02/21

$\mathcal{Y} = \mathbb{R}$, $x_{row} \in \mathcal{X} = \{\text{red}, \text{green}\}$

price

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

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

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

OLS soln

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2} = r\frac{s_y}{s_x}$$

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

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

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

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

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

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

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

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

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

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

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

Now $p=2$ $\quad x_{raw} \in \mathcal{X} = \{red, green, blue\}$

$$g(x_{raw}) = \begin{cases} \text{if } x_1 = 1 \Rightarrow x_{raw} = green \\ \text{if } x_2 = 1 \Rightarrow x_{raw} = blue \\ \text{if } x_1 = x_2 = 0 \Rightarrow x_{raw} = red \end{cases}$$

"reform" or "baseline" Categ

The dummy for $<= 3$

$x_1 \in \mathcal{X} = \{0,1\}$ $\qquad x_2 \in \{0,1\}$

Because intersept $\Rightarrow$ $L-1$ dummies

$$p+1=3=L$$

Alternative parametrization

(no intersept)

$$\mathcal{H} = \{ w_1 x_1 + w_2 x_2 + w_3 x_3 : w_1, w_2, w_3 \in \mathbb{R} \}$$

$x_1 = 1$ if red

$x_2 = 1$ if green

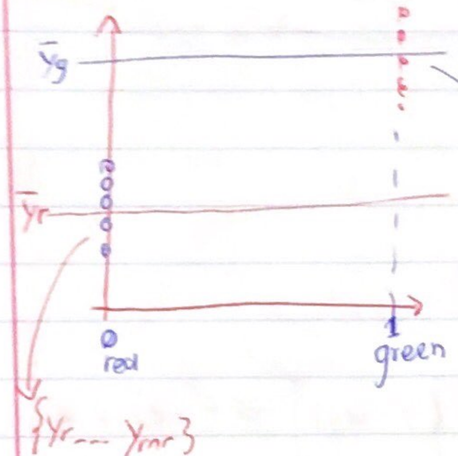
$x_3 = 1$ if blue

$$g(x) = \bar{y}_r x_1 + \bar{y}_g x_2 + \bar{y}_b x_3$$

$\downarrow$ this is clear

$$\underbrace{g(x) = \underbrace{\bar{y}_r}_{b_0} + \underbrace{(\bar{y}_g - \bar{y}_r)}_{b_1} + \underbrace{(\bar{y}_b - \bar{y}_r)}_{b_2}}_{\text{here you see change}}$$

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




Consider two random variabls $X, Y$

they are "dependent" if

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

or "associate"

Linear Correlation

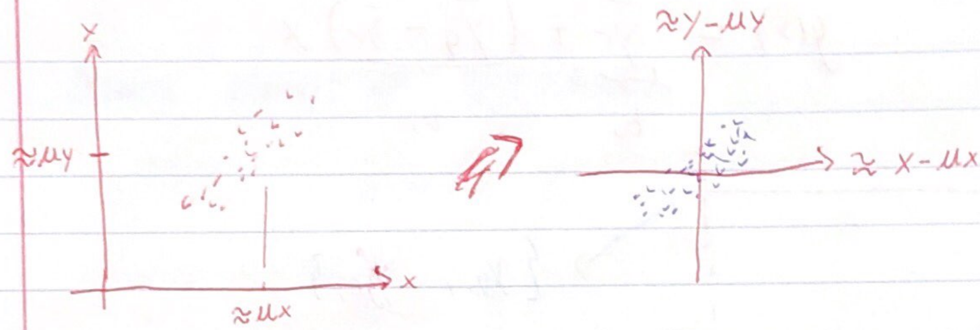
$$\rho := \mathrm{Corr}[X,Y] = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \in [-1,1] \rightarrow \text{estimated by } r$$

Unitless

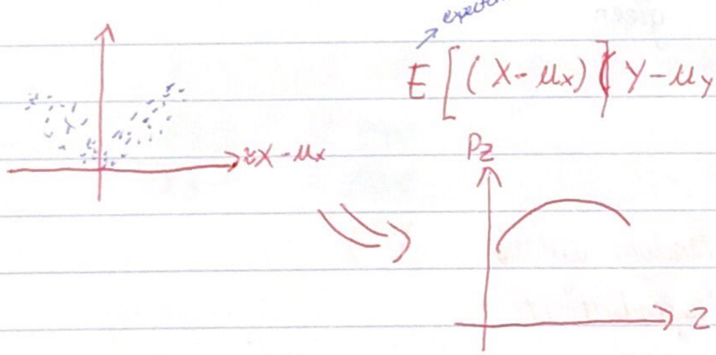
$$\sigma_{xy} = E[(X-\mu_x)(Y-\mu_y)] \rightarrow \text{estimated by } S_{xy}$$

"$\mathrm{Cov}[X,Y]$"

Covariance




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

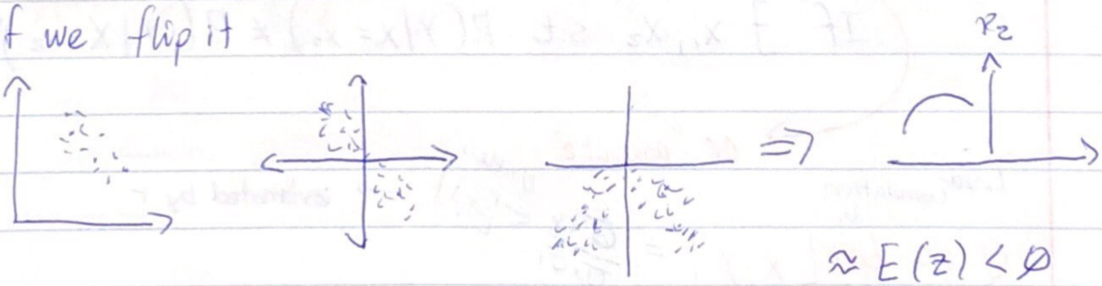



expectation

$E\left[(X-\mu_x)(Y-\mu_y)\right] \approx E(z) > 0$

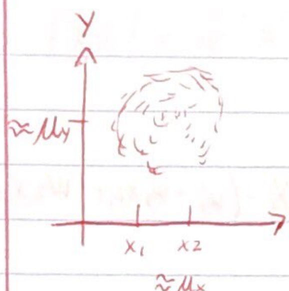
$\Rightarrow \rho > 0$

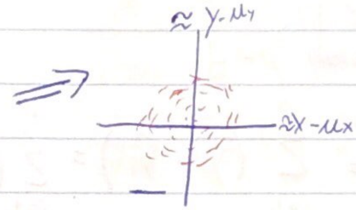
if we flip it



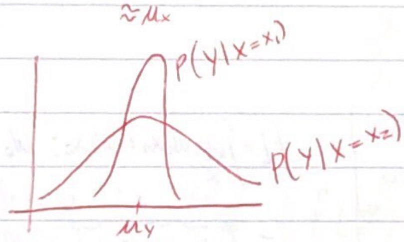

$\approx E(z) < 0$

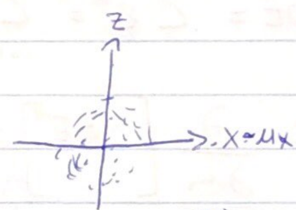
$\Rightarrow \rho < 0$



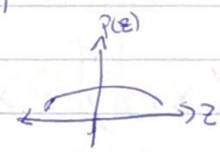




Correlation coefitien does not tell us if are dependent or independent




$E(z)=0$




$\varepsilon = 0$

Association ≠> Linear correlation

Linear correlation ~~≠~~ is a type of association }

$$\boxed{\begin{array}{c} R^2 = r^2 \\ \text{if } \rho = 1 \end{array}}$$

MIDTERM 2

$\mathcal{Y} = \mathbb{R}$

as when $p=2$

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

$$b_2 : \frac{\partial}{\partial m_1} \left[ SSE \right] \overset{set}{=} 0$$

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

$$\mathcal{H} = \{ w_0 + w_1 x_1 + w_2 x_2 : w_0, w_1, w_2 \in \mathbb{R} \}$$
$$= \{ \vec{w} \cdot \vec{x} : \vec{w} \in \mathbb{R}^3 \}$$

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

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

$$e_i^2 \rightarrow \vec{e}^T \vec{e}$$

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

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

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

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

WE ARE MINIMIZIN THE ERROR $= \emptyset$

$(X\vec{w})^T = \vec{w}^T X^T$

$$\frac{\partial}{\partial w_0}\left[SSE\right] \overset{set}{=} 0$$

$$\frac{\partial}{\partial w_1}\left[SSE\right] \overset{set}{=} 0$$

$$\vdots$$

$$\frac{\partial}{\partial w_p}\left[SSE\right] \overset{set}{=} 0$$

---

$c \in \mathbb{R}$ wrt $\vec{x}$

Define : $\frac{\partial}{\partial \vec{x}}\left[c f(\vec{x})\right] = \begin{bmatrix} \frac{\partial}{\partial x_1}\left[f(\vec{x})\right] \\ \frac{\partial}{\partial x_n}\left[f(\vec{x})\right] \end{bmatrix} =$

$\dim(\vec{x}) = n$

$a \in \mathbb{R}^n$ constan wrt $\vec{x}$

$$\frac{\partial}{\partial \vec{x}}\left[\vec{a}^T \vec{x}\right] = \begin{bmatrix} \frac{\partial}{\partial x_1}\left[a_1x_1 + a_2x_2 \cdots + a_nx_n\right] \\ \frac{\partial}{\partial x_n}\left[a_1x_1 + a_2x_2 \cdots + a_nx_n\right] \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \vec{a} \neq \vec{a}^T$$

this is always column vector

$b, c \in \mathbb{R}$

$$\frac{\partial}{\partial \vec{x}}\left[b f(\vec{x}) + c\, g(\vec{x})\right] = \begin{bmatrix} \frac{\partial}{\partial x_1}\left[b + f(\vec{x}) + c g(\vec{x})\right] \\ \frac{\partial}{\partial x_n}\left[\right. \end{bmatrix} = b\frac{\partial}{\partial \vec{x}}\left[f(\vec{x})\right] + c\frac{\partial}{\partial \vec{x}}\left[g(\vec{x})\right]$$

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

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

Quadratic form

$$A\vec{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \quad a_{2n}x_n \\ \vdots \\ a_{n1}x_n + x_{n2}x_2 + \quad a_{nn}x_n \end{bmatrix}$$

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

$$\vec{x}^T (A\vec{x}) = [x_1 \cdots x_n] \begin{bmatrix} a_1 \cdot \vec{x} \\ a_2 \cdot \vec{x} \\ \vdots \\ a_n \cdot \vec{x} \end{bmatrix} =$$

$$\Rightarrow x_1(\vec{a}_1 \vec{x}) + x_2(\vec{a}_2\vec{x}) + \quad + x_n(\vec{a}_n\vec{x})$$

$$= x_1(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n) \cdots$$

$$= a_{11}x_1^2 + a_{12}x_2x_1$$

Take derivative $x_1^2$

$$\frac{\partial}{\partial x_1}[\cdots] = (2(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n) + (a_{21}x_2) + \cdots + (a_{n1}x_1) \Rightarrow$$

$$\frac{\partial}{\partial x_2}[\cdots] = (a_{1n}x_1) + (a_{21}x_1 + 2a_{22}x_2 + \cdots + a_{2n}x_n) + \cdots + (a_{n2}x_2) \Rightarrow$$

$$\vdots$$

$$\frac{\partial}{\partial x_n}[\cdots] = (a_{1n}x_1) + (a_{2n}x_2) + \cdots + (a_{n1}x_n + a_{n2}x_2 + \cdots + 2a_{nn}x_n) \Rightarrow$$

$$\Rightarrow 2\left(a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n\right) = 2\vec{a}_1 \cdot \vec{x}$$

$$\Rightarrow 2\cdot\left(a_{12}x_1 + a_{22}x_2 + \; + a_{2n}x_n\right) = 2\vec{a}_2 \cdot \vec{x} \quad = 2A\vec{x}$$

$$\Rightarrow 2\left( \qquad\qquad\qquad \right) = 2\vec{a}_n \cdot \vec{x}$$

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

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

$$\Rightarrow \underset{p+1}{\vec{0}} \quad - 2\underset{p+1\times n}{X^T}\underset{n\times p+1}{\vec{y}} \quad + 2X^T X\vec{w} \overset{set}{=} \vec{0}_{p+1}$$

dimention

assume $X^T X$ symetry

$$(X^T X)^T = \cancel{X}^T (X^T)^T = X^T X$$

We want to solve from $\vec{w}$

$$X^T X\vec{w} = X^T\vec{y} \Rightarrow (X^T X)^{-1}(X^T X)\vec{w} = (X^T X)^{-1} X^T\vec{y}$$

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

OLS soln

Assumption $X^T X$ is invertible

if rank $[X] = p+1 \Rightarrow X^T X$ is invertible

$\Updownarrow$

dim$[\text{colsp}(X)] = p+1$

dim$[\text{Nullsp}[X]] = 0$

$p$ features and In(-) are linearly independent is not duplicates

If $\text{rank}\left[X\right] = p+1 \Rightarrow X^T X$ is inversible

Proof

$X^T x$ not invertible $\Rightarrow$ not full rank $\Rightarrow$ Nullity $\neq 0$

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

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

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

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

WTS $\text{rank}(x) = p+1 \Rightarrow x^T x$ invertible

$x^T x$ noninvertible $\qquad \text{rank}(x) \neq p+1$

$\exists$

full rank

pretend $\text{rank}(x) = p+1$

$Xb = \bar{a} \neq \vec{0}$

$x^T($