ECE 544 NA
9/16/2013

TODAY

- SEPARABILITY
- GENERALIZED LINEAR SEPARATRIX
- ERROR: MEAN SQUARED ERROR

SEPARABILITY

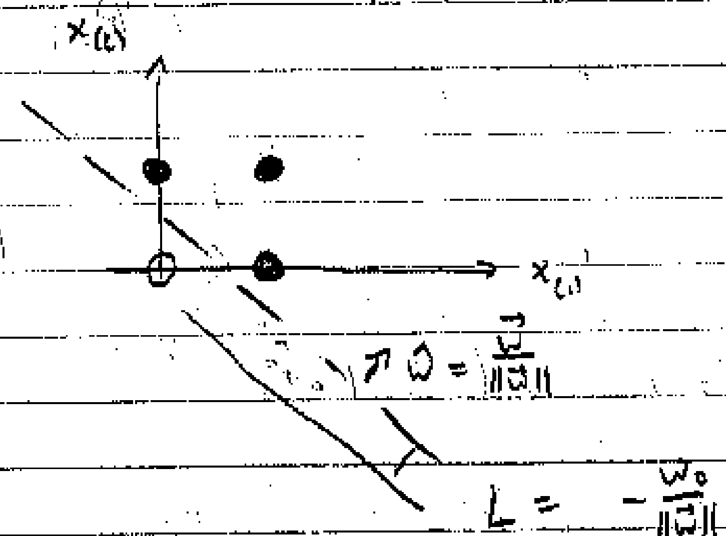
TOKEN SETS $\mathcal{X} = \{\vec{x}^1, \ldots, \vec{x}^N\}$, $\vec{x}^n \in \mathbb{R}^D$
AND $\mathcal{Z} = \{\vec{z}^1, \ldots, \vec{z}^N\}$, $\vec{z}^n \in \mathbb{R}^D$

ARE LINEARLY SEPARABLE IFF

$\exists\, \vec{w}, w_0$ s.t.

$$\vec{w}^T\vec{x}^n + w_0 > 0 \quad \forall\, \vec{x}^n \in \mathcal{X}$$

$$\vec{w}^T\vec{z}^n + w_0 < 0 \quad \forall\, \vec{z}^n \in \mathcal{Z}$$



ACADEMICS LOVE: $p(C_k \mid \vec{x}) = \dfrac{1}{1 + e^{-\vec{w}^T\vec{x} + w_0}}$

CONSUMERS WANT: 100% ACCURACY

$\Leftrightarrow$ PERFECT SEPARABILITY

CONVEX HULL THEOREM

$\mathcal{X}$ AND $\mathcal{Z}$ ARE LINEARLY SEPARABLE

IFF

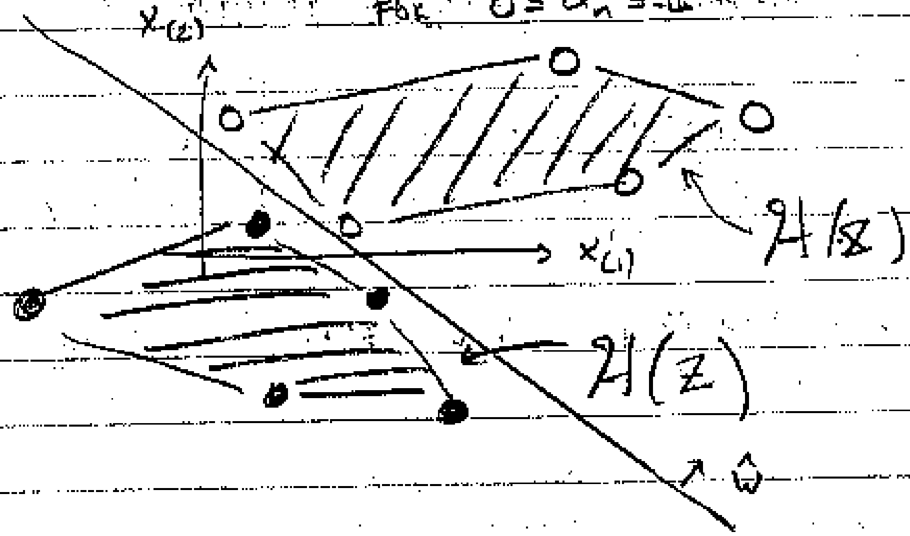
$\mathcal{H}(\mathcal{X}) \cap \mathcal{H}(\mathcal{Z}) = \emptyset$

IFF

$\nexists$ NO $\vec{x}^n \in \mathcal{H}(\mathcal{Z})$ AND $\nexists$ NO $\vec{z}^n \in \mathcal{H}(\mathcal{X})$

CONVEX HULL $\mathcal{H}(\mathcal{X}) = \left\{ \vec{x} : \vec{x} = \sum_{n=1}^{N} \alpha_n \vec{x}^n \right\}$

FOR $0 \leq \alpha_n \leq 1$



PROOF: $\mathcal{H}(\mathcal{Z}) \cap \mathcal{H}(\mathcal{X}) \neq \emptyset$

IFF

$\exists\ \vec{x} = \sum_{n=1}^{N} \alpha_n \vec{x}^n = \sum_{n=1}^{N} \beta_m \vec{z}^n$

$0 \leq \alpha_n \leq 1,\ 0 \leq \beta_m \leq 1$

THEN: $\vec{w}^T \vec{x} + w_0 \geq 0$

IFF $\exists\ \vec{z}^m$ S.T. $\vec{w}^T \vec{z}^n + w_0 \geq 0$

GENERALIZED LINEAR SEPARATRIX

SUPPOSE $\mathcal{X}$, $\mathcal{Z}$ NOT LINEARLY SEPARABLE.

CONSTRUCT NONLINEAR FUNCTIONS $\phi_j(\vec{x}^n)$

THEN

$$y(\vec{x}^n) = w_0 + \sum_{j=1}^{M} \phi_j(\vec{x}^n)$$

ALGORITHM TO GUARANTEE

GENERALIZED LINEAR SEPARABILITY

I. INITIALIZE START WITH $\vec{\phi}^n = \begin{bmatrix} 1 \\ \phi_1(\vec{x}^n) \\ \vdots \\ \phi_M(\vec{x}^n) \end{bmatrix}$

AND $\vec{\psi}^n = \begin{bmatrix} 1 \\ \phi_1(\vec{z}^n) \\ \vdots \\ \phi_M(\vec{z}^n) \end{bmatrix}$, E.G. START WITH $\phi_j(\vec{x}) = x_{(j)}$.

II. ITERATE

A. CHECK IF CONVEX HULLS OVERLAP

FOR EACH $\vec{\phi}^n$, TRY TO FIND $\vec{\phi}^n = \sum_{k=1}^{N} \beta_k^n \vec{\psi}^k$

S.T. $0 \leq \beta_k^n \leq 1$ (USE LAGRANGIAN).

IF NO SUCH β_k^n EXIST, OK.

IF SUCH β_k^n EXIST, $\vec{\phi}^n$ IS A "POTENTIAL ERROR".

B. ADD AN $(M+1)^{ST}$ FEATURE

CREATE A $\phi_{M+1}(\vec{x})$ THAT MOVES AT LEAST ONE "POTENTIAL ERROR" OUT OF THE OPPOSING CONVEX HULL.

HOW TO FIND BETAS:

LAGRANGIAN

MINIMIZE $\quad \mathcal{E} = \frac{1}{2}\left\| \sum_{k=1}^{N} \beta_k \vec{\Psi}^k - \vec{\Phi} \right\|^2 = \frac{1}{2}\left\| \vec{\Phi} - \Psi \vec{\beta}^T \right\|_2^2$

$$\vec{\beta} \in R^N \quad \vec{\Phi} \in R^{M+1} \quad \Psi \in R^{(M+1) \times N}$$

$\Longleftrightarrow$ MINIMIZE

$$\mathcal{L} = \frac{1}{2}\left\| \vec{\Phi} - \Psi \vec{\beta} \right\|_2^2 + \sum_{k=1}^{N} \lambda_k \left[\min(0, \beta_k - 1) - \max(0, \beta_k) \right]$$

$\mathcal{L} = 0$: $\vec{\Phi}$ NOT IN $\mathcal{H}(\mathcal{Z})$

$\mathcal{L} > 0$: $\vec{\Phi}$ IN $\mathcal{H}(\mathcal{Z})$

$$0 = \frac{\partial \mathcal{L}}{\partial \beta^l} = \begin{cases} \vec{\Psi}^{l\,T}\left(\vec{\Phi} - \Psi \vec{\beta}^T\right) & 0 \le \beta^l \le 1 \\ \text{---}\;''\;\text{---} + \lambda_l & \beta^l \ge 1 \\ \text{---}\;''\;\text{---} - \lambda_l & \beta^l \le 0 \end{cases}$$

$$\nabla_{\vec{\beta}} \mathcal{L} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial \beta_1} \\ \vdots \\ \frac{\partial \mathcal{L}}{\partial \beta_N} \end{bmatrix} = \left(\Psi^T \vec{\Phi} - \Psi^T \Psi \vec{\beta}^T\right) + \vec{\lambda}[\vec{\beta} \ge 1] - \vec{\lambda}[\vec{\beta} \le 0]$$

ARBITRARY CONSTANTS

— CHOOSE SO ALL OF $0 \le \beta_l \le 1$

$$\vec{\beta}_{opt} = \max\left(0, \min\left(1, (\Psi^T \Psi)^{-1} \Psi^T \vec{\Phi}\right)\right)$$