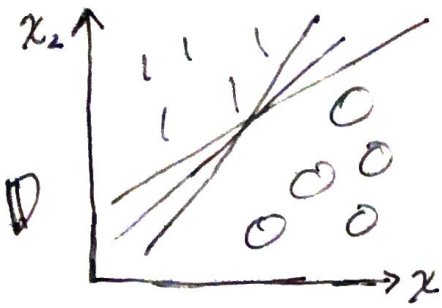
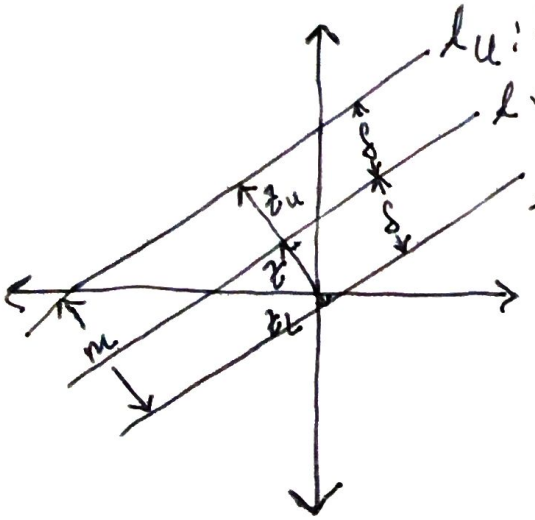
Math 342W Lecture 5



$\Rightarrow$ Perceptron gives us some line to separate the two regions



$\Rightarrow$ Check Previous lecture for m and $\vec{z}$

Recall that the Hesse Normal form is overparameterized.

↳ $\forall c \neq 0$, $c(\vec{w}\cdot\vec{x} - b) = 0 \Rightarrow$ represents ℓ

Let $c = \frac{1}{\delta}$, $\exists\ \vec{w}_1, b$

$\ell = \vec{w}\cdot\vec{x} - b = 0$
$\ell_2 = \vec{w}\cdot\vec{x} - (b-1) = 0$
$\ell_u = \vec{w}\cdot\vec{x} - (b+1) = 0$

Goal: Make m as large as possible $\equiv$ $\|\vec{w}\|$ small as possible
↳ "No points in the way", s.t

a) $\forall y_i = 1$, $\vec{x}_i \in R_u$

b) $\forall y_i = 0$, $\vec{x}_i \in R_2$

a) $\vec{x}_i$ is above l_u, which means

↳ x_{i2} > the corresponding x_{i2} on l_u at x_{i1}

↳ $l_u: \vec{w} \cdot \vec{x}_i - (b-1) = 0 \Rightarrow w \cdot x_i = b+1$

$\Rightarrow w_1 x_{i1} + w_2 x_{i2} = b+1$

$\Rightarrow x_{i2} = \frac{1}{w_2}(b+1 - w_1 x_{i1})$

if $\vec{x}_i \in B_u \Rightarrow x_{i2} \geq \frac{1}{w_2}(b+1-w_1 x_{i1})$

⋮

By "same" process

$B_u: \vec{w} \cdot \vec{x}_i - (b+1) \geq 0$ (a)

$B_l: \vec{w} \cdot \vec{x}_i - (b-1) \leq 0$ (b)

(a) $\forall y_i = 1$,

$\vec{w} \cdot \vec{x}_i - b - 1 \geq 0 \Rightarrow \vec{w} \cdot \vec{x}_i - b \geq 1$

$\Rightarrow \frac{1}{2}(\vec{w} \cdot \vec{x}_i - b) \geq \frac{1}{2} \Rightarrow (y_i - \frac{1}{2})(\vec{w} \cdot \vec{x}_i - b) \geq \frac{1}{2}$

(b) $\forall y_i = 0$,

$\vec{w} \cdot \vec{x}_i - b + 1 \leq 0 \Rightarrow \vec{w} \cdot \vec{x}_i - b \leq 1$

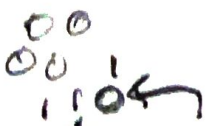
$\Rightarrow \frac{1}{2}(\vec{w} \cdot \vec{x}_i - b) \leq -\frac{1}{2} \Rightarrow (y_i - \frac{1}{2})(\vec{w} \cdot \vec{x}_i - b) \geq \frac{1}{2}$

A_{sum}: Find $\vec{w}, b$ s.t

$\forall_i (y_i - \frac{1}{2})(\vec{w} \cdot \vec{x}_i - b) \geq \frac{1}{2}$, where $\|\vec{w}\|$ is minimal

Notes.

1) No analytical soln', need optimization heuristic
2) This can be shown to be a "quadratic problem"
3) General for all $p \geq 2$
4) Wikipedia have "1" instead of a "$\frac{1}{2}$"
   ↳ $y = \{-1, 1\}$ ↳ $y = \{0, \frac{1}{2}\}$



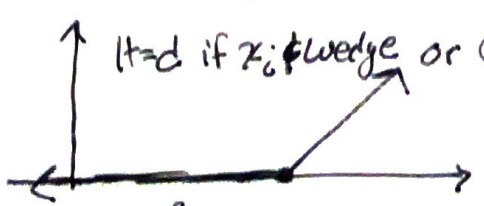
Assume $\vec{x}_{17}$ is the point where $\{0, \frac{1}{2}\}$ "sticks out" from the group.

↳ SVM doesn't work now, so we need to fix this.

$\Rightarrow (y_i - \frac{1}{2})(\vec{w} \cdot \vec{x}_i - b) \not\geq \frac{1}{2}$

$\Rightarrow \exists\, d > 0$, s.t $(y_i - \frac{1}{2})(\vec{w} \cdot \vec{x}_i - b) + d = \frac{1}{2}$ ↗ d is a measure of the violation

Consider "Hinge Loss" H



$\Rightarrow H_i = \max\{0, \frac{1}{2} - (y_i - \frac{1}{2})(\vec{w} \cdot \vec{x} - b)\}$

$SHE = \sum_{i=1}^{n} H_i$

We want to...
1) Minimize SHE
2) Maximize wedge width

We have two objective functions

Vapnik (1963)

$$\underset{\vec{w}, b}{\operatorname{argmin}} \left\{ \frac{1}{n} SHE + \lambda \|\vec{w}\|^2 \right\}$$

↳ Let's interpret this in words

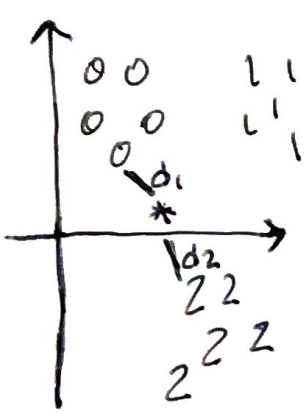
$\frac{1}{n}$ SHE → Average (Sum Hinge Error)

λ = Hyperparameter or "tuning parameter"

f_{Vapnik} which you specify

$g = f(\mathbb{D}, \mathcal{H}, \lambda)$ or $g = f_\lambda(\mathbb{D}, \mathcal{H})$

$y = \{0, 1, 2, \ldots, L\}$, $L \geq 2$



$g: \mathcal{X} \to y$

$* \Rightarrow d_1 < d_2 \Rightarrow 0$

Here we used 1 neighbor, do this with $2, 3, 4, \ldots \frac{k}{5}$

Nearest Neighbors

1-Nearest Neighbor Algorithm

0) Find function $d: \mathcal{X} \times \mathcal{X} \to \mathbb{R}_{\geq 0}$

1) Calculate $d_i = d(\vec{x}^*, \vec{x}_i) \; \forall i$

2) Find $i^* = \operatorname{argmin}(d_i)$

3) Return y_{i^*}

K-Nearest Neighbor Algorithm

0) " "

1) " "

2) $i_1^*, \ldots, i_n^*$ of d_i's

3) Return mode $[y_{i_1}^*, y_{i_2}^*, \ldots, y_{i_n}^*]$ => Ties Go Randomly

These are dependent on Euclidean Distance

$$\Rightarrow d = \sqrt{\sum_{j=1}^{p} (x_0^* - x_{ij})^2}$$