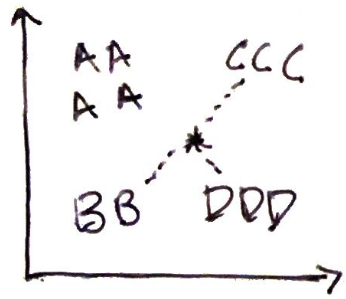
Math 342W Lecture 6

$y = \{A, B, C, D, \ldots\}$



Now, $y = \mathbb{R}$ (output space is $\in \mathbb{R}$)

Models for this Response are called "Regression"

Definition: To go backwards

Null Model: $g_0 = \bar{y}$

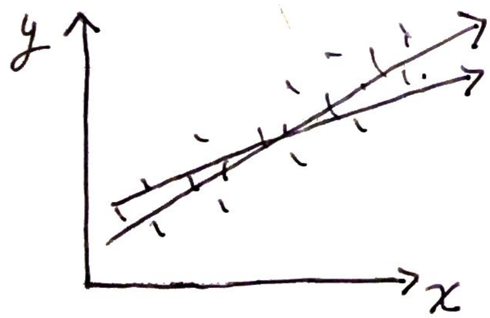
Let $\vec{x} \in \mathcal{X} = \mathbb{R}^p \implies \mathcal{H} = \{ \vec{w} \cdot \vec{x} : \vec{w} \in \mathbb{R}^{p+1} \}$

(Linear Model) $\quad \hookrightarrow = w_0 + w_1 x_1 + \ldots + w_p x_p$

$h^*(\vec{x}) = \underbrace{w_0^*}_{\beta_0} + \underbrace{w_1^* x_1}_{\beta_1} + \ldots + \underbrace{w_p^* x_p}_{\beta_p} \implies p$ dimensional line/plane

$\beta_0 + \beta_1 + \ldots + \beta_p \implies$ "Betas" (Econometrics)

We can visualize $p = 1$ (1 parameter)



$\rightarrow h^*(x) \implies \beta_0 + \beta_1 x \implies$ Perfect (Unobtainable)

$\rightarrow g(x) \implies b_0 + b_1 x \implies$ Imperfect (Obtainable)

Error Functions

$\sum_{i=1}^{n} |y_i - \hat{y}_i| \implies$ SAE (Sum Absolute Error)

$\underset{w_0, w_1 \in \mathbb{R}}{\arg\min} \{SAE\} = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$

$\frac{\partial}{\partial w_0} \left[\sum_{i=1}^{n} |y_i - \hat{y}_i| \right] \quad \frac{\partial}{\partial w_1} \left[\sum_{i=1}^{n} |y_i - \hat{y}_i| \right]$

THIS DOESN'T WORK

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \Rightarrow SSE \text{ (Sum Square Error)}$$

$$\underset{w_0, w_1 \in \mathbb{R}}{\operatorname{argmin}} \{SSE\}$$

$$\frac{\partial}{\partial w_0}\left[\sum_{i=1}^{n} (y_i - \hat{y}_i)^2\right] = \frac{\partial}{\partial w_0}\left[\sum_{i=1}^{n} y_i^2 + w_0^2 + w_1^2 x_i^2 - 2y_i w_0 - 2y_i w_1 x_i + 2w_0 w_1 x_i\right]$$

$$= \sum y_i^2 + n w_0^2 + w_1^2 \sum x_i^2 - 2n w_0 \bar{y} - 2w_1 \sum x_i y_i + 2n w_0 w_1 \bar{x}$$

$$\left\{ \bar{x} n = \sum x_i \text{ b/c } \bar{x} = \frac{1}{n} \sum x_i \right\}$$

$$= w_0 = \bar{y} + w_1 \bar{x}$$

$$\Rightarrow b_0 = \bar{y} + b_1 \bar{x}$$

For $\frac{\partial}{\partial w_1} \Rightarrow b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$

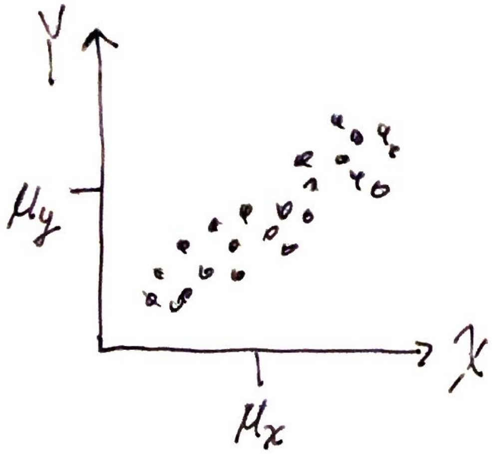
$\sigma_x^2 = \text{Var}[x]$ estimate by $s_x^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

$$= \frac{1}{n-1} \sum x_i^2 - 2x_i \bar{x} + \bar{x}^2 = \frac{1}{n-1} \sum x_i^2 - 2n\bar{x}^2 + n\bar{x}^2$$

$$\Rightarrow (n-1) s_x^2 = \sum x_i^2 - n\bar{x}^2$$

$$\text{Corr}[X,Y] = \frac{\text{Cov}[X,Y]}{SD[X]SD[Y]} = \frac{E[(X-\mu_x)(Y-\mu_y)]}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \overset{\text{Eventually}}{\Rightarrow} \frac{S_{xy}}{S_x^2}$$

Covariance



Ind. | $P(Y|X=x_1) = P(Y|X=x_2)\ \forall x_1, x_2$

Dep. | $P(Y|X=x_1) \neq P(Y|X=x_2)\ \exists x_1, x_2$

Covariance is a measure of <u>Linear</u> Dependence.

b_0 and b_1 are the least squares estimates (OLS) ordinary least squares

Sample Points the more you add is $O(n^{-\frac{1}{2}})$ error rate

→ Convergence of error to 0.

$X = \{A, B\}$, $y = \mathbb{R}$

$A \to 0$, $B \to 1$

$g(0) = \bar{y}_0$
$g(1) = \bar{y}_1$ ⇒ OLS Estimates as well