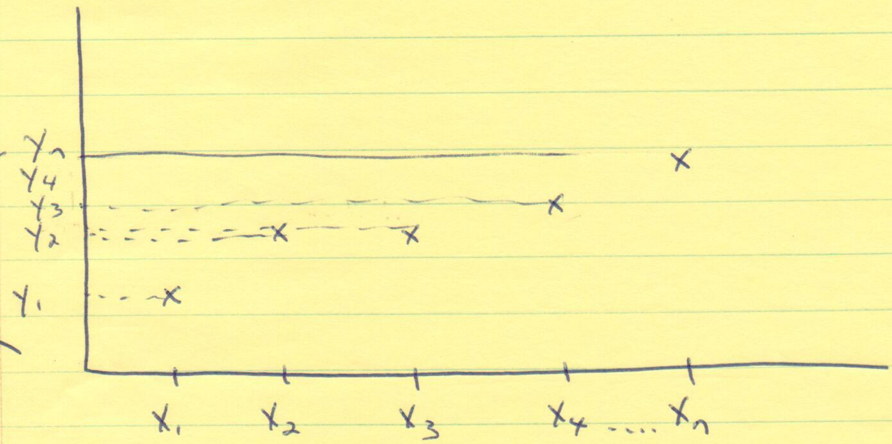
Calibration + Regression - Part 1

ME330 did linear regression. We will use a matrix approach that will eventually take sample size in to error calculation

* Must start by fitting straight line to data *

linear

Y dependant



measure these

x - independant

data table

Y	X
y_1	x_1
y_2	x_2
y_n	x_n

set these - assume <u>no</u> uncertainty

equation for a line

$$y = \beta_0 + \beta_1 x$$

choose β_0 + β_1 to best fit data

Put table in matrix format - cause were good engineers

$$\{y\} = \begin{Bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{Bmatrix} \qquad [X] = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_n \end{bmatrix}$$

column vector

$[n \times 2]$ rectangle vector

$\{y\}^T$ = row vector

Calibration + Regression part 1

consider the value $\{y\} - [X]\{\beta\}$ → broken down below

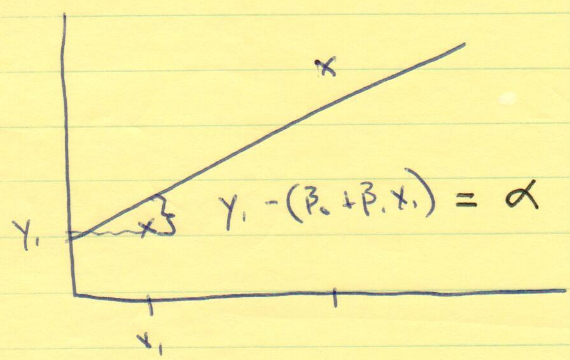
$$\{\beta\} = \begin{Bmatrix} \beta_0 \\ \beta_1 \end{Bmatrix}$$

$$[X]\{\beta\} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_n \end{bmatrix} \begin{Bmatrix} \beta_0 \\ \beta_1 \end{Bmatrix}$$

$$[n \times 2] \; \{2 \times 1\} = n \times 1$$

$$= \begin{Bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \beta_0 + \beta_1 x_3 \\ \beta_0 + \beta_1 x_n \end{Bmatrix}$$

$$\{y\} - [X]\{\beta\} \longrightarrow \begin{Bmatrix} y_1 - (\beta_0 + \beta_1 x_1) \\ y_2 - (\beta_0 + \beta_1 x_2) \\ y_3 - (\beta_0 + \beta_1 x_3) \\ y_n - (\beta_0 + \beta_1 x_n) \end{Bmatrix} = \underbrace{\begin{Bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_n \end{Bmatrix}}_{\text{residuals}}$$



$\beta_0 + \beta_1$ still unknown

$y_1 - (\beta_0 + \beta_1 x_1) = \alpha$ = difference from each data point to the prediction

= residuals

square each element of α and add up. choose $\beta_0 + \beta_1$ to minimize this value

"Best fit" mean least squares

if we have vector $\{\alpha\}$, the L^2 (least square) norm is defined as

$$\|\{\alpha\}\|_2 = \sum_{i=1}^{n} (\alpha_i)^2$$

will choose $\{\beta\}$ to minimize $\|\{\alpha\}\|_2$