Review

02/21/19

$$y = \mathbb{R}, \quad x_{raw} \in \mathcal{X} = \{ red, green \}.$$

$$\Downarrow \qquad \downarrow \qquad \downarrow$$

$$x \in \mathcal{X} = \{0, 1\}$$

$$g(x_{raw}) = \begin{cases} \bar{y}_r & \text{if } x = red \\ y_g & \text{if } x = green \end{cases}$$

$$g(x) = \underbrace{\bar{y}_r}_{b_0} + \underbrace{(\bar{y}_g - \bar{y}_r)}_{b_1} x$$

OLS solution.

$$b_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} \qquad \bar{x} = \frac{\sum x}{n}$$

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = r\frac{s_y}{s_x} \Rightarrow \frac{n_g\bar{y}_g - np_g\bar{y}}{n_g - np_g^2} = \frac{p_g n\bar{y}_g - np_g\bar{y}}{p_g n - np_g^2} = \frac{\bar{y}_g - \bar{y}}{1-p_g}$$

$\bar{x} = \frac{1}{n}\sum x_i = p_g$ (proportion green in $\mathbb{D}$)

$$\Rightarrow \frac{\bar{y}_g - (p_g\bar{y}_g + (1-p_g)\bar{y}_r)}{1-p_g}$$

$\sum x_i^2 = n_g$ (# of greens) w/ $n_g = p_g n$

$$= \frac{(1-p_g)\bar{y}_g - (1-p_g)\bar{y}_r}{1-p_g}$$

$\sum x_i y_i = n_g\bar{y}_g = y_{g_1} + \dots + y_{g n_g}$

$$= \bar{y}_g - \bar{y}_r$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{(y_{g_1} + \dots + y_{g n_g})}{n} + \frac{(y_{r_1} + \dots + y_{r n_r})}{n}$$

$$= \frac{n_g\bar{y}_g}{n} + \frac{n_r \bar{y}_r}{n} = p_g\bar{y}_g + (1-p_g)\bar{y}_r \qquad b_1 = \frac{\Delta y}{\Delta x} \quad \Delta x = 1 \Rightarrow \Delta y$$

$$b_0 = \bar{y} - b_1\bar{x} = \bar{y} - (\bar{y}_g - \bar{y}_r)p_g$$

$$= (p_g\bar{y}_g + (1-p_g)\bar{y}_r) - p_g(\bar{y}_g - \bar{y}_r)$$

$$= \cancel{p_g\bar{y}_g} + (1-p_g)\bar{y}_r - \cancel{p_g\bar{y}_g} + p_g\bar{y}_r$$

$$= \bar{y}_r$$

If there are three separate variables there are two dimentions

$y = \mathbb{R}$, $x_{raw} \in \mathcal{X} = \{red, green, blue\}$

$p = 2$ $\quad x_1 \in \mathcal{X} = \{0, 1\}$ $\quad x_2 \in \{0, 1\}$

three dimention for

$L = 3$

if $x_1 = 1 \Rightarrow x_{raw} = green$

if $x_2 = 1 \Rightarrow x_{raw} = blue$

$\underbrace{\text{if } x_1 = x_2 = 0 \Rightarrow x_{raw} = red}_{\text{"reference" or "base" category}}$

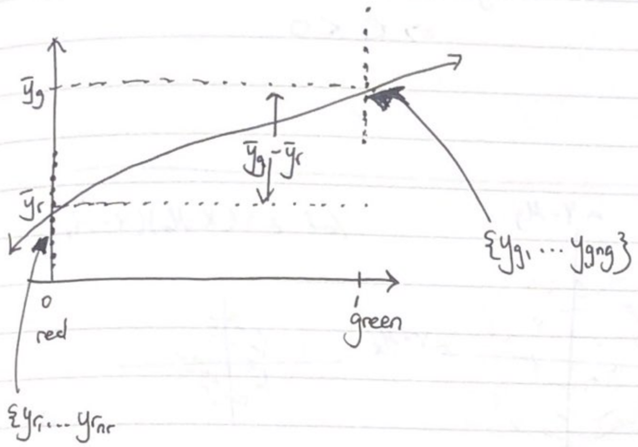
$$g(x) = \underbrace{\bar{y}_r}_{b_0} + \underbrace{(\bar{y}_g - \bar{y}_r)}_{b_1}x_1 + \underbrace{(\bar{y}_b - \bar{y}_r)}_{b_2}x_2$$

because intercept $\Rightarrow L-1$ dimention

$p + 1 = 3 = L$

Alternative Para (no intercepts)

$\mathcal{L} = \{ \omega_1 x_1 + \omega_2 x_2$



Consider two r.vs X, Y they are "dependents" if (or "associated")

If $\exists\, x_1, x_2$ s.t $P(Y \mid X = x_1) \neq P(Y \mid X = x_2)$

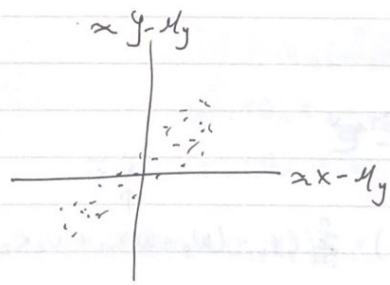
ρ = ~~Covariance~~ linear correlation $[X, Y] = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \in [-1, 1]$ unitless estimated by r

$\sigma_{xy} = \mathbb{E}[(X - \mu_x)(Y - \mu_y)]$ estimated by S_{xy}

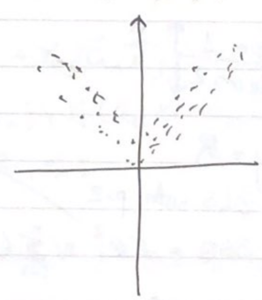
"Cov $[X, Y]$"

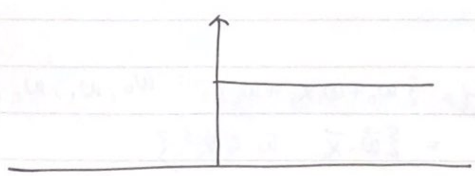
covariance

possitive linear asso.
possitive correlation.



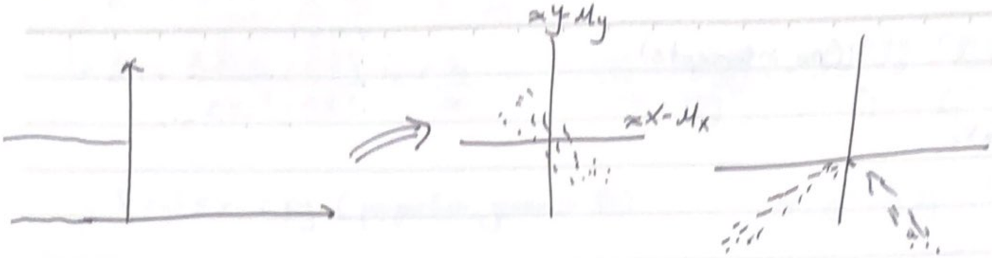
let $z = (x - \mu_x)(y - \mu_y)$

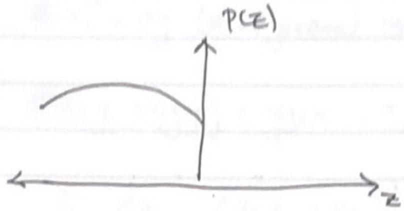




$\mathbb{E}[(X - \mu_x)(Y - \mu_y)] \approx \mathbb{E}(z) > 0$

$\Rightarrow \rho > 0$



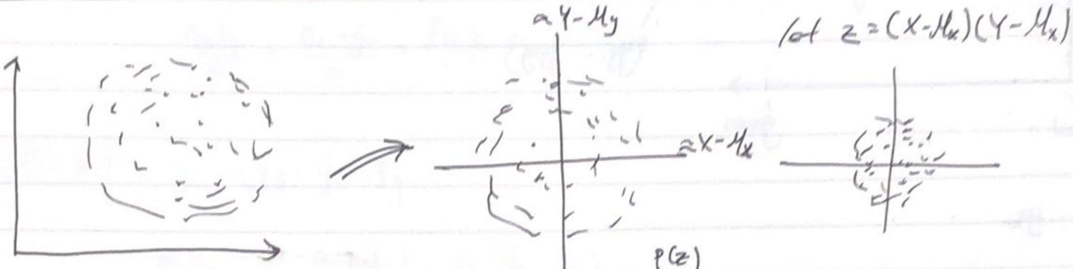


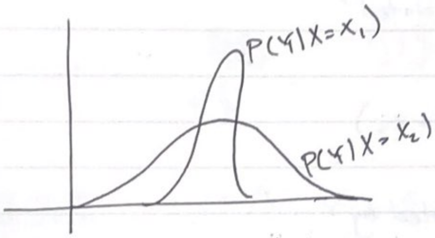
$E[(X-\mu_x)(Y-\mu_y)] \approx E(z) < 0$

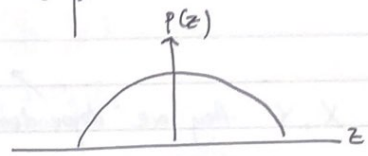
$\Rightarrow e < 0$

---

let $z = (X-\mu_x)(Y-\mu_x)$







$E[(X-\mu_x)(Y-\mu_y)] \approx E[z] = 0$

$\Rightarrow e = 0$

Association $\xrightarrow{\text{does not imply}}$ linear correlation.

$\boxed{\text{Prove } R^2 = r^2 \text{ if } p=1}$

Linear correlation $\overset{\text{is a type of}}{\in}$ {association}

$\frac{\text{midterm 1} \uparrow}{\text{midterm 2} \downarrow}$

$y = \mathbb{R}$ $\quad \vec{e}^T\vec{e}$

OLS with p=2

$SSE = \sum e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i) = \sum_{i=1}^{n} (y_i - (w_0 + w_1 x_{i1} + w_2 x_{i2}))^2 = \sum (y_i - \vec{w}\cdot\vec{x}_i)^2$

$\vec{w}\cdot\vec{x}_i$

$b_2 : \frac{\partial}{\partial}[SSE] = 0$

$\mathbb{D} = \langle X, y \rangle \quad X = \underbrace{\begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} \end{bmatrix}}_{p} \Big\} n$

$\mathcal{H} = \{ w_0 + w_1 x_1 + w_2 x_2 : w_0, w_1, w_2 \in \mathbb{R} \}$

$= \{ \vec{w}\cdot\vec{x} : \vec{w} \in \mathbb{R}^3 \}$

$$\hat{y} = \underset{(n\times(p+1))}{X} \vec{w} = \begin{bmatrix} w_0 + w_1 x_{11} + w_2 x_{12} \\ w_0 + w_1 x_{21} + w_2 x_{22} \\ \vdots \\ w_0 + w_1 w_{n1} + w_2 x_{n2} \end{bmatrix} \in \mathbb{R}^{n\times 1}$$

Note: $(\vec{a}+\vec{b})^T = \vec{a}^T + \vec{b}^T$

$(\vec{a}+\vec{b})^T = \vec{b}^T\vec{a}$

$(x\vec{w})^T = \vec{w} + x^T$

$$\Rightarrow SSE = \Sigma e_i^2 = \vec{e}^T e = (\vec{y}-\vec{\hat{y}})^T(\vec{y}-\vec{\hat{y}})$$

$\vec{e}^T\vec{e}$

$$\Rightarrow (\vec{y}^T - \vec{\hat{y}}^T)(\vec{y}-\vec{\hat{y}}) = \vec{y}^T\vec{y} - \vec{\hat{y}}\vec{y} - \vec{y}^T\vec{\hat{y}} + \vec{\hat{y}}^T\vec{\hat{y}}$$

$$= \vec{y}^T\vec{y} - 2\vec{\hat{y}}^T\vec{y} + \vec{\hat{y}} + \vec{\hat{y}} = \vec{y} + \vec{y} - 2\vec{w}^T x^T\vec{y} + \vec{w}^T x^T X\vec{w}$$

$$\frac{\partial}{\partial\vec{w}}[SSE] = 0_{p+1}$$

col vector

$a \in \mathbb{R}^n$

Let A be an $n\times n$ symmetric matrix constant w.r.t $\vec{x}$

$$\frac{\partial}{\partial\vec{x}}[\underbrace{\vec{x}^T A\vec{x}}_{\text{Quadratic Form}}] = \begin{bmatrix} \frac{\partial}{\partial x_1}[bf(\vec{x}) + cg(\vec{x})] \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} = b\frac{\partial}{\partial\vec{x}}[f(\vec{x})] + c\frac{\partial}{\partial\vec{x}}[g(\vec{x})]$$

$$A\vec{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \quad\quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{n}x_n \end{bmatrix} = \begin{bmatrix} \leftarrow \vec{a}_1 \rightarrow \\ \leftarrow \vec{a}_2 \rightarrow \\ \vdots \\ \leftarrow \vec{a}_n \rightarrow \end{bmatrix}\begin{bmatrix} \uparrow \\ \vec{x} \\ \downarrow \end{bmatrix} = \begin{matrix} \vec{a}_1\vec{x} \\ \vec{a}_2\vec{x} \\ \vdots \\ \vec{a}_n\vec{x} \end{matrix}$$

$$\vec{x}^T(A\vec{x}) = [x_1 \dots x_n]\begin{bmatrix} \vec{a}_1\vec{x} \\ \vec{a}_2\vec{x} \\ \vdots \\ \vec{a}_n\vec{x} \end{bmatrix} = x_1(\vec{a}_1\vec{x}) + x_2(\vec{a}_2\vec{x}) + \dots + x_n(\vec{a}_n\vec{x})$$

$$= x_1(a_{11}x_1 + a_{12}x_2 + \dots + a_n x_n) + x_2(a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x) + x_n(a_{n1}x_1 + \dots + a_{n1}x_n)$$

$$\Rightarrow a_{11}x_1^2 + a_{12}x_2x_1 + \dots + a_{1n}x_nx_1 + a_{21}x_1x_2 + a_{22}x_2^2) + \dots + (a_{2n}x_nx_2) + \dots + (a_{n}x_1x_n + a_{n2}x_2x_n + \dots + a_{nn}x_1^2)$$

$$\frac{\partial}{\partial\vec{x}}[\vec{x}^T A\vec{x}] =$$

$$\frac{\partial}{\partial x_1}[\dots] = (2a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) + (a_{21}x_2) + \dots (a_{n1}x_n) = 2(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) = 2\vec{a}_1\vec{x}$$

$$\frac{\partial}{\partial x_2}[\dots] = (a_{12}x_1 + 2a_{22}x_2 + \dots + a_{2n}x_n) + \dots + (a_{n2}x_n) = 2(a_{12}x_1 + a_{22}x_2 + \dots + a_{2n}x_n) = 2\vec{a}_2\vec{x}$$

$\vdots$

$$\frac{\partial}{\partial x_n}[\dots] = (a_{1n}x_1) + (a_{2n}x_2) + \dots + (a_{n1}x_n + a_{n2}x_2 + \dots + 2a_{nn}x_n) = 2(a_{n1}x_1 + a_{n2}x_2 + \dots a_{nn}x_n) = 2\vec{a}_n x$$

$$= 2\begin{bmatrix} \leftarrow \vec{a}_1 \rightarrow \\ \leftarrow \vec{a}_2 \rightarrow \\ \vdots \\ \leftarrow \vec{a}_n \rightarrow \end{bmatrix}\begin{bmatrix} \uparrow \\ \vec{x} \\ \downarrow \end{bmatrix} = 2A\vec{x}$$

Assuming $X^T X$ symmetric.
$= X^T (X^T)^T = X^T X$ ✓

$$\frac{\partial}{\partial \vec{w}}\left[\vec{y} + \vec{y} - 2\vec{w}^T X^T \vec{y} + w^T X^T X \vec{w}\right] =$$

$$= \frac{\partial}{\partial \vec{w}}\left[\vec{y} + \vec{y}\right] - 2\frac{\partial}{\partial \vec{w}}\left[\vec{w}^T X^T \vec{y}\right] + \frac{\partial}{\partial \vec{w}}\left[\vec{w}^T X^T X \vec{w}\right] = \vec{0}_{p+1} - 2X^T\vec{y} + 2X^T X\vec{w} \overset{set}{=} \vec{0}_{p+1}$$

$$X^T X\vec{w} = X^T\vec{y} \Rightarrow \cancel{(X^TX)^{-1}(X^Ty)} = (X^TX)^{-1}X^T\vec{y} \Rightarrow \boxed{\vec{b} = (X^TX)^{-1}X^T\vec{y}}$$

OLS solution.

Assuming X^TX is invertible

If $\text{rank}[X] = p+1 \Rightarrow X^TX$ is invertible.

$\parallel$

$\dim[\text{col}\ [X]] = p+1$

$\Downarrow$

p features and T_n (

are linearly independent. i.e. no duplicate information.

<u>Proof.</u>

X^TX not invertible $\Rightarrow$ not full rank $\Rightarrow$ nullity $\neq 0 \Rightarrow \dim[\text{Nullspace}[X^TX]] > 0$

$\Rightarrow \exists \vec{v} \in \mathbb{R}^n$ non zero such that

$\Rightarrow X^T \underbrace{X\vec{v}}_{\vec{0}} = 0$

$X^T\vec{0} = \vec{0}$

$\vec{0}\ X = \vec{0}$