MATH 462 Lect 2

Refer to posted notes.

SS Ch 9 Linear Predictors.

Data $x \in \mathcal{X}$

$\hookrightarrow$ $x \in \mathbb{R}^d$ $\quad$ d dim of data

$x \xrightarrow{f} f(x) \in \mathbb{R}^d$ $\quad$ f_i features

$x \xrightarrow{\psi} \psi(x) \in \mathbb{R}^d$

Linear Model

$$y(x) = h_w(x) = w \cdot x = \sum_{i=1}^{d} w_i x_i$$

Affine Model $y(x) = h_{w,b}(x) = w \cdot x + b$

$$\tilde{x} = (x, 1) \qquad \hat{w} = (w, b) \qquad \hat{w} \cdot x = w \cdot x + b$$

features

polynomial regression.

$x \in \mathcal{X} = \mathbb{R}$

$f(x) \in \mathbb{R}^d$

$f(x) = (1, x, x^2, x^3) \in \mathbb{R}^4$

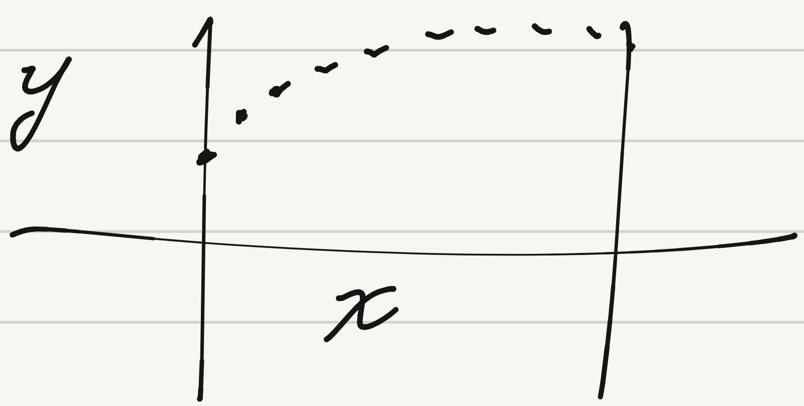
$w \cdot f(x) = w_1 + w_2 x + w_3 x^2 + w_4 x^3$

Fit $S_m = \{(x_1, y_1), \dots, (x_m, y_m)\}$

with model linear in w

features nonlinear.

see notes!



$d = 4$

$f(x) = (1, x, x^2, x^3)$

$x_1 \leftarrow f(x_1) = f_1$

$x_i \quad f_i$

$2 \rightarrow (1, 2, 4, 8)$

abstraction

$$F = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}$$

Model

$$y = h_w = F \cdot w = \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix}$$

Loss $e = F \cdot w - y$

Regression $L = \cancel{e \cdot e}/2 = \|e\|/2$

$\min \frac{\|e\|^2}{2}$

$y = (y_1, \dots, y_m)$

abstract loss

$l(e)$ component-wise

$$e = \begin{bmatrix} .1 \\ .2 \\ .0 \\ .5 \end{bmatrix} \quad e^2 = \begin{bmatrix} 0.01 \\ 0.04 \\ 0 \\ .25 \end{bmatrix} \quad |e| = \begin{bmatrix} .1 \\ .2 \\ 0 \\ .5 \end{bmatrix}$$

Quadratic Regression

$$e = Fw - y$$

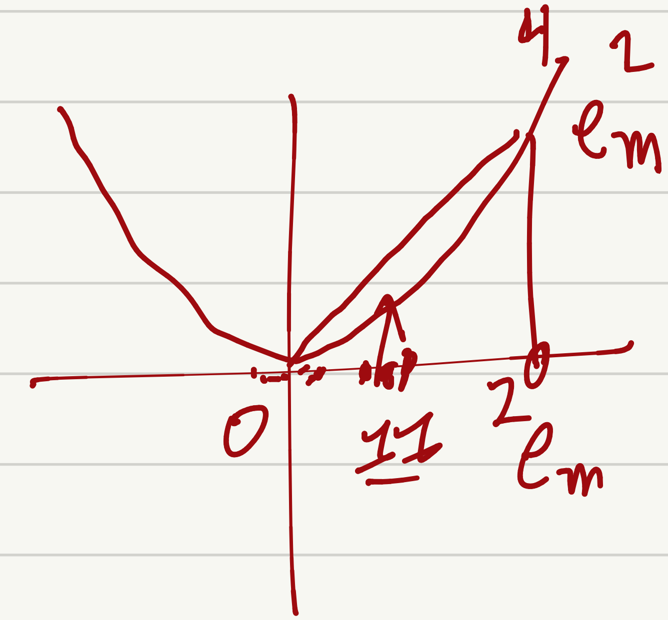
$$e^2 = \|Fw - y\|^2$$

<u>ANS</u> reem.

$$\min \|Fw - y\|^2 \xRightarrow{\text{Vector calc}} F^T(Fw - y) = 0$$

$$F = \underbrace{\begin{bmatrix} \quad & \quad \\ & \\ & \end{bmatrix}}_{d} m \qquad F^TF = {}_m\begin{bmatrix} \quad \end{bmatrix}^m \begin{bmatrix} w \end{bmatrix}_m = F^Ty = \begin{bmatrix} \quad \end{bmatrix}_m$$

$$w = \frac{\sum_{i=1}^{m} y_i x_i}{\sum_{i=1}^{m} x_i^2}$$



HW Linear regression stability.

$$S_m = \{(x_1, y_1), \dots, (x_m, y_m)\}$$

$$S'_m = \{(x_1, y_1), \dots, (x_{m-1}, y_{m-1}), (x_m, \tilde{y}_m)\}$$

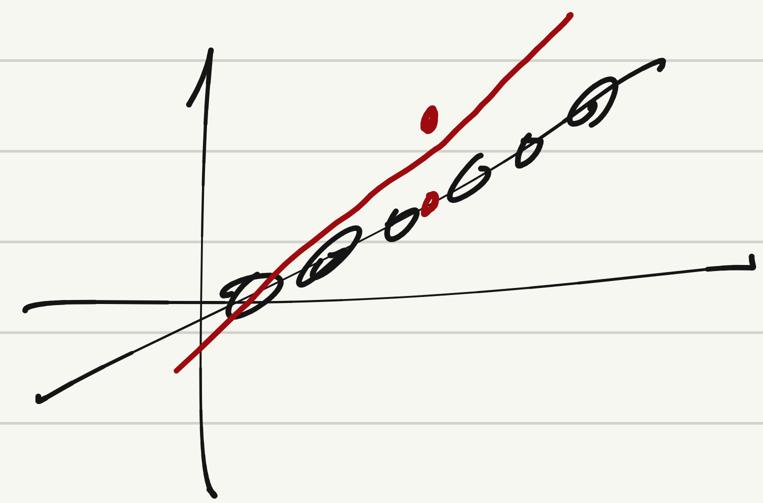
$$\tilde{y}_m = y_m + e$$

New $\tilde{w}$: changed.

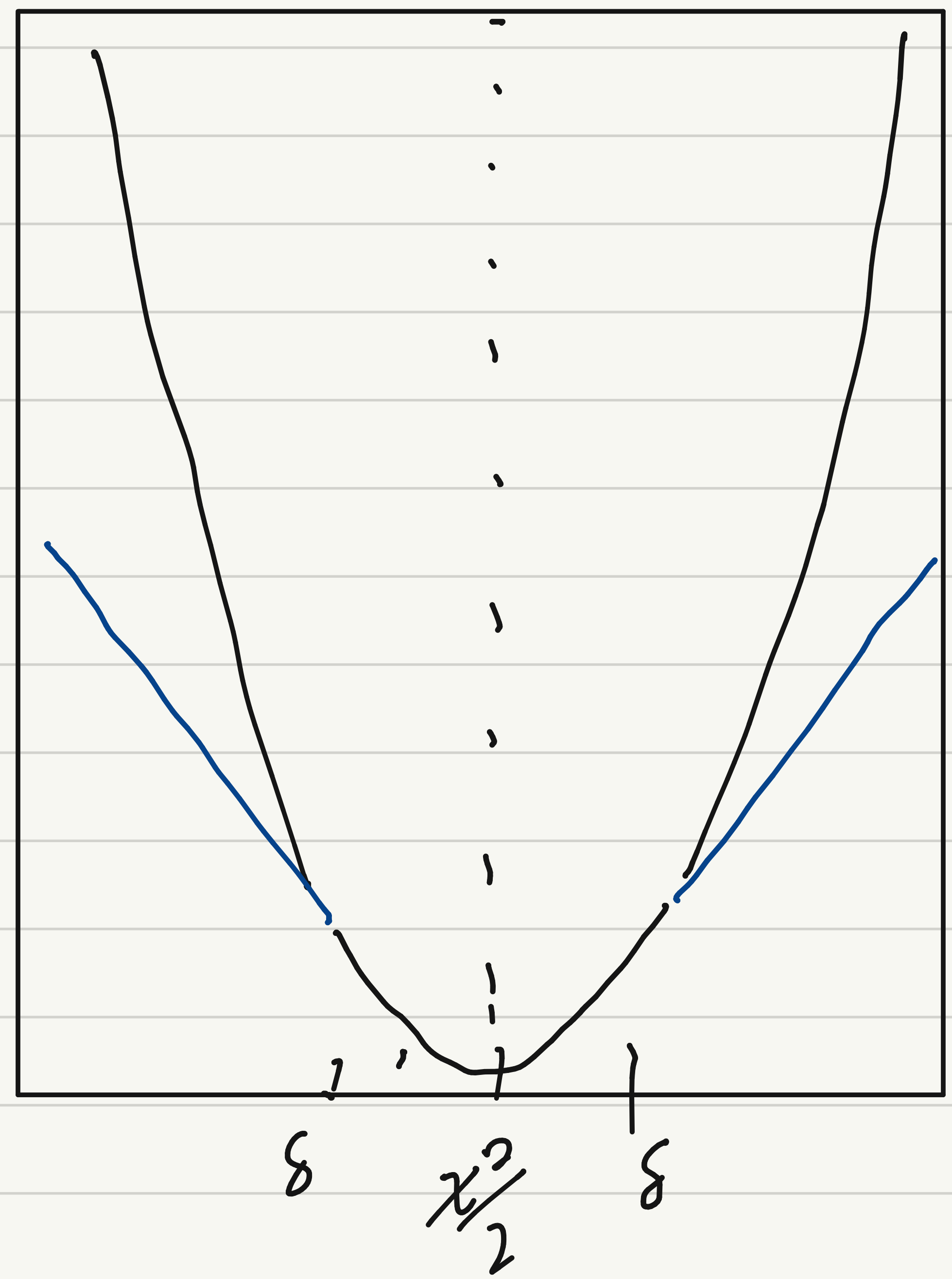
How much?

answer in terms of e:

"sensitivity to outliers"

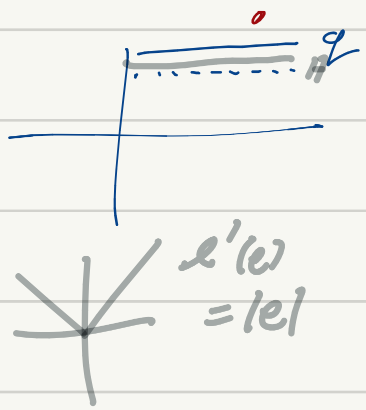


Huber Loss



Huber loss
designed to be
less sensitive to
outliers

3 losses
$e^2/2$
$|e|$
$H(e)$

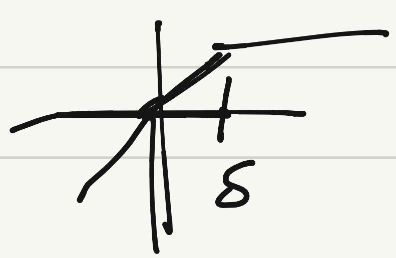


Check

cts: $t = \delta$ $\quad \delta^2/2$
$\delta(\delta/2)$ ✓

slopes $\quad h'_\delta(t) = \begin{cases} t \\ \delta \end{cases}$ $\quad t = \delta$ cts, $t = -\delta$

$$h_\delta(t) = \begin{cases} t^2/2 & |t| < \delta \\ \delta(t - \delta/2) & \text{otherwise} \end{cases}$$

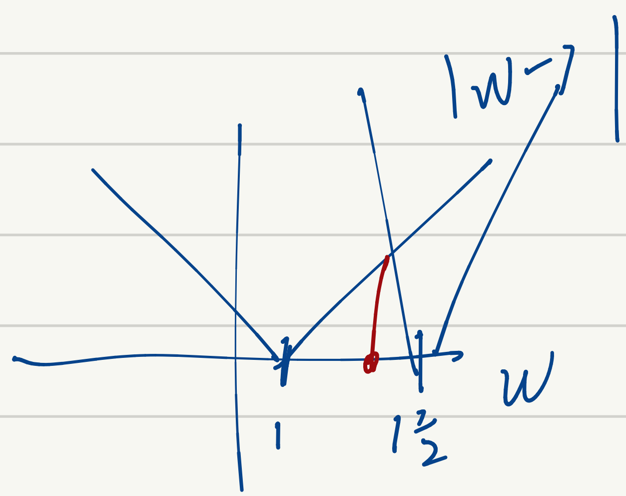


Refer to HW

EX ℓ^1 loss

$$\min_w L(w) = \left\{ \frac{1}{2}\left(|w-1|\right) + |2w-3| \right\}$$

HW



ℓ^1 loss

no formula for w^*

yes abstract soln e.g. pw. linear & convex loss

$\Rightarrow$ "Theory" $\Rightarrow$ soln.

Loss design

"soft" grading policy 5 HW.

Hard → HW grade avey of 5 HW.

soft average of top 4 grades

e_4 = average of errors top 4 $\in [0,5]$ 1/5

e_5 = error on worst $\in [0,5]$ 2/5

$$\hat{e}_5 = \begin{cases} e_4 - 1 & \text{if } e_5 > 2e_4 \\ \text{between} & \text{o.w.} \\ e_4 & \text{if } e_5 = e_4 \end{cases}$$

continue in HW