02/14

$\mathcal{Y} = \{0, 1\} \Rightarrow$ Binary Classification Model.

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

$\downarrow$

converts all $h \in \mathcal{H}$ to have range $\mathcal{Y}$

what if $\mathcal{Y} = \mathbb{R}$ or $\mathcal{Y} \subset \mathbb{R}$

ie the response

is continuous

$\rightarrow$ Regression Models. $\Rightarrow$ shoul be continuous model.

$\downarrow$

Best way to minimize the errors just for continuous

Null Model

$\hookrightarrow g_0 = \bar{y}$ sample average

$g_0 = g_{null}$

$\mathbb{D} = \langle X, y \rangle$

$\mathbb{D}_0 = y$

OUR FIRST NON TRIVIAL MODEL

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

$= \{ w_0 + w_1 x_1 + \dots + w_p x_p : w_0, w_1, \dots, w_p \in \mathbb{R} \}$

if $p = 1$

$\mathcal{H} = \{ w_0 + w_1 x = w_0, w_1 \in \mathbb{R} \}$

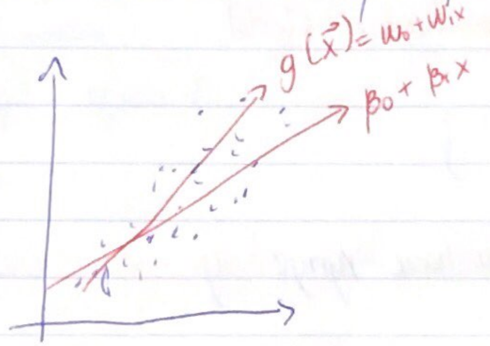
$h^*(\vec{x}) = w_0^* + w_1^* x_1 + \dots + w_p^* x_p$

$= \beta_0 + \beta_0 x_1 + \dots \beta_p x_p$ econometrics

this represents the true linear model

$$Y = h^*(\vec{x}) + \varepsilon = \beta_0 + \beta_1 x_1 + \dots + \beta_p^* p + \varepsilon$$

misspecification & ignorance error



How to "fit" g. ie such an element of $\mathcal{H}$ which is a good model? then

first we need to specify an objective function that reflects error in the model. Then, select the g and minimize the error

$$* \quad SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Sum square errors

this is our metric to measure error

p = 1 We want minimal SSE linear function

$$\hat{w}_0\ \hat{w}_1 = \underset{w_0, w_1 \in \mathbb{R}}{\text{argmin}} \{SSE\} \qquad \hat{y}_i = g(x_i)$$

$$= \text{argmin} \left\{ \sum (Y - (w_0 + w_1 x_i))^2 \right\}$$

$$\bar{y} = \frac{1}{n} \Sigma y$$

$$\Sigma (y_i - w_0 - w_1 x_i)^2$$

$$= \Sigma \; y^2 - w_0^2 - w_1^2 x_i^2 - 2 y_i w_0 - 2 y_i w_1 x_i + 2 w_0 w_1 x_i$$

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

$$\frac{\partial}{\partial w_0}[SSE] = 2 n w_0 - 2 n \bar{y} + 2 w_1 n x = 0$$

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

$$\frac{\partial}{\partial w_1}[SSE] = 2 w_1 \Sigma x_i^2 - 2 \Sigma x_i y_i + 2 w_0 n \bar{x} = 0$$

$$w_1 \Sigma x_i^2 = \Sigma x_i y_i - w_0 n \bar{x}$$

$$w_1 \Sigma x_i^2 = \Sigma x_i y_i - (\bar{y} - w_1 \bar{x}) n \bar{x} =$$

$$\Sigma x_i y_i - n \bar{x} \bar{y} + w_1 n \bar{x}^2$$

$$\Rightarrow w_1 (\Sigma x_i^2 - n \bar{x}^2) = \Sigma x_i y_i - n \bar{x} \bar{y}$$

$$\Rightarrow \hat{w}_1 = \frac{\Sigma x_i y_i - n \bar{x} \bar{y}}{\Sigma x_i^2 - n \bar{x}^2}$$

$$\rho : \text{corr}[x, y] = \frac{\text{cov}[x, y]}{SE[x]\, SE[y]}$$ $\Rightarrow$ this is unitless

correlation

$$\sigma_{xy} := \text{cov}[x, y] := E[(x - \mu_x)(y - \mu_y)] \Rightarrow$$

sigma

Simple covariance $\rightarrow S_{xy} := \frac{1}{n-1} \sum (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} \Big(\rightarrow$

$$\rightarrow \Big(\sum x_i y_i - \underbrace{\sum \bar{x} y_i}_{-n\bar{x}\bar{y}} - \underbrace{\sum x_i \bar{y}}_{-n\bar{x}\bar{y}} + \sum \bar{x}\bar{y} \Big) = \frac{1}{n-1} \Big(\sum x_i y_i + n\bar{x}\bar{y} \Big)$$

$\sigma_x^2 = Var(x)$ est by $S_x^2 := \frac{1}{n-1} \sum (x_i - \bar{x})^2$

$$= \frac{1}{n-1} \Big(\sum x_i^2 - 2 \underbrace{\sum x_i \bar{x}}_{n\bar{x}^2} + \underbrace{\sum \bar{x}^2}_{n\bar{x}^2} \Big)$$

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

estimated by

$r = \frac{S_{xy}}{S_x S_y}$

$$\hat{\beta}_1 = b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{(n-1) S_{xy}}{(n-1) S_x^2} = \frac{S_{xy}}{S_x^2} = r \frac{S_y}{S_x}$$

$$\lim_{n \to \infty} b_0 = \beta_0, \quad \lim_{n \to \infty} b_1 = \beta_1$$

$$\rightarrow g(x) \rightarrow h^*(x)$$

How well does g predict?

p = y = Data
deltay_i = Y[i] - Yhat[i]

```
for ( j in 1:(p+1)) {

    W[j] = W[j] + delta_y_i[j] }
```

02/19

AKA

$p = 1$

$\mathcal{Y} = \mathbb{R}$, p covariants

$\mathcal{H} = \{ \vec{w} \cdot \vec{x} : \vec{w} \in \mathbb{R}^{p+1} \} = \{ w_0 + w_1 x : w_0, w_1 \in \mathbb{R} \}$

$A = OLS$

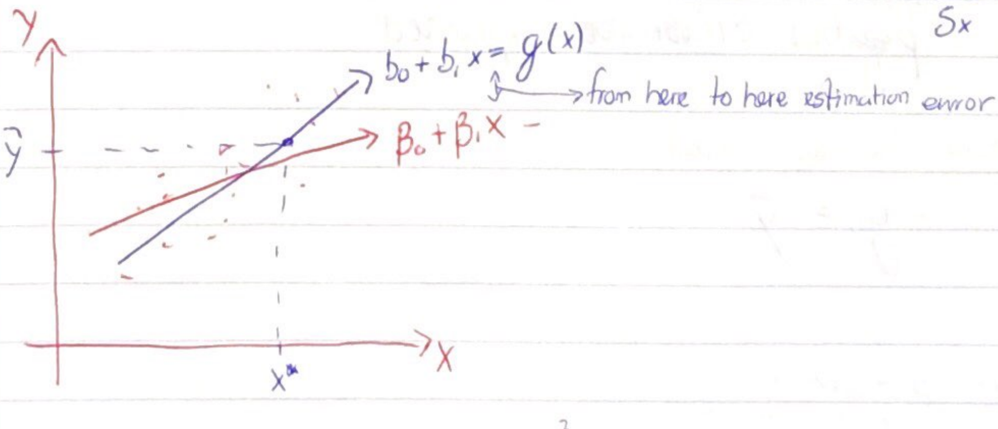
algorithm minimize SSE

$\xrightarrow{A} g(x) = b_0 + b_1$

$b_0 = \bar{y} - r \frac{s_y}{s_x} \bar{x}$

$h^*(x) = \beta_0 + \beta_1 x$

$b_1 = r \frac{s_y}{s_x}$



$\frac{4+6}{4}$ $\frac{12}{4} = 3$