C(A)$.

The projection matrix associated to A: $\boxed{P = A(A^TA)^{-1}A^T}$ $(m\times m)$

If $\vec{b}$ is in $\mathbb{R}^m$, $P\vec{b}$ is projection of $\vec{b}$ onto $V = C(A)$.

Can check: $P^2 = P$ (exercise).

What if we apply the Rank-Nullity Theorem to $P: \mathbb{R}^m \to \mathbb{R}^m$?

$$\dim(\mathrm{im}(P)) + \dim(N(P)) = \dim(\mathbb{R}^m)$$

image $\mathrm{im}(P)$ = column space $C(P)$ = {possible outputs of P}

= V (since an output is a projection onto V!)

nullspace $N(P) = \{ \vec{v} \text{ in } \mathbb{R}^m \mid \text{projection of } \vec{v} \text{ onto } V \text{ is } \vec{0} \} = V^\perp$

$$\longrightarrow \boxed{\dim(V) + \dim(V^\perp) = m.}$$

<u>Example</u> Let $\vec{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\vec{a}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ be a basis for a plane $V \subset \mathbb{R}^3$.

Find the associated projection matrix P.

$$A = \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 1 \\ 1 & 2 \end{bmatrix}$$

$$(A^TA)^{-1} = \frac{1}{(6)(2)-(1)(1)} \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix}$$

$$\rightarrow P = A(A^TA)^{-1}A^T = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix} \left(\frac{1}{11} \begin{bmatrix} 2 & -1 \\ -1 & 6 \end{bmatrix} \right) \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

$$= \frac{1}{11} \begin{bmatrix} 3 & -7 \\ 3 & 4 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 0 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 10 & -1 & 3 \\ -1 & 10 & 3 \\ 3 & 3 & 2 \end{bmatrix}$$

Back to the general setup.

$\underset{m \times n}{A}$ Suppose we want to solve $A\vec{x} = \vec{b}$ ($\vec{x}$ unknown)

Recall: $A\vec{x} = \vec{b}$ has a solution $\Leftrightarrow$ $\vec{b}$ is in $C(A)$, column space.

What if $\vec{b}$ is not in $C(A)$?

Then $A\vec{x}=\vec{b}$ has no solution, but we can look for the $\vec{x}$ which come "closest" to solving $A\vec{x}=\vec{b}$.

Basic idea: project everything to $C(A)$.

let P = projection matrix associated to A $= A(A^TA)^{-1}A^T$

(Here we should retain the assumption that columns of A are indep.)

Then projecting everything to $C(A)$ means applying P:

$$A\vec{x}=\vec{b} \rightarrow PA\vec{x}=P\vec{b}$$

$$\rightarrow \underbrace{A(A^TA)^{-1}}\,\underbrace{A^TA}_{}\,\vec{x} = A(A^TA)^{-1}A^T\vec{b}$$

cancel

$$\rightarrow A\vec{x} = A(A^TA)^{-1}A^T\vec{b}$$

multiply by A^T

$$\rightarrow A^TA\vec{x} = \cancel{(A^TA)(A^TA)^{-1}}A^T\vec{b}$$

$$\rightarrow A^TA\vec{x} = A^T\vec{b}$$

$$\rightarrow \vec{x} = (A^TA)^{-1}A^T\vec{b}.$$

We have obtained: $\boxed{\vec{x} = (A^TA)^{-1}A^T\vec{b}}$

This is the "least squares approximation" to $A\vec{x}=\vec{b}$.

It is called this because it's the vector $\vec{x}$ which minimizes the "error" quantity

$$\|A\vec{x}-\vec{b}\|^2 = \sum_{i=1}^{m}\left(b_i - \sum_{j=1}^{n} a_{ij}x_j\right)^2$$

where a_{ij} is the (i,j)-entry of A.

Example Find the line that best approximates the data

$(1,1)$ $(2,-1)$ $(-1,1)$

We're looking for $y = ax+b$ (a, b the unknowns) so that the line best approximates the data.

$$\left.\begin{array}{ll} (1,1) \text{ on the line:} & 1 = a+b \\ (2,-1) \text{ on the line:} & -1 = 2a+b \\ (-1,1) \text{ on the line:} & 1 = -a+b \end{array}\right\} \text{equiv. to: } \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}}_{\vec{b}}$$

We cannot actually solve $A\vec{x} = \vec{b}$ so we look for the least squares approximation, which is

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

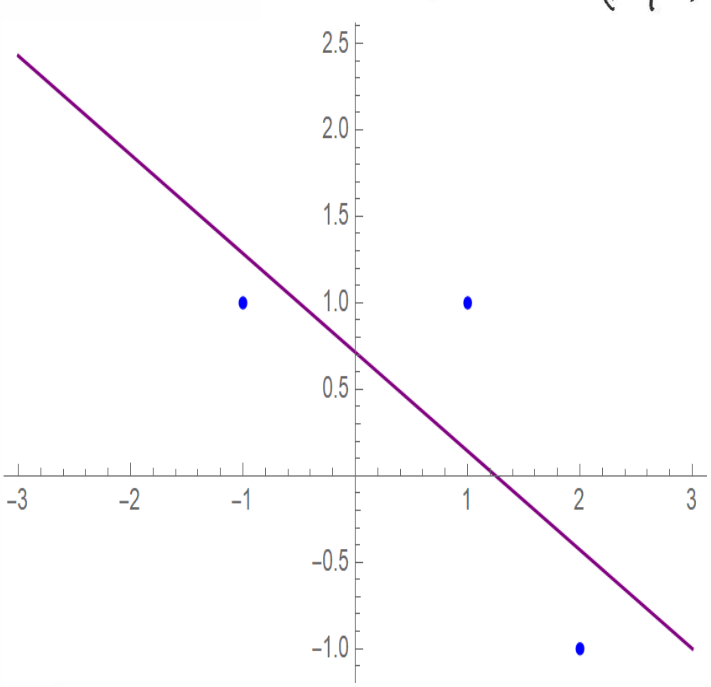
$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix}$ $\quad A^T = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ $\quad AA^T = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$

$$(A^T A)^{-1} = \frac{1}{(6)(3)-(2)(2)} \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix}$$

So $\vec{x} = \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{14} \begin{bmatrix} 3 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$$= \frac{1}{14} \begin{bmatrix} -8 \\ 10 \end{bmatrix} = \frac{1}{7} \begin{bmatrix} -4 \\ 5 \end{bmatrix}$$

Thus $y = \left(\frac{-4}{7}\right) x + \left(\frac{5}{7}\right)$ is the line we want.

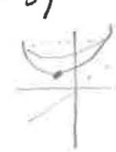


Other viewpoint:

$$\|\vec{b} - A\vec{x}\| = \left\| \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \right\|^2$$

$$= (1-a-b)^2 + (-1-2a-b)^2 + (1+a+b)^2$$

This is the "error" term.



It's minimized at $a = \frac{-4}{7}$, $b = 5/7$