Zoom Workplace Meeting View Edit Window Help

Boosting — Sequential

— Additive Model

— Boosting reduces Bias

Stage 0 — Predicts y

$D^{1000} \rightarrow \{x_i, y_i\}_{i=1}^{n}$ $\hat{y} = f_0(x)$ $\boxed{y = f(x)}$

$E_0 = y - \hat{y}$

$= y - f_0(x)$

Base Models

High Bias + Low Variance

Weak Learners

Stage 1 — Predicts E_0

$\{x_i, E_{0i}\}_{i=1}^{n}$ $\hat{E}_0 = f_1(x)$

$E_1 = E_0 - \hat{E}_0$

$= y - f_0(x) - f_1(x)$

Stage 2 - predicts E_1

$\{x_i, E_{1i}\}_{i=1}^{n}$ $\hat{E}_1 = f_2(x)$

$E_2 = E_1 - \hat{E}_1$

$= y - f_0(x) - f_1(x) - f_2(x)$

$y = f_0(x) + f_1(x) + f_2(x) +$

Stage 0

x_1 Area	x_2 Bedrooms	y Price	$\hat{y}$ $f_0(x)$	E_0
1500	2	50	32	18
1800	2	70	58	12
2200	3	83	80	7
2800	3	98	89	9

Stage 2

Area	Bedrom	E_1	$\hat{E}_1$ $f_2(x)$	E_2
1500	2	9	3	6
1800	2	6	2	4
2200	3	5	2	3
2800	3	6	3	3

$y = 32 + 9 + 3$

$= 44$

Stage 1

Area	Bedrooms	E_0	$\hat{E}_0$ $f(x)$	E_1
1500	2	18	9	9
1800	2	12	6	6
2200	3	7	2	5
2800	3	9	3	6

53

→ $\boxed{50}$ → 49.9

→ 43

→ 41.8

→ 42.5

→ 43

→ 41

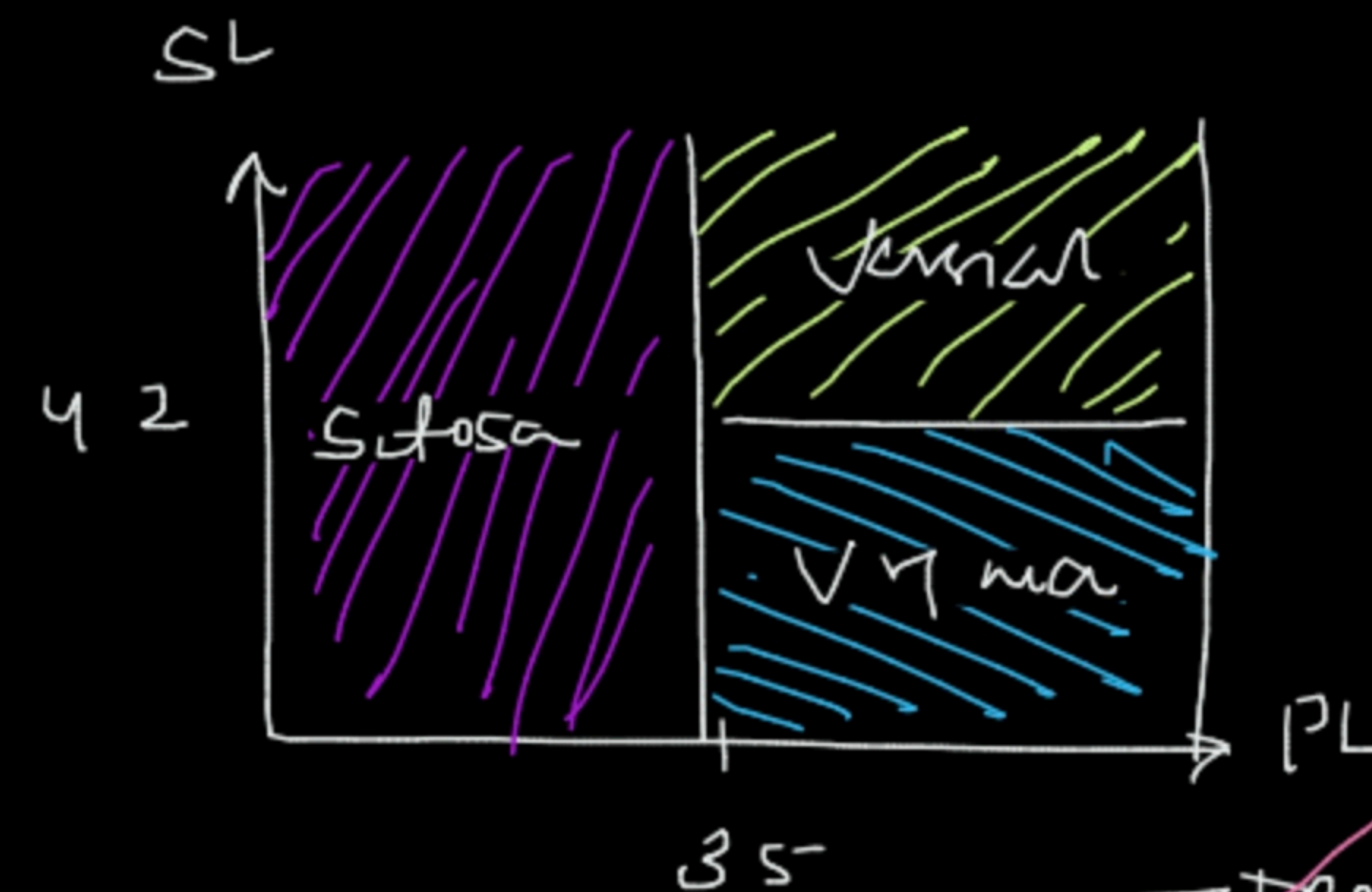
→ 32 Rough prediction of y

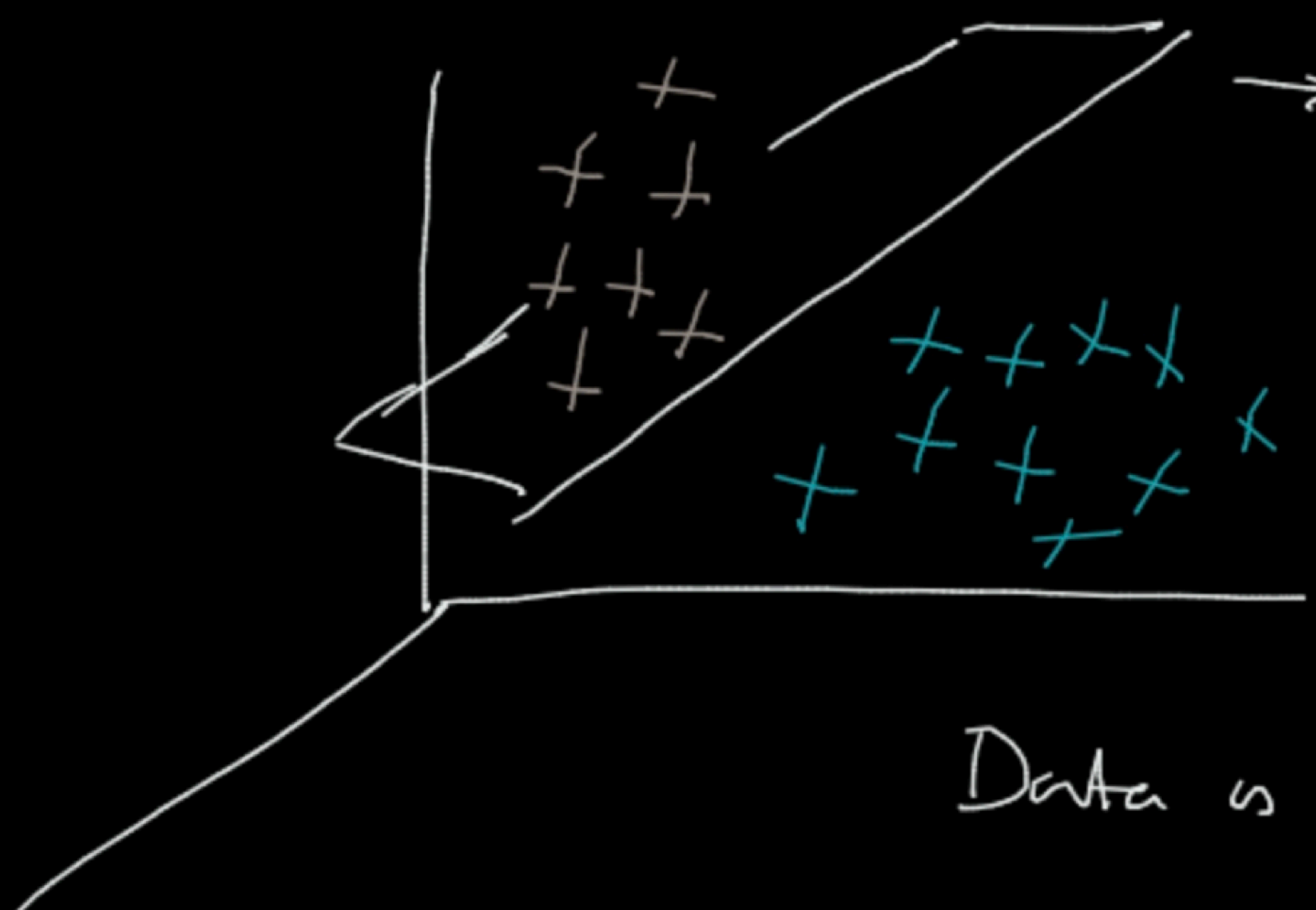
Classification

Partition feature space into pure regions assigned to each class

PL < 3 5
    Setosa
Else
    PL > 4 2
        Versicolor
    else
        Virginica

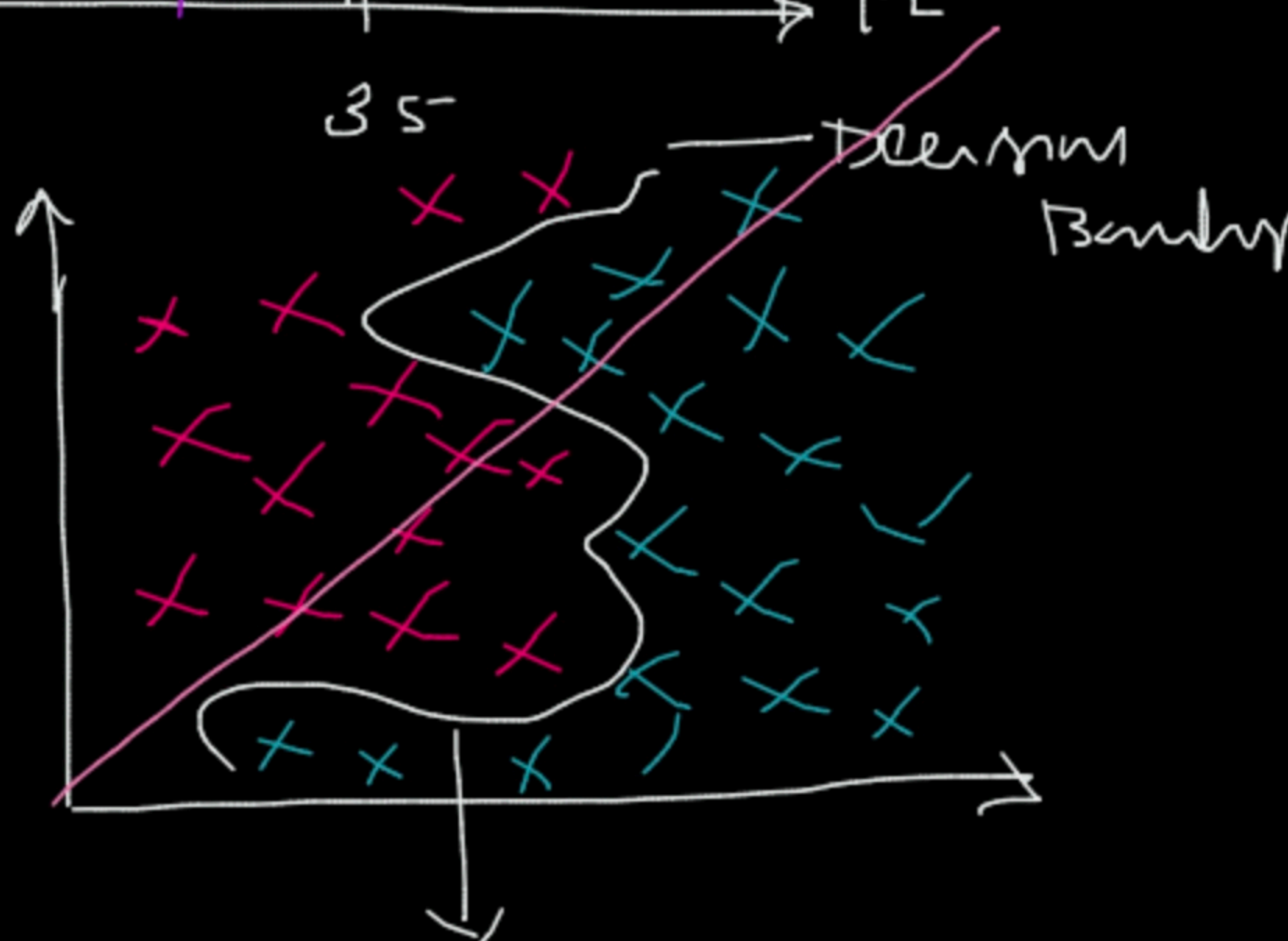
D Tree → Axes parallel





→ Decision Boundary

Data is linearly separable



Non linear

0.8
$\uparrow$
$M_1 \quad M_2 \quad M_3 \quad M_4$
$\downarrow$
0.02

1 Assigns weightage to models based on their performance

2 Every Data point is given a weightage

		x_1	x_2	y	
0.13 ↓	0.2	—	—	0	0
0.3 ↑	0.2	—	—	1	0 —
0.13 ↓	0.2	—	—	1	1
0.13 ↓	0.2	—	—	0	0
0.3 ↑	0.2	—	—	1	0
1.0	1.0				

$y \to f_0(x) + f_2(x) -$

Pip install xgboost

XGBoost

Extreme ← | ↓ | └ Boosting

— parallelize

— CS

Gradient → Focuses on error from prev step

$y - \hat{y}$

$\frac{dy}{dx} \rightarrow \frac{\Delta y}{\Delta x}$

Loss Fn → Fn of error

→ Squared Error

$y = x^2$

$= 2x^1$

$$\text{Loss Fn}(L) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

"We can customize the loss fn provided it is differentiable"

↑

$$\text{Loss}(L) = (y - \hat{y})^2$$

$$\frac{\partial L}{\partial \hat{y}} = \text{Partial differential}$$

$$\partial L / \partial \hat{y} = -2(y - \hat{y})$$

Pseudo residual → $-\frac{1}{2}\, \partial L / \partial \hat{y} = (y - \hat{y})$ ↳ residual

X Dosage	Y Efficiency	$\hat{y}$	E_0	$\hat{y}$	E_1	$\hat{y}$	E_2 ↗ pseudo
15 → 12	−10	0.5	−10.5	−2.65	−7.35	−2.88	−7.12
→ 18	7	0.5	6.5	+2.6	4.4		
30 27	8	0.5	7.5	+2.6	5.4		
33	−7	0.5	−7.5	−1.75	−5.25		

$\alpha_1 = 0.3$

Gain = Left Sim + Right Sim − Parent Sim

$= 110.25 + 14.08 - 4$

$= 120.33$

$f_0(x) = 0.5$

$f_1(x) \rightarrow \hat{E_0}$

$$Sim = \frac{(\Sigma \text{ residual})^2}{\text{no of res} + \lambda} \quad \lambda = 0$$

$\alpha_2 = 0.5$

$\boxed{-10.5 \quad 6.5 \quad 7.5 \quad -7.5}$ $\quad Sim = \frac{(-4)^2}{4} = 4.0$

Dosage < 15 (yes / No)

yes: $\boxed{-10.5}$

$Sim = (-10.5)^2$

$= 110.25$

$0.3 \times (-10.5)$

-3.15

No: $\boxed{6.5 \quad 7.5 \quad -7.5}$ $\quad Sim = \frac{(6.5)^2}{3} = 14.08$

Dosage < 30 (yes / No)

yes: $\boxed{6.5, 7.5}$ $\quad +7.0$

No: $\boxed{-7.5}$

$\hat{y} = 0.5 + 0.3(-7.5)$

$= 0.5 - 2.25$

$\hat{y} = 0.5 + 0.3(7.0)$

$= +2.6$

$\hat{y} = f_0(x) + \overset{\alpha_1}{f_1(x)}$

$= 0.5 + (-3.15)$

$\boxed{-7.35 \quad 4.4 \quad 5.4 \quad -5.25}$

Dosage < 23

$\boxed{-7.35, 4.4}$ $\quad -1.47$

$\boxed{5.4, -5.25}$ $\quad 0.075$

$\hat{y} = 0.5 + (-2.65) + 0.5(-0.1)$

$= 0.5 + (-2.65) + 0.5(-1.47)$

$= -2.88$

$$\hat{y} = f_0(x) + \alpha_1 f_1(x) + \alpha_2 f_2(x) +$$

$y - \hat{y} \rightarrow$ pseudo residual