MATH 462

1. Biz . Office Hours { Thurs 2 pm Gabriela zoom
   Weds 4p - 4:45 }

   Notes & HW

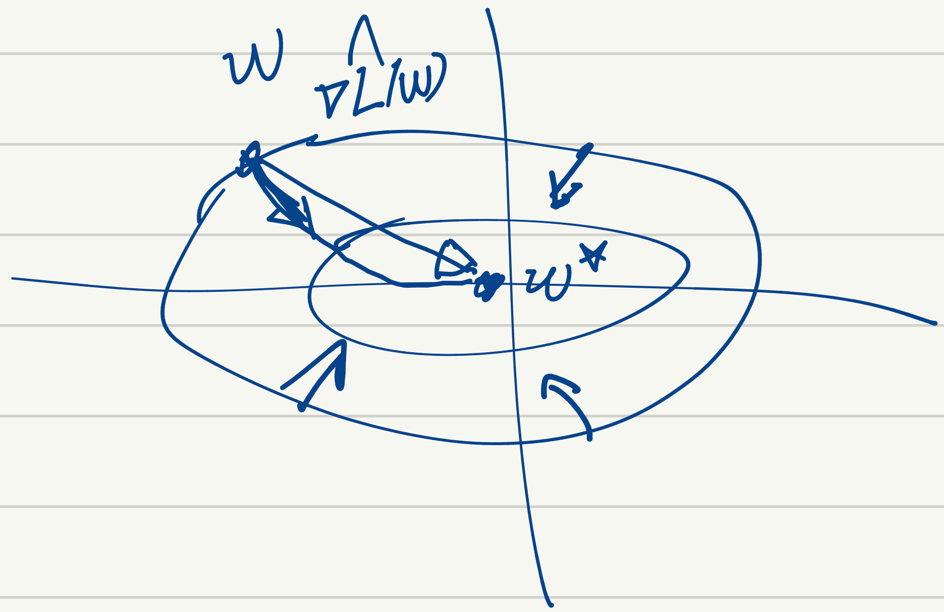
2. Review

3. Losses.

Gradient $\quad w \in \mathbb{R}^d$

$\hat{L}(w)$ convex

$$\nabla \hat{L}(w) = \begin{bmatrix} \frac{\partial L}{\partial w_1} \\ \vdots \\ \frac{\partial L}{\partial w_d} \end{bmatrix}$$



$$\boxed{\nabla L(w) = 0}$$

e.g. $\hat{L}(w) = \left\| \begin{pmatrix} 3w_1 \\ 4w_2 \end{pmatrix} - \begin{pmatrix} 6 \\ 7 \end{pmatrix} \right\|^2 / 2 = \frac{(3w_1 - 6)^2 + (4w_2 - 7)^2}{2}$

$$\frac{\partial \hat{L}(w)}{\partial w_1} = \frac{2}{2}(3w_1 - 6) \cdot 3 = 9w_1 - 18$$

$$\frac{\partial \hat{L}(w)}{\partial w_2} = \frac{2}{2}(4w_2 - 7)4 = 16w_2 - 28$$

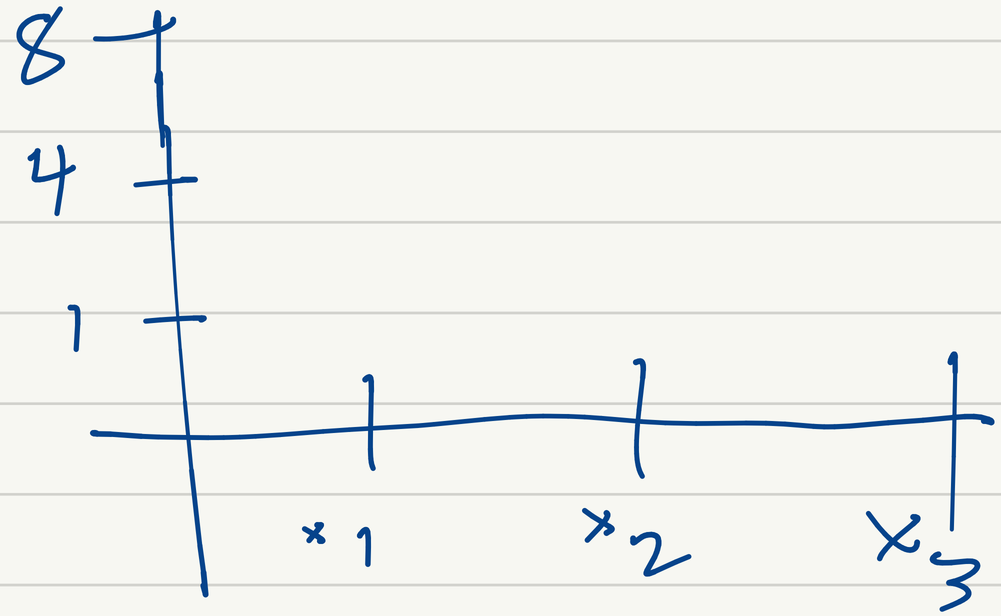
$$\nabla L(w) = \begin{bmatrix} 9w_1 - 18 \\ 16w_2 - 28 \end{bmatrix}$$

$$0 = \nabla L(w) \Rightarrow \begin{matrix} 9w_1 = 18 \\ 16w_2 = 28 \end{matrix} \Rightarrow \begin{pmatrix} w_1 = 2 \\ w_2 = 7/4 \end{pmatrix}$$

Understanding Losses.

How to they balance errors?

- Huber loss is less sensitive to outlier
- Analyze loss.



model $w \in \mathbb{R}$

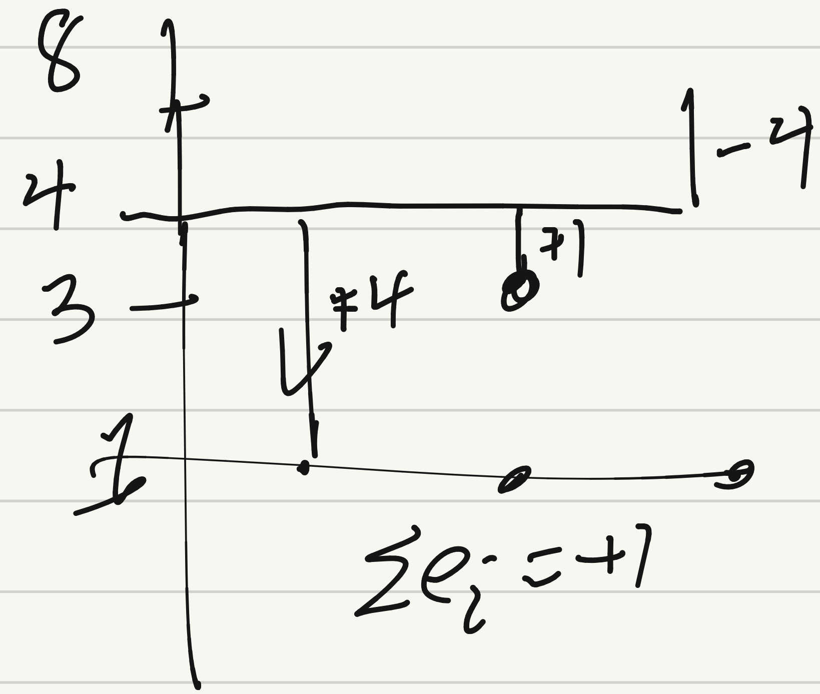
data $\{y_1, y_2, y_3, \dots, y_m\}$

$y_i \in \mathbb{R}$

ELM $\min_{w} \frac{1}{m} \sum_{i=1}^{m} \ell(e_i = w - y_i)$

for ℓ_2 ℓ_1 ℓ_H

$$\min_{w} \hat{L}(w) = \frac{1}{m} \sum_{i=1}^{m} \ell(e_i = w - y_i)$$



1. $$\hat{L}'(w) = \frac{1}{m} \sum_{i=1}^{m} \ell'(e_i)$$

$e_i = w - y_i$

CASE $\ell(e_i) = e_i^2/2$ $\quad \ell'(e_i) = e_i$

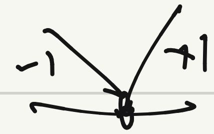
$$0 = \frac{1}{m} \sum_{i=1}^{m} e_i = \frac{1}{m} \sum_{i=1}^{m} (w - y_i)$$

$w = 3.5$

$e: 2.5, .5, -4.5$

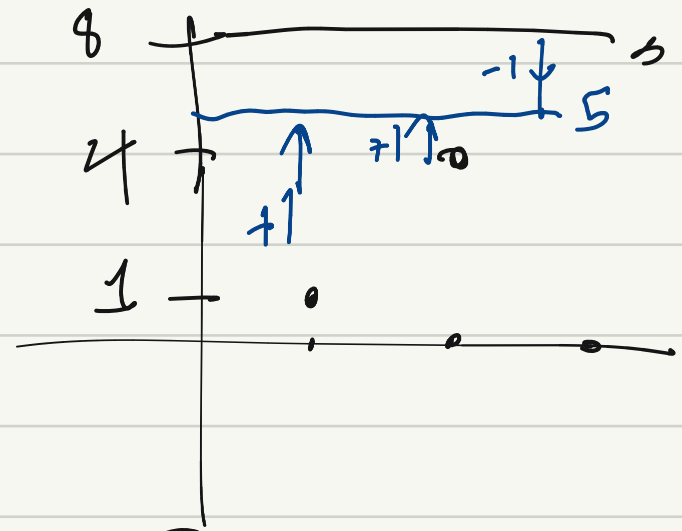
$$\Rightarrow \underline{w = \frac{1}{m} \sum_{i=1}^{m} y_i}$$

Loss $\ell(e) = |e|$



$\hat{L}(w) = \frac{1}{m} \mathbf{1} \cdot \ell(e)$

$\hat{L}'(w) = \frac{1}{m} \sum_{i=1}^{m} \ell'(e_i)$



$w = 8$

$\ell(e) = (\ell(e_1), \dots, \ell(e_m))$

$\ell'(e) = (\ell'(e_1), \dots, \ell'(e_m))$

$\ell'(e) = \begin{cases} -1 & e < 0 \\ +1 & e > 0 \\ \text{undef} & e = 0 \end{cases}$

$e = (8-1, 8-4, 8-8)$
$= (7, 4, 0)$

$L(e) = \|e\| = 11$

$w = 4$ $e = (4-1, 4-4, 4-8)$
$= (3, 0, -4)$
$\ell'(e) = (+1, ?, -1)$

$e = (5-1, 5-4, 5-8)$
$(4, 1, -3)$
$\ell'(e) = (+1, +1, -1)$

check $w^* = 4$ $\hat{L}(4) = 3 + 0 + 4 = 7$

prove median is soln for ℓ' loss. $m = \text{odd}$

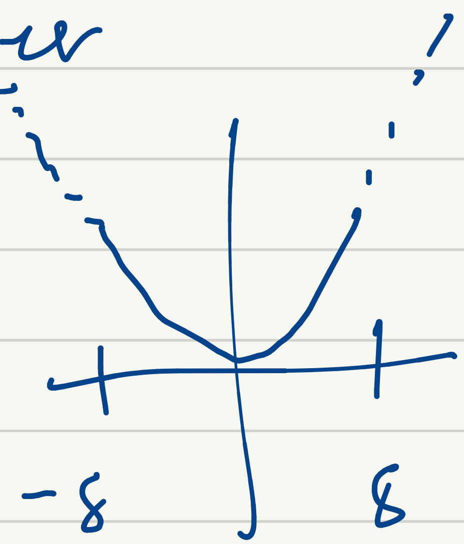
sort y_i in order. $w^* =$ "middle" value

average of middle two. $m = \text{even}$

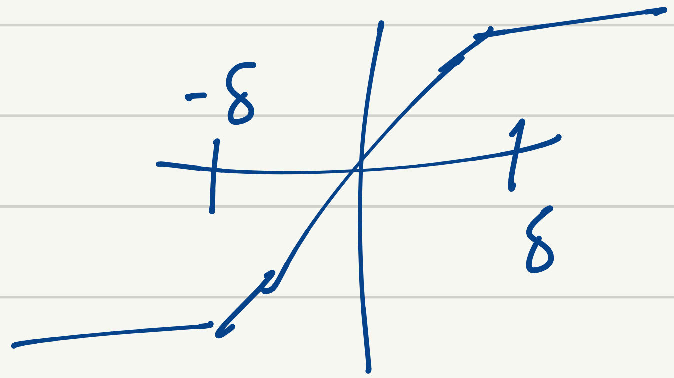
EX $y_1, y_2 = (1, 3)$ 1.5
$1 \,.\, 2 \,.\, 3 \,.$

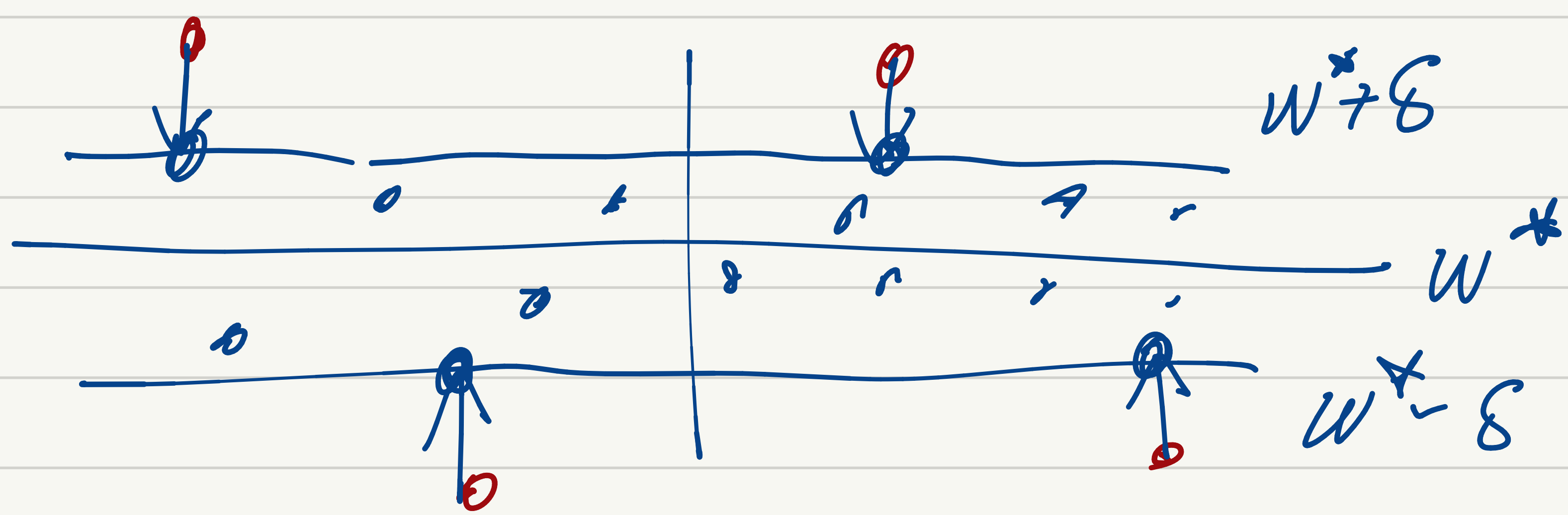
Huber

$$h_\delta(e) = \begin{cases} \frac{e^2}{2}, & |e| < \delta \\ \delta\left(|e| - \frac{\delta}{2}\right), & |e| \ge \delta \end{cases}$$



$$h'_\delta(e) = \begin{cases} e, & |e| < \delta \\ \delta, & e > \delta \\ -\delta & e < -\delta \end{cases}$$

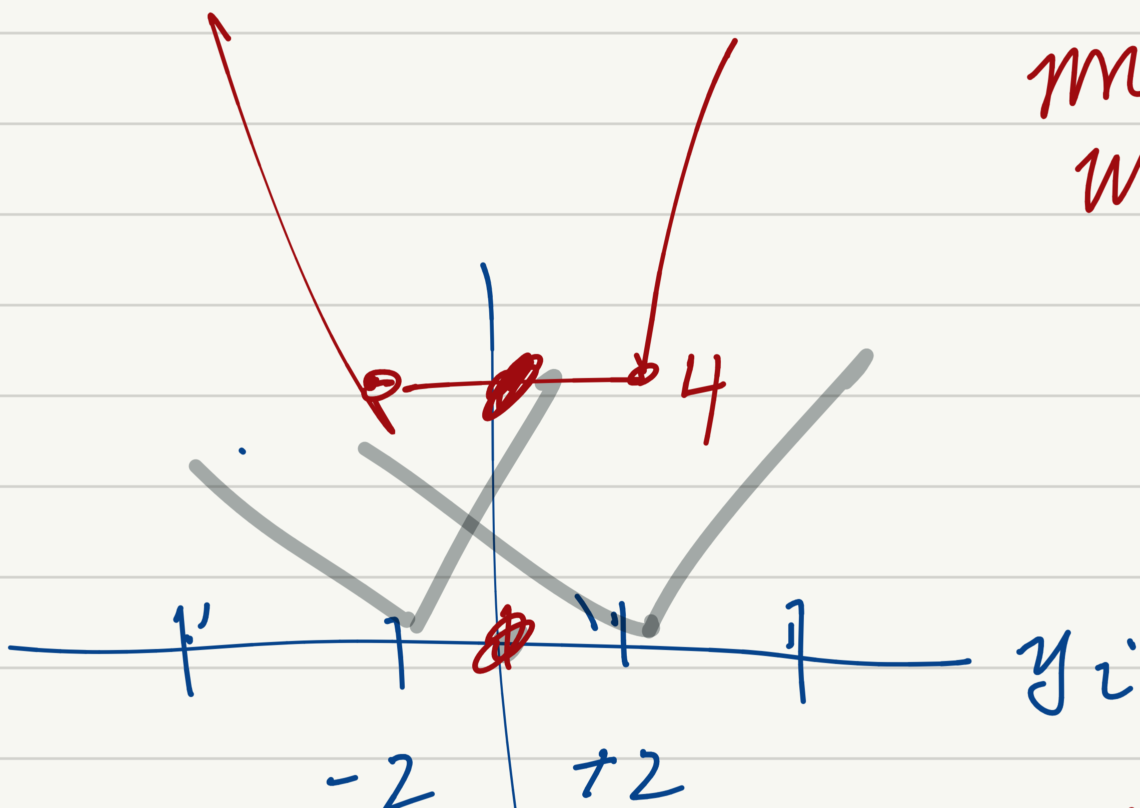




① ℓ^1 loss

4 data points

$L(w)$



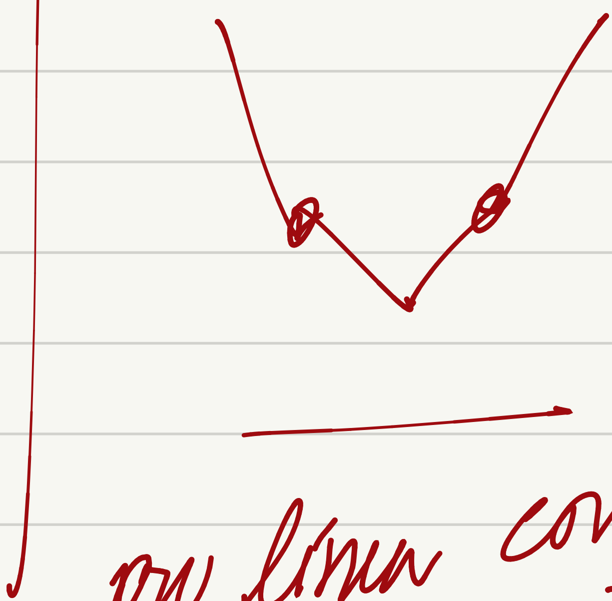
$\min_{w} L'(w)$ = any $w \in [-2, 2]$

median average of middle

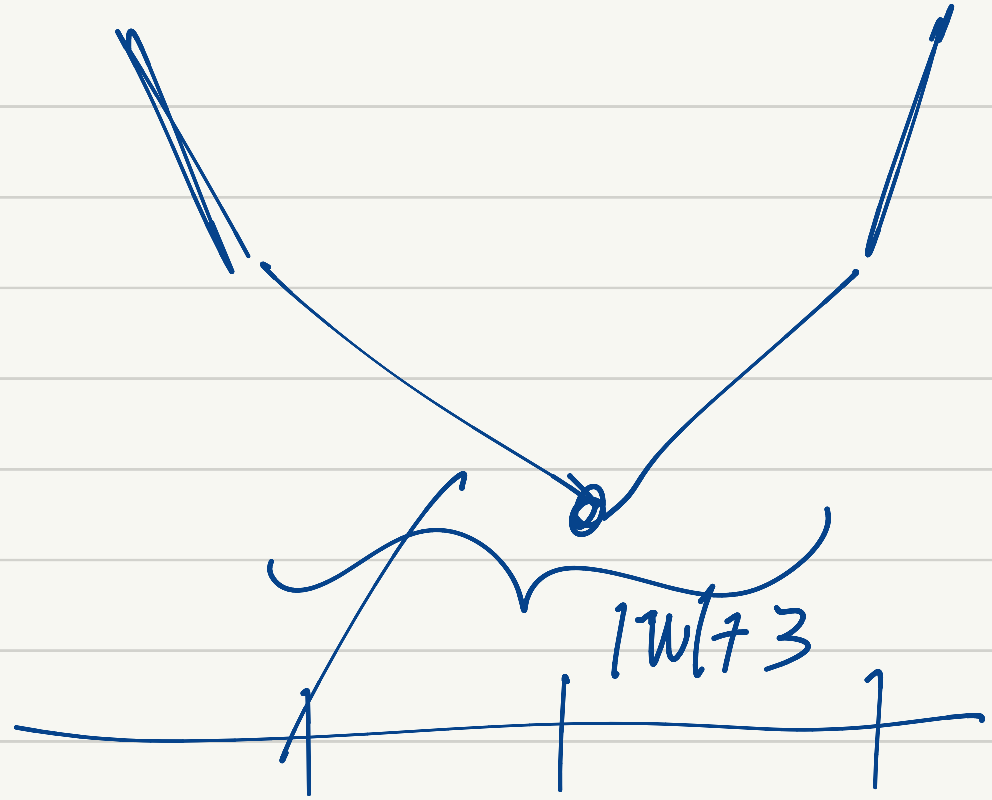
$$|x+2| + |x-2| = \begin{cases} 4-2x & x<-2 \\ 4 & |x| \le 2 \\ 4+2x & x>2 \end{cases}$$

e.g. $x+2+2-x=4$

3 values



pw linear corner at median



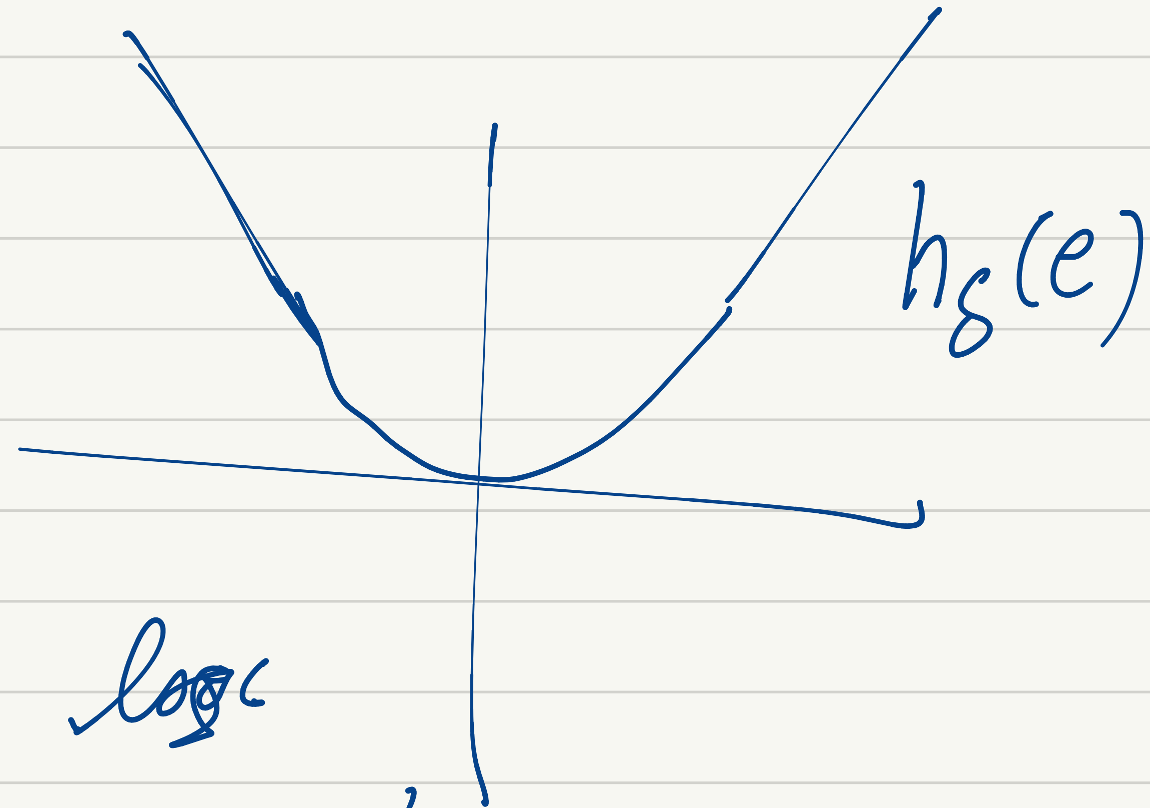
$-1 \qquad 0 \qquad +2$

$$L(w) = |w+1| + |w| + |w-2|$$

$$= w+1 + |w| + 2 - w = 3 + |w|$$

$$-1 \le w < 2$$

w^* - middle value



$\delta \to \infty$ $\quad l_2$ loss

everything inlier.

or $\quad$ all data $\in [-\delta, \delta]$

$\delta \to 0 \iff$ all data outlier

$$\frac{1}{\delta} h_\delta(e) = \frac{\delta(|e| - \frac{\delta}{2})}{\delta} = \frac{\delta|e| - \frac{\delta^2}{2}}{\delta}$$

$\iff \delta \cdot l_1$ loss