$$\frac{\partial SSE}{\partial \beta_0} = \boxed{-2(y - \beta_0 - \beta_1 x) = 0}$$

$$\frac{\partial SSE}{\partial \beta_0} = (y - \beta_0 - \beta_1 x)^2$$

$$= 2(y - \beta_0 - \beta_1 x)$$

β_0 / β_1

$\beta_0 = ?$

$\beta_0\ \beta_1 \rightarrow SSE \downarrow$

$f - f(g(x))$

$$\frac{\partial f}{\partial \beta_0} = y - \beta_0 - \beta_1 x = -1$$

$$\frac{\partial SSE}{\partial \beta_1} = \boxed{-2x(y - \beta_0 - \beta_1 x) = 0}$$

∂ const $= 0$

$$x^n = n x^{n-1}$$

$$x^2 = 2x^{2-1}$$

$$= 2x$$

$$\beta_0^1$$

$$1 \beta_0^{1-1} = 1$$

$\frac{\partial SSE}{\partial \beta_0}$ = eq1 = 0 $\underline{2(y - \beta_0 - \beta_1 x) = 0}$

$\frac{\partial SSE}{\partial \beta_1}$ = eq^n = 0 $\Rightarrow$ $\underline{2x(y - \beta_0 - \beta_1 x) = 0}$

β_0 & β_1

$\beta_0 + \beta_1 x = y$

$\underline{\text{2 variable}}$ $\rightarrow$ $\underline{\text{150 variable}}$ $\rightarrow$ $\underline{750}$

$\underline{\text{150 equ}^n}$

( )

OLS

$\overset{n}{\text{equ}} = 1\,90 \rightarrow \beta_0 \;\; \beta_1 - \beta_m$

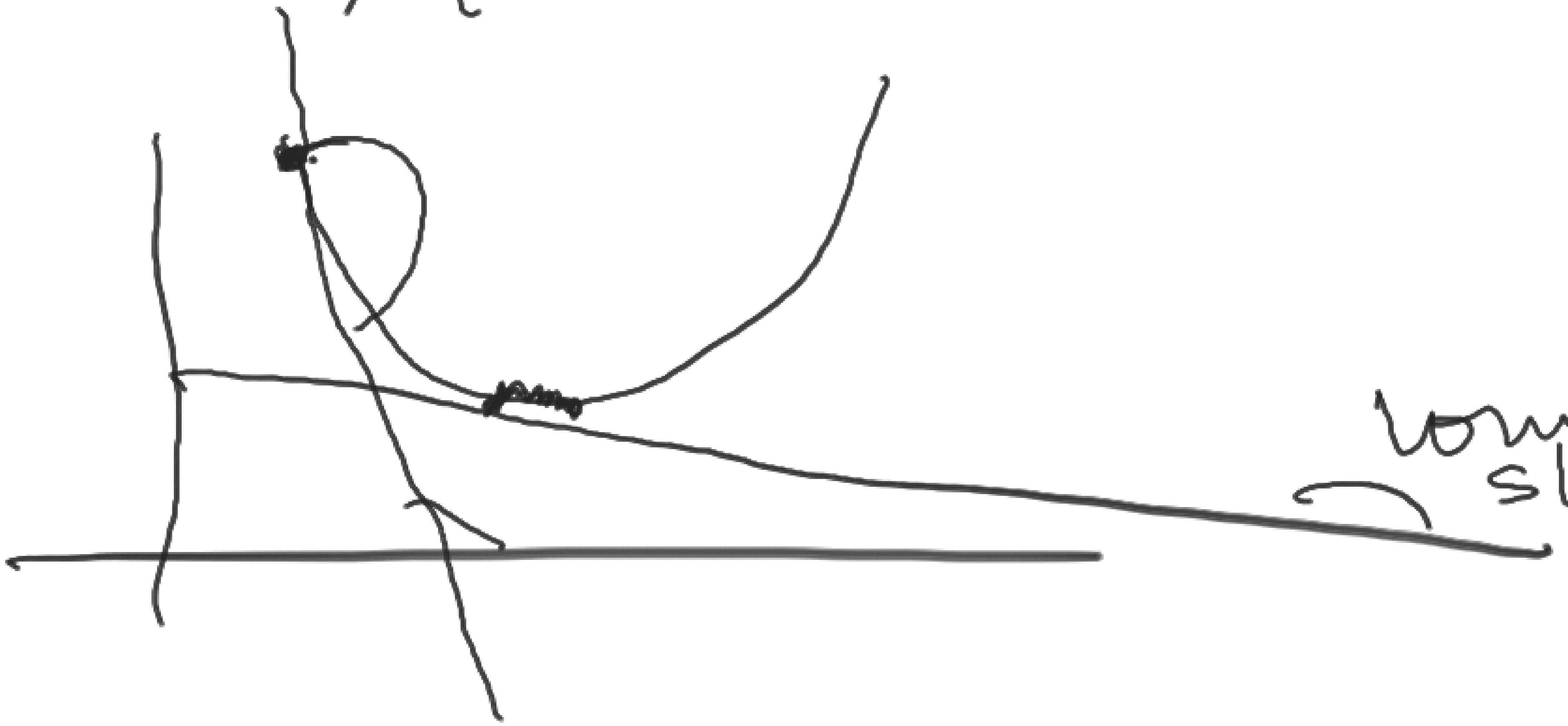
750 equ

GD

$\beta_i = \text{random}$

$\text{iter} = \beta_i - \beta_i + \alpha \dfrac{\partial SSE}{\partial \beta_i}$

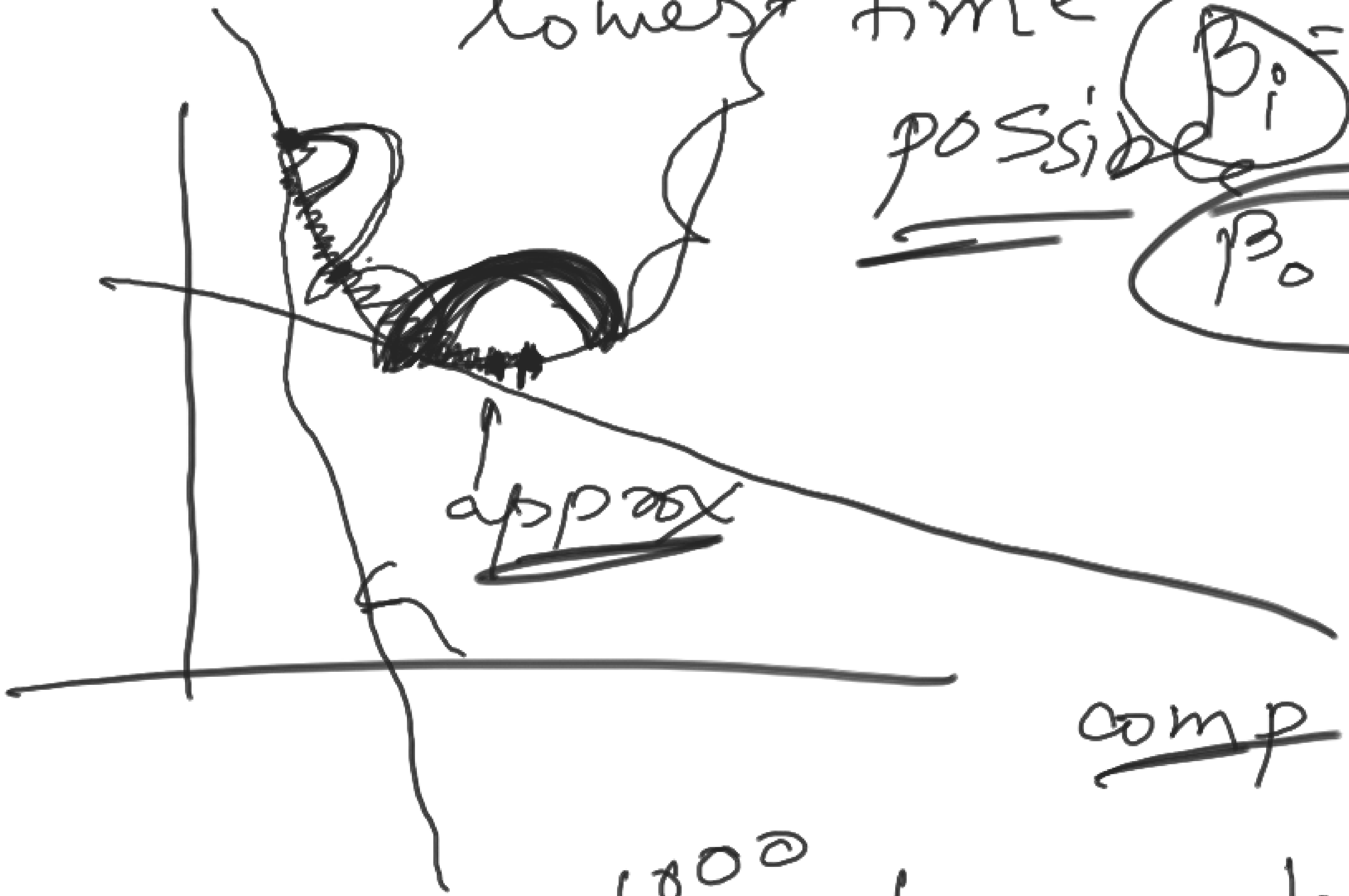
updating



lower slope

$\dfrac{dy}{dx}$

lowest time possible



$$\beta_i = \beta_i + \alpha \boxed{\frac{\partial SSE}{\partial \beta_i}}$$

$\beta_0 \quad \beta_1$

slope

$\alpha \uparrow$

comp $\quad \alpha \downarrow$

iterations = 1000 elements

$SSE_i - SSE_{i-1} < x$

~~hetero~~

p no.

error can be model

Boosting

XGboost

$\leftarrow$ error = f(sin)

homosceda

linear

Algo → Boosting

SSE → model

residual Best outcome

linear

$$1(y-\hat{y})^2 =$$

0	1
$\frac{1}{2}$	2
1	3
$\sqrt{2}$	
$\frac{\sqrt{3}}{2}$	4
1	5
$\frac{\sqrt{3}}{2}$	6
$\frac{1}{\sqrt{2}}$	7

hetero

100%

$y = ht \quad wt$

lower share

$BMI = f(ht, wt)$ x_3

share

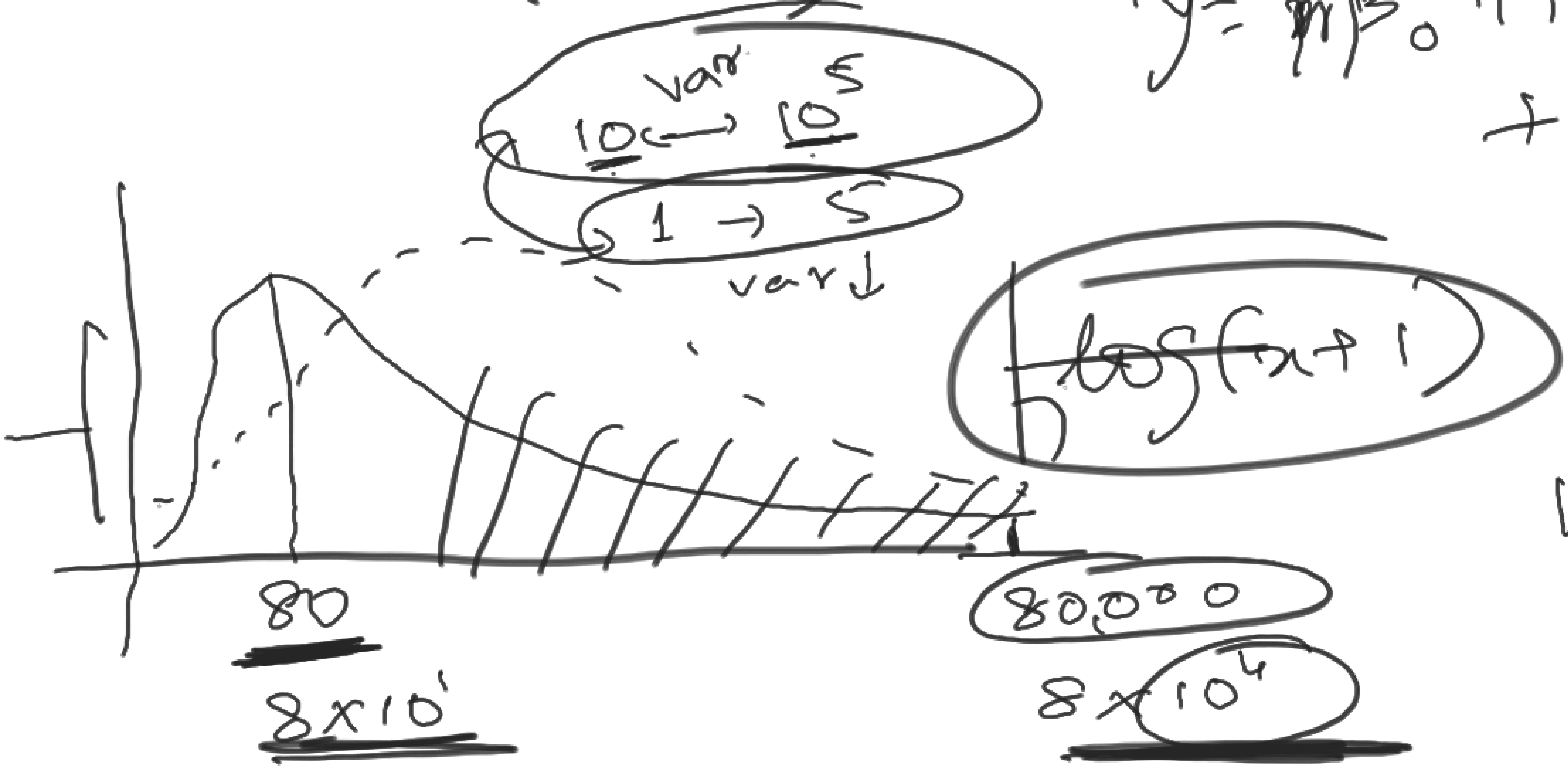
$y = m_1 ht + m_2 wt + m_3 \frac{ht}{wt}$

ht

$= m_1 ht \left(1 + \frac{m}{wt} \right)$

multicollinear → 2 indepen

$\sin(\quad)$



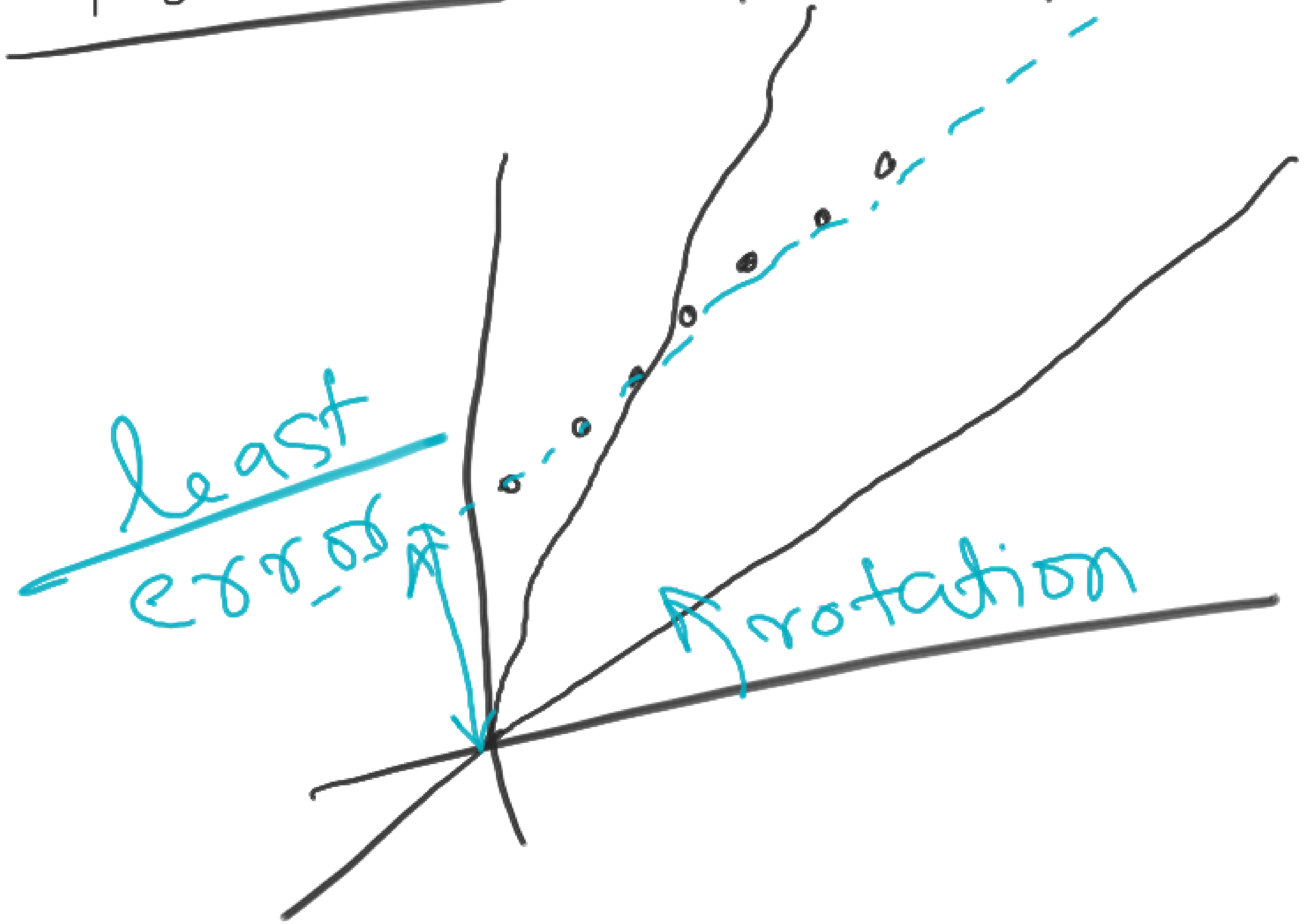
$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

$$\log_{10} 10^5 = 5$$

$$\log_{10} 10 = 1$$

Guided $\longrightarrow$ VIF

$\beta_0 = 0 \longrightarrow$ funć passes through origin



$\beta = 0$

minimize error

salary

projects

-ve

complaints

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

Base Salary

+ve

$\uparrow y \propto \beta_1 \uparrow \rightarrow$ +ve correla

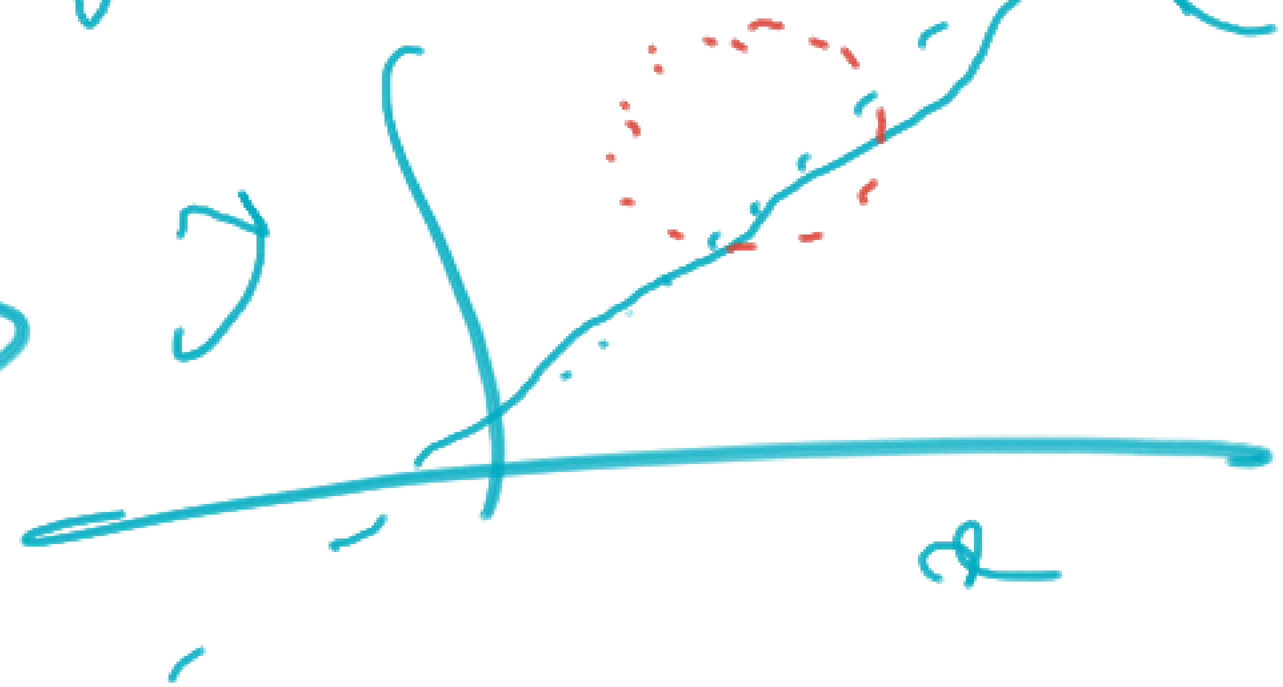
$\downarrow y$

$x_2 \uparrow$

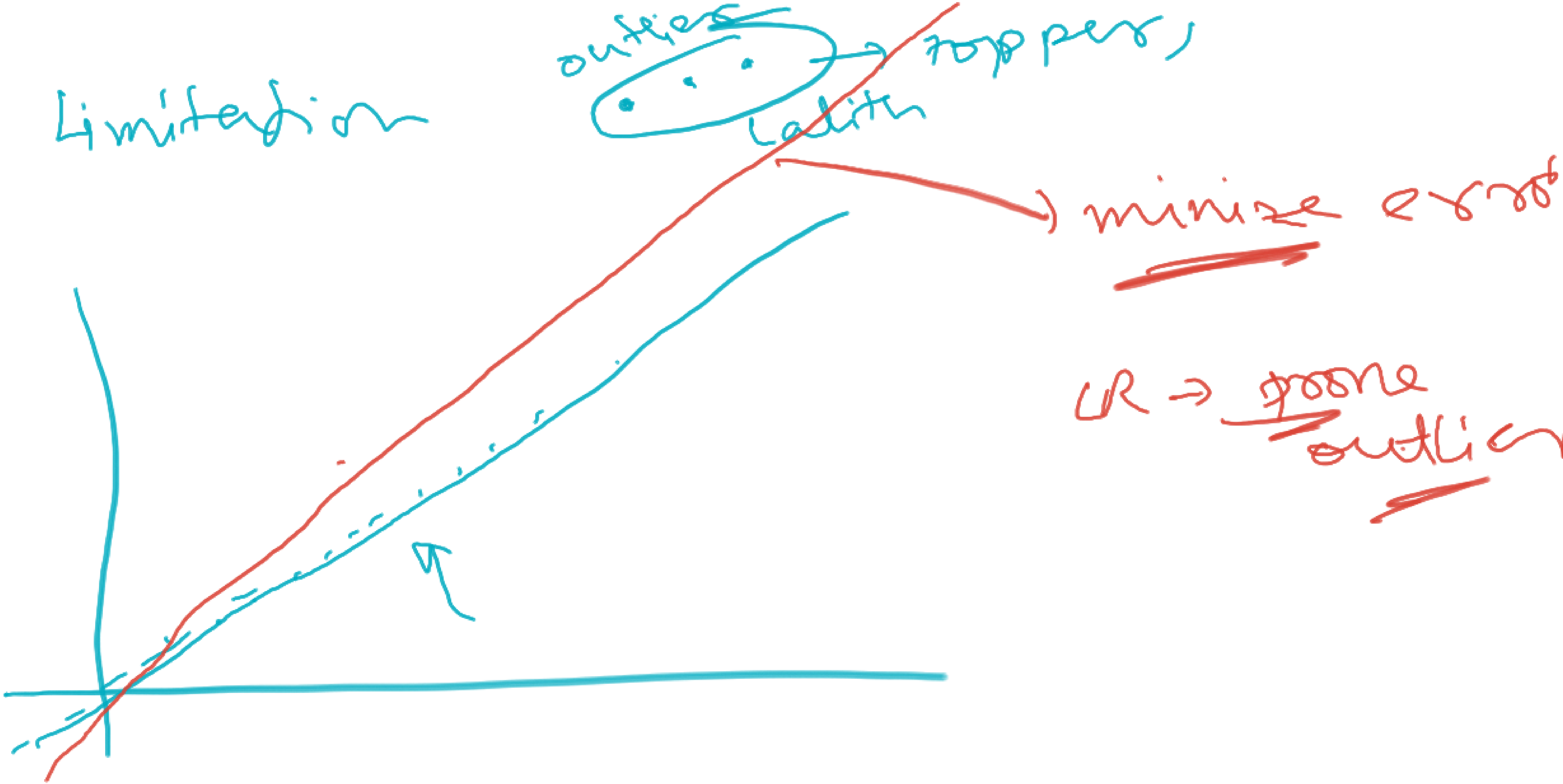
$\hat{y} \leftarrow x_1 \mid x_2$

y lin x

EDA

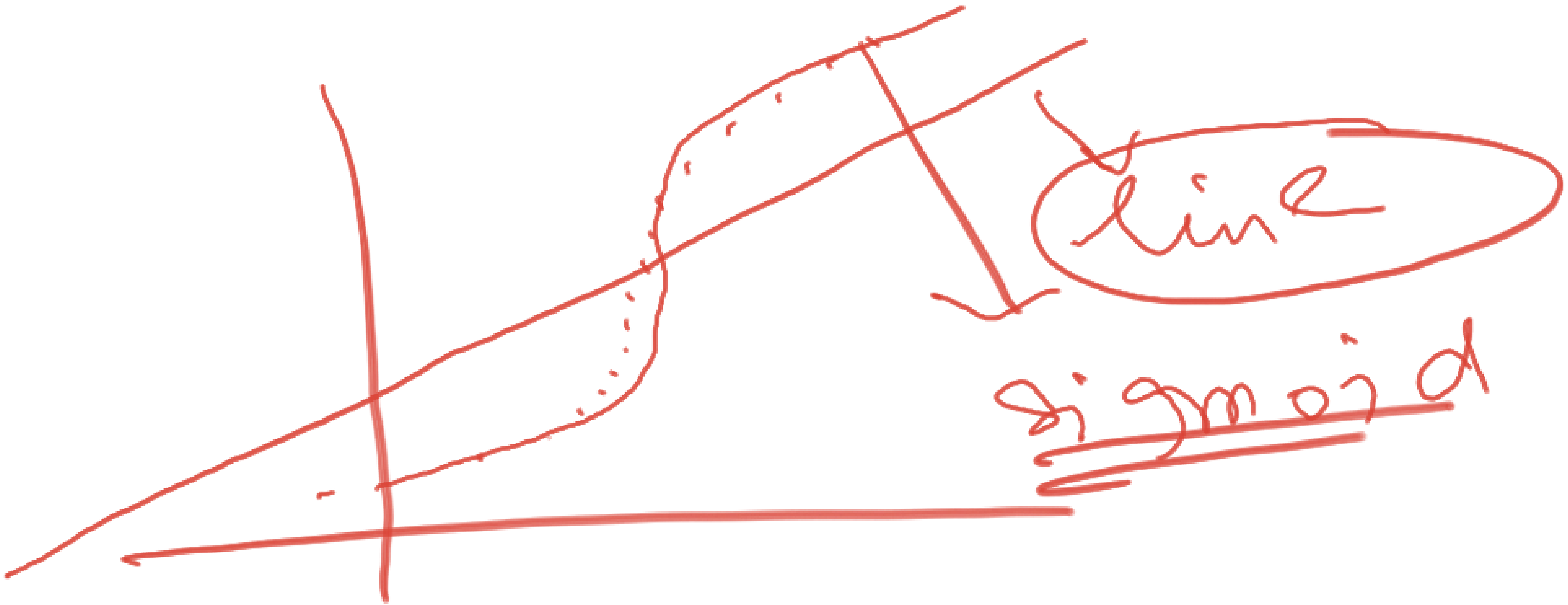


Limitation

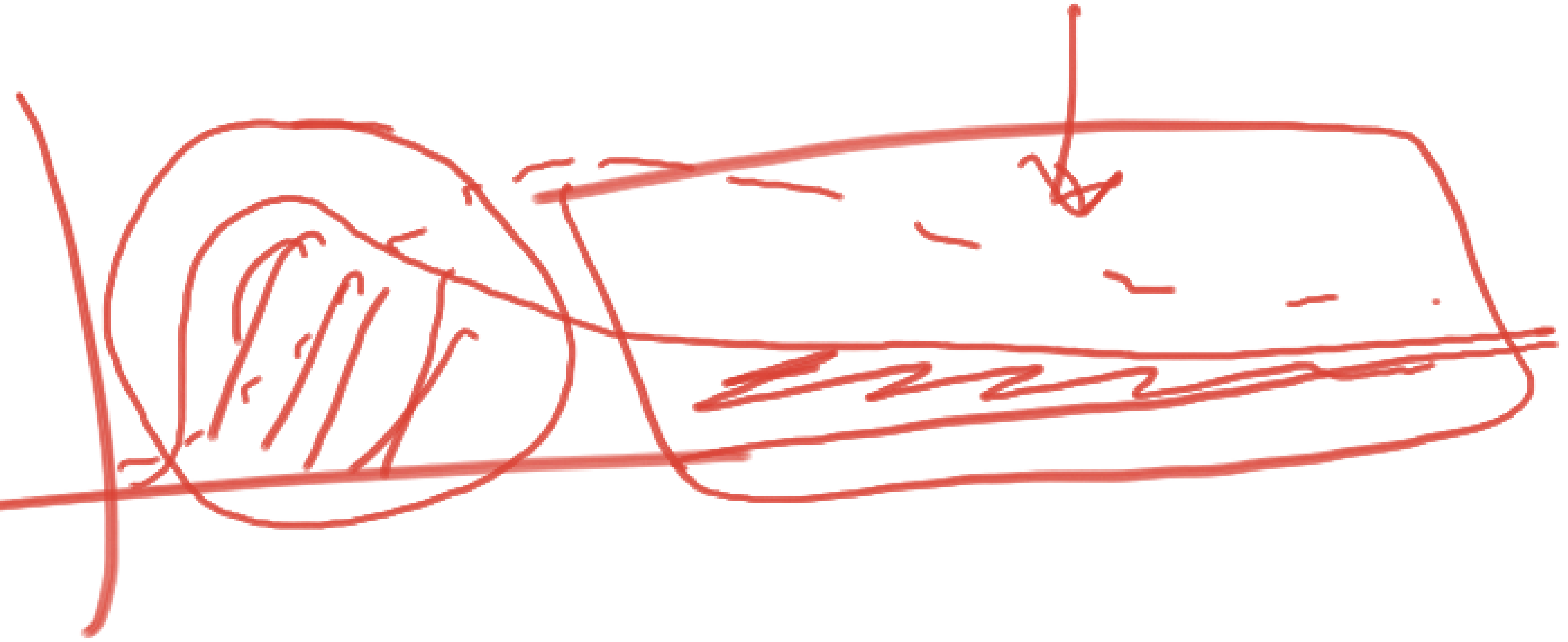


LR → prone outlier

underfitting, small → adapt



residual → normally data



$0.25 \xleftrightarrow{25\times} 5$

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

age

$25 \xleftrightarrow{25\times} 1$

$y = \beta_1 \longleftrightarrow \beta_2$

$6{,}00{,}000 \longleftrightarrow 30$

(salary) age

$\downarrow \beta_1 \longrightarrow \beta_2$

Evalution Metric model

MSE , MAPE

RMSE

R^2 adj R^2

MAE

test → RS , NCERT

evalu

train → model

%, question

d since

$\frac{1}{n} \frac{\Sigma (y-\hat{y})^2}{5}$

$(y-\hat{y})^2$

$(y-\hat{y})^2$

mean

$$R^2 \rightarrow 1 - \frac{MSE(model)}{MSE(no\text{-}brain)}$$

mod

y	$\hat{y}$
y_1	y_{mean}
y_2	y_{mean}
y_3	y_{mean}
y_4	y_{mean}
y_5	y_{mean}

MSE model

$MSE_{NB} \rightarrow$ better

$MSE_m \lll MSE_{NB}$

x y baseline MSE.Base mse

P 1,00,000 → 87,000

Y 75K 87,000

R 70K 87K $1 - \dfrac{MSE_{model}}{MSE_{baseline}}$

L 90K 87K

A 95K R^2 $1 - 0 \to 1$

assum

model ⇓ $R_1 \to \dfrac{MSE_m}{MS_b}$

10	8	2
9	8	1
8	8	0
7	8	-1
6	8	-2

Adj

2 vars

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$$

$y - \hat{y}$

MSE

$R^2 = 1 - \frac{MSE_{model}}{MSE_{basel}}$ ↓

$B_E = y - \hat{y}'$

$B = (y - \hat{y}) + (\hat{y} - \hat{y}')$

MSE MSE − MSE ↓

1000 → 10 990

0.9

0.92

$R^2 \uparrow$

0.92

0.9

lot of info

MSE

RMSE

$(y-\hat{y})^2 \longleftrightarrow$ penalize deviation $\uparrow$

$(5-2)^2 \leftarrow 9$ more error

$(5-1)^2 \leftarrow 16$

penalize deviation

unit MSE $\rightarrow Rs^2$

RMSE = Rs

avg salary

\$100K

\$10 wrong

\$99K

MAE MAPE

penalize → dev

avg

$\uparrow\downarrow$

SEM

Diff $\sqrt{\Sigma(y-\hat{y})^2}$

R^2 → goodness of fit

tm

$\frac{MSE_m}{MSE_b}$

adj → model complex