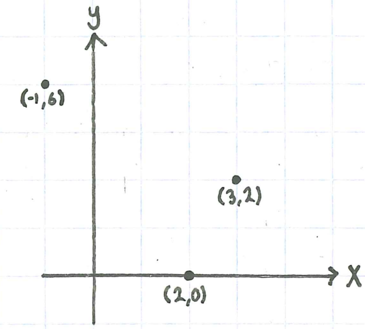
ESC103 - ENGINEERING MATHEMATICS and COMPUTATION

Solving the Least Squares Problem

It is often the case that $A\vec{x}=\vec{b}$ has no solution. This most commonly arises with tall and thin systems ($m>n$). In such cases, we are often still interested in finding a solution for $\vec{x}$ such that $A\vec{x}\approx\vec{b}$

Example:

3 Data Points $\{(-1,6),(2,0),(3,2)\}$



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

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}\underbrace{\begin{bmatrix} a \\ b \\ c \end{bmatrix}}_{\text{unknowns}} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \qquad A\vec{x}=\vec{b}$$

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 2 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 6 \\ 1 & 2 & 4 & 0 \\ 1 & 3 & 9 & 2 \end{array}\right] \xrightarrow{GE} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Square ($m=n=3$)

$r=3 \therefore A\vec{x}=\vec{b}$ has 1 solution

$a=2$

$b=-3$

$c=1$

$$y = 2-3x+x^2$$

What about $y=a+bx$?

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}\underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\text{unknowns}} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

Tall and Thin ($m=3>n=2$)

$r=2$

$A\vec{x}=\vec{b}$ has 0 solutions

$$\left[\begin{array}{cc|c} 1 & -1 & 6 \\ 1 & 2 & 0 \\ 1 & 3 & 2 \end{array}\right] \xrightarrow{GE} \left[\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right]$$

Transpose of a Matrix

If Matrix A is $m \times n$, the transpose of Matrix A, denoted A^T, is the $n \times m$ Matrix where the rows of A^T are the columns of Matrix A written in the same order.

Example:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 5 & 0 \end{bmatrix} \quad (2 \times 3)$$

$$A^T = \begin{bmatrix} 1 & 0 \\ 2 & 5 \\ -1 & 6 \end{bmatrix} \quad (3 \times 2)$$

Properties:

$P_1 : (A^T)^T = A$

$P_2 : (cA)^T = cA^T$, c is a scalar

$P_3 : (A+B)^T = A^T + B^T$

$P_4 : (AB)^T = B^T A^T$

Formulating the Least Squares Problem

$$\underset{(m \times n)}{A}\underset{(n \times 1)}{\vec{x}} = \underset{(m \times 1)}{\vec{b}}$$

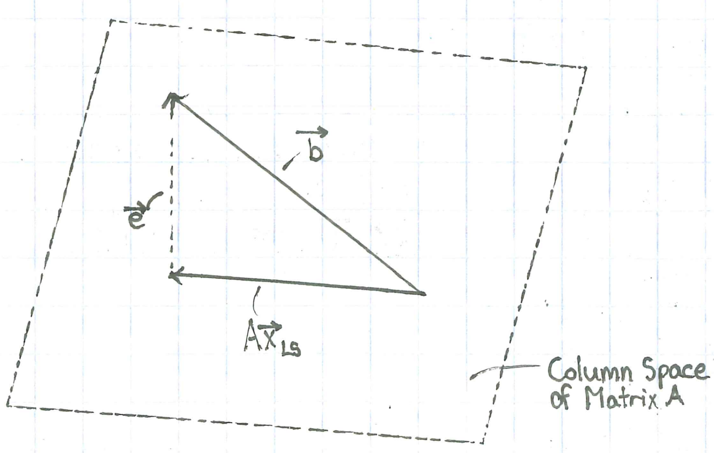
$$A = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

$$A\vec{x} = x_1 \begin{bmatrix} \uparrow \\ \vec{a}_1 \\ \downarrow \end{bmatrix} + x_2 \begin{bmatrix} \uparrow \\ \vec{a}_2 \\ \downarrow \end{bmatrix} + \cdots + x_n \begin{bmatrix} \uparrow \\ \vec{a}_n \\ \downarrow \end{bmatrix}$$

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

There is a solution to $A\vec{x} = \vec{b}$ if $\vec{b}$ lies in the column space of Matrix A.
There is no solution to $A\vec{x} = \vec{b}$ if $\vec{b}$ does not lie in the column space of Matrix A.

To solve the latter problem, we may still want to find a vector X such that $A\vec{x}$ is closest to $\vec{b}$.



Let's define an error vector $\vec{e}$ where

$$\vec{e} = \vec{b} - A\vec{x}_{LS} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} \quad (m \times 1)$$

When we project $\vec{b}$ onto the column space of Matrix A, the error vector will be orthogonal to every column vector of Matrix A.

$$\therefore \begin{bmatrix} \uparrow \\ \vec{a}_1 \\ \downarrow \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} = 0 \qquad \begin{bmatrix} \uparrow \\ \vec{a}_n \\ \downarrow \end{bmatrix} \cdot \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} = 0$$

$$\underset{(1 \times m)}{\begin{bmatrix} \longleftarrow \vec{a}_1 \longrightarrow \end{bmatrix}} \underset{(m \times 1)}{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}} = 0 \qquad \begin{bmatrix} \longleftarrow \vec{a}_n \longrightarrow \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} = 0$$

$$\underset{(n \times m)}{\begin{bmatrix} \longleftarrow \vec{a}_1 \longrightarrow \\ \longleftarrow \vec{a}_2 \longrightarrow \\ \vdots \\ \longleftarrow \vec{a}_n \longrightarrow \end{bmatrix}} \underset{(m \times 1)}{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}} = \underset{(n \times 1)}{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}} = \vec{0}$$

$$A^T\vec{e} = A^T(\vec{b} - A\vec{x}_{LS}) = \vec{0}$$

$$\underbrace{A^TA}_{(n \times n)} \underbrace{\vec{x}_{LS}}_{(n \times 1)} = \underbrace{A^T\vec{b}}_{(n \times 1)}$$

Let $A^* = A^TA \qquad \vec{b}^* = A^T\vec{b}$

$A^*\vec{x}_{LS} = \vec{b}^*$ (Just Another $A\vec{x} = \vec{b}$ Type Problem!)

Matrix A^* is a square system $(n \times n)$

If $\text{rank}(A^*) = n$, then there is a unique solution for x_{LS}

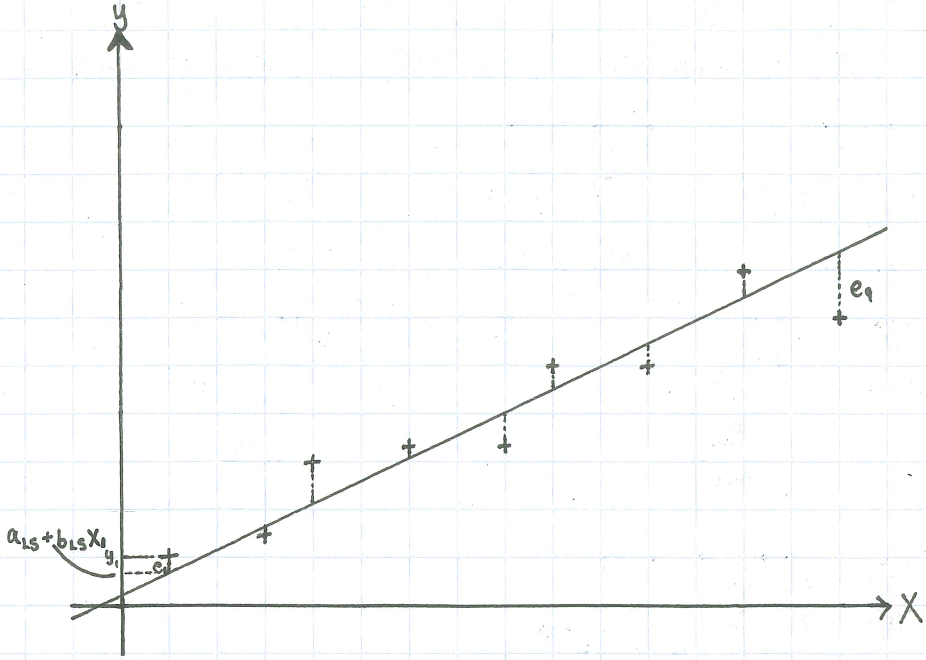
Why do we call it least squares?

When we project $\vec{b}$ onto the column space of Matrix A, we are finding the shortest vector $\vec{e}$

$$\|\vec{e}\| = \sqrt{e_1^2 + e_2^2 + \cdots + e_m^2}$$

it turns out that in this case minimizing $\|\vec{e}\|$ is equivalent to minimizing $\|\vec{e}\|^2$

$$\|\vec{e}\|^2 = e_1^2 + e_2^2 + \cdots + e_m^2$$



Why is that e?

Recall: $\begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$

$$\vec{e} = \vec{b} - A\vec{X}_{LS}$$

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} - \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \end{bmatrix} \begin{bmatrix} a_{LS} \\ b_{LS} \end{bmatrix}$$

$$e_1 = y_1 - (a_{LS} + b_{LS} X_1)$$
$$e_2 = y_2 - (a_{LS} + b_{LS} X_2)$$
$$e_3 = y_3 - (a_{LS} + b_{LS} X_3)$$