30 Linear Regression

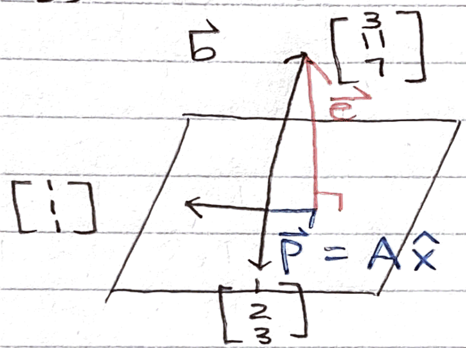
$SSE = e_1^2 + e_2^2 + e_3^2$ where e is the points' residual

$y = C + Dt$ $(1,3)$ $(2,11)$ $(3,7)$

1) $3 = C + D(1)$
2) $11 = C + D(2)$
3) $7 = C + D(3)$

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 3 \\ 11 \\ 7 \end{bmatrix}}_{\vec{b}}$$

no solution



$\vec{p} \equiv$ projection of $\vec{b}$ onto the column space (A)

$\hat{x} \equiv$ the linear combination that gets a vector $\vec{p}$ that lies right underneath $\vec{b}$, $\vec{e}$ as small as possible

$\vec{e}$ is the error vector,

$\vec{e} \perp \vec{p}$ then $\vec{p}^T \vec{e} = 0$

$\vec{p} + \vec{e} = \vec{b}$ $\quad \vec{e} = \vec{b} - \vec{p} = (\vec{b} - A\hat{x})$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = A \quad \begin{bmatrix} 3 \\ 11 \\ 7 \end{bmatrix} = \vec{b}$$

$(A\hat{x})^T (\vec{b} - A\hat{x}) = 0$

$A^T A \hat{x} = A^T \vec{b}$ $\quad \hat{x}^T A^T (\vec{b} - A\hat{x}) = 0$ $\quad$ if $\hat{x}^T \neq \vec{0}$, then

$\hat{x}^T (A^T \vec{b} - A^T A \hat{x}) = 0$ $\quad A^T \vec{b} - A^T A \hat{x} = 0$

SEE R

$\boxed{A^T A \hat{x} = A^T \vec{b}}$

$\hat{x} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ $\quad y = 3 + 2t$ $\quad \hat{x} = \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix}$

minimize $\|\vec{e}\|$

$\vec{p} = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}$ $\quad \vec{e} = \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix}$

$\vec{p} = A\hat{x}$

See R

$\vec{e} = \vec{b} - \vec{p}$