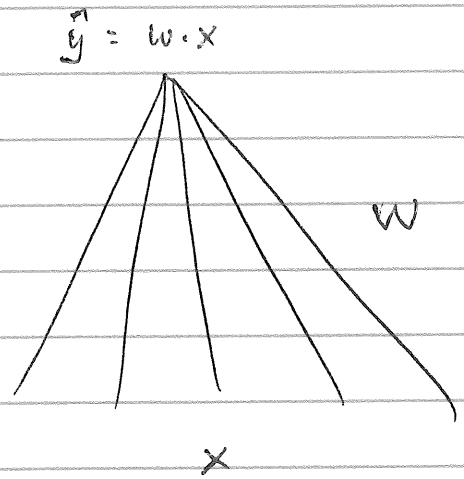
LINEAR REGRESSION

$(\vec{x}, y)$ $\in \mathbb{R}$ example



linear neuron

square Loss $(\hat{y} - y)^2 = (w \cdot x - y)^2$

$\in \mathbb{R}$ $\in \mathbb{R}$

LOGISTIC REGRESSION

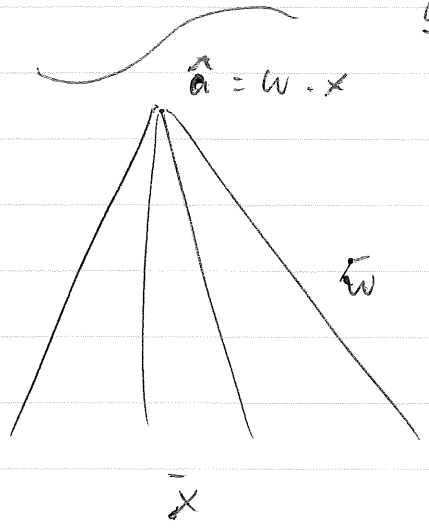
$(\bar{x}, y)$ example

$\in \{0,1\}$

or $\in [0,1]$

$\hat{y} = \sigma(w \cdot x) \in [0,1]$

"prob. of 1"



σ as transfer function

$$\sigma(x) = \frac{e^x}{1+e^x} \qquad 1-\sigma(x) = \frac{1}{1+e^x}$$

Logistic Loss

$$L(y,\hat{y}) = y \ln \frac{y}{\sigma(w \cdot x)} + (1-y) \ln \frac{1-y}{1-\sigma(w \cdot x)}$$

when y binary:

$$L(0,\hat{y}) = -\ln 1-\sigma(w \cdot x)$$

$$L(1,\hat{y}) = -\ln \sigma(w \cdot x)$$

"soft classification because $\hat{y}$ is probability

$$L(y,\hat{y}) = y \ln y + (1-y) \ln (1-y) - y\, w \cdot x + \ln(1+e^{w \cdot x})$$