Linear Classification

Wednesday, Feb17

$y \in \mathbb{R}$

$y \in \{\text{red}, \text{blue}, \text{green}\}$

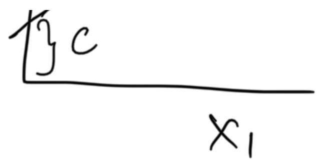
$$w^T x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix}$$

$$\boxed{\begin{aligned} & w_0 + w^T x \\ & = w_0 + \sum_{j=1}^{d} w_j x_j \end{aligned}}$$

$\boxed{y = mx + c}$

$x_2 = m x_1 + c$



$$\boxed{w_1x_1 + w_2x_2 + w_0 = 0}$$

$$x_2 = -\frac{w_1}{w_2}x_1 - \frac{w_0}{w_2}$$

$$\frac{w_1}{2}x_1 + \frac{w_2}{2}x_2 + \frac{w_0}{2} = 0$$

$$w = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \qquad w_0 = 1$$

$$\boxed{4x_1 + 2x_2 + 1 = 0}$$

If $x_1, x_2 = (0, -\frac{1}{2})$

$$4 * 0 + 2 * (-\tfrac{1}{2}) + 1$$

$$= 0$$

$\mathbf{w}^T\mathbf{x} + w_0$ will be 0

if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ ~~is $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$~~ lies on the line.

$x_1, x_2 = (2, 2)$

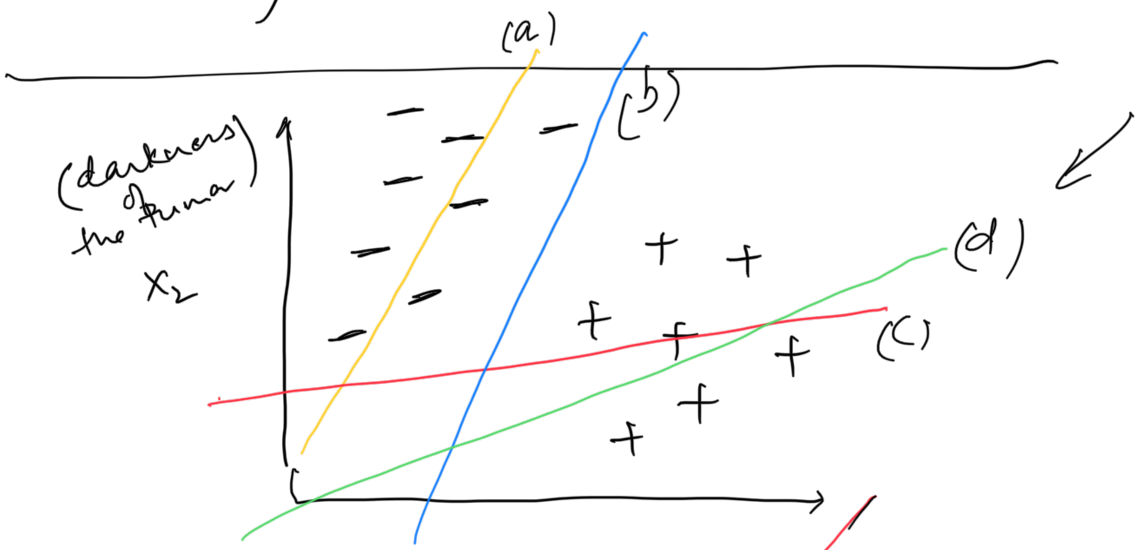
$$w^Tx + w_0 = 8 + 4 + 1 = 13$$

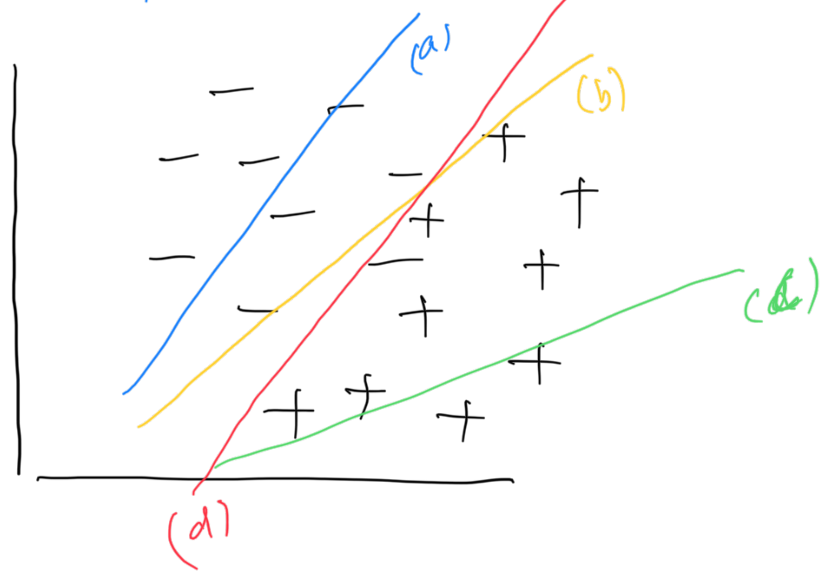
$x_1, x_2 = (-4, 2)$

$$w^Tx + w_0 = -8 + 4 + 1 = -3$$

A line can be used as a decision

boundary (or a decision rule)



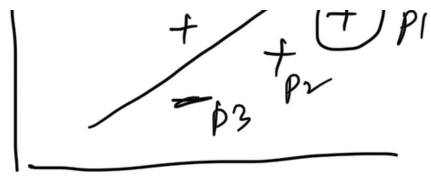


$$\mathbb{I}(\text{true}) = 1$$

$$\mathbb{I}(\text{false}) = 0$$

for p1 : (x_1, y_1)

$y_1(w^Tx_1 + w_0) > 0$



For p4 : $(x_4, \overset{-1}{y_4})$

$\underline{w^Tx_4 + w_0} < 0$

But $y_4(w^Tx_4 + w_0) > 0$

for p3 : $(x_3, \underline{\overset{-1}{y_3}})$

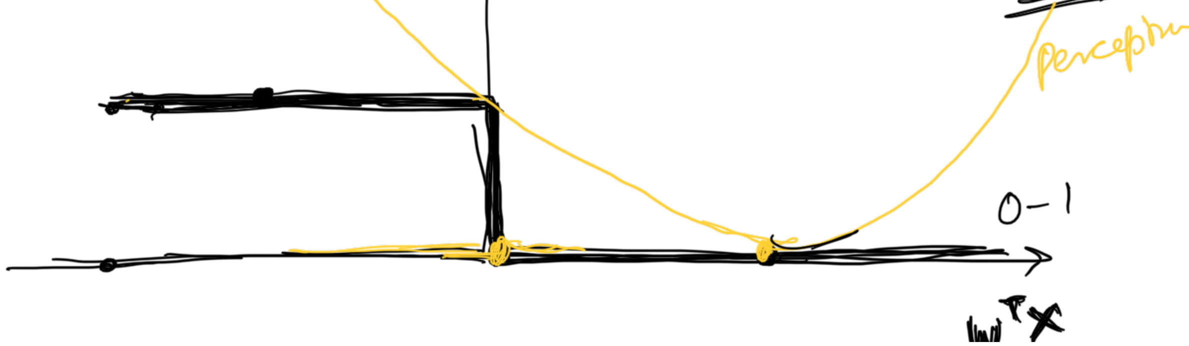
$\underline{w^Tx_3 + w_0} > 0$

But $y_3(w^Tx_3 + w_0) < 0$

$$\boxed{\begin{aligned} \underline{\mathbf{w}}^T x + w_0 \geq 0 &\rightarrow y = +1 \\ \text{else } y &= -1 \end{aligned}}$$

Consider a dataset with only one example, <x, y>

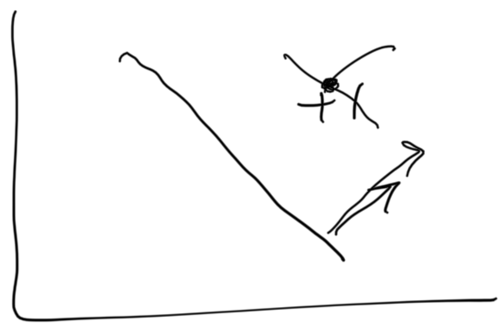
let $\underline{y = +1}$

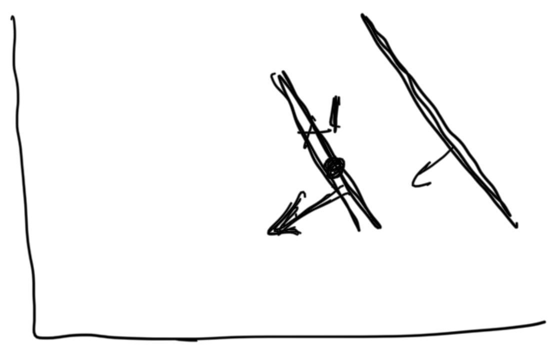


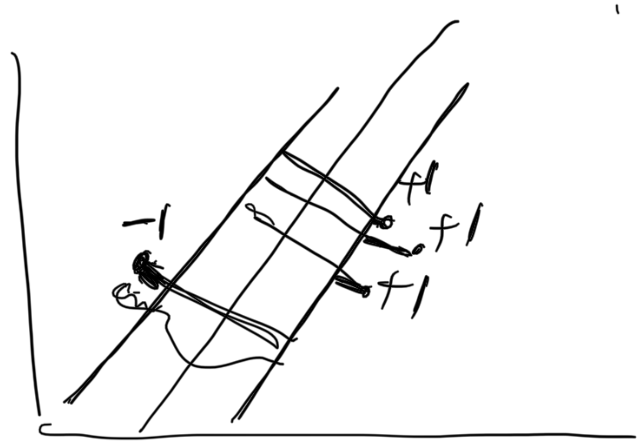
for same w: $w^Tx = -8$

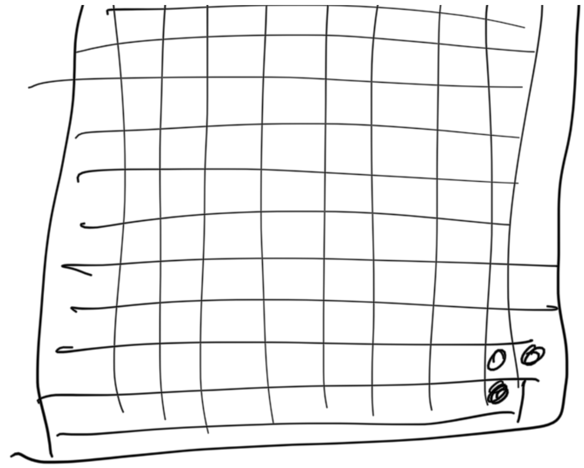
For perceptron: $\frac{1}{2}(- w^Tx)^2$

Fri Feb 19









$$W = W - \eta \nabla J_w$$

$$\frac{P(y=+1)}{P(y=-1)} = \exp(w^T x)$$

$$e^{w^T x}$$

$$\frac{P(y=+1)}{1-P(y=+1)} = \exp(w^T x)$$

$$\cancel{\frac{1}{1-P(y=+1)} = 1 + \exp(w^T x)}$$

$$1 \cancel{- P(y=+1)} =$$

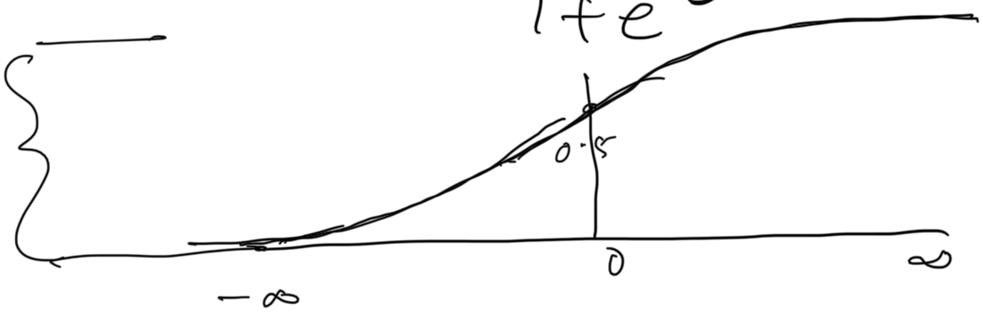
$$1 - P(y=+1) = \exp(-w^T x)$$

$$\frac{1}{p(y=+1)} = 1 + \exp(-w^T x)$$

$$\boxed{p(y=+1) = \frac{1}{1+\exp(-w^T x)}}$$

Sigmoid function

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



If $p(y=+1) \geq 0.5 \quad y=+1$

$\qquad\qquad\quad < 0.5 \quad y=-1$

If $w^T x \geq 0 \quad y=+1$

$\qquad\quad < 0 \quad y=-1$

equivalent

$\frac{1}{\cdots} \geq 0.5 = \frac{1}{\cdots}$

$(1+\exp(-w^T x))$ $\qquad$ 2

$$\cancel{w^T x \geq 0.5}$$

$$2 \geq 1 + \exp(-w^T x)$$

$$\exp(-w^T x) \leq 1$$

$$-w^T x \leq 0$$

$$\text{or} \quad \underline{w^T x \geq 0}$$

Find w such that

p_1

p_2

prob $\rightarrow$ p_3 $\quad \max \sum_{i=1}^{N} \log p_i$

$p_i(y_i = +1)$

$\vdots \quad \min -\frac{1}{N} \sum_{i=1}^{N} \underline{\log p_i}$

p_N

$$-\frac{1}{N} \sum \log(1 + \exp(-y_i w^T x_i))$$

$$\frac{1}{N} \sum \log(1 + \exp(-y_i w^T x_i$$

$= \frac{1}{N}$

How to optimize for W?

Direct Minimization

$\underline{\nabla J = 0}$ $\times$

$$\nabla J = \frac{d}{dW} J$$

$$H = \frac{d}{dW} \nabla J = \nabla^2 J$$