ECE 544NA

9/12/2013

TODAY

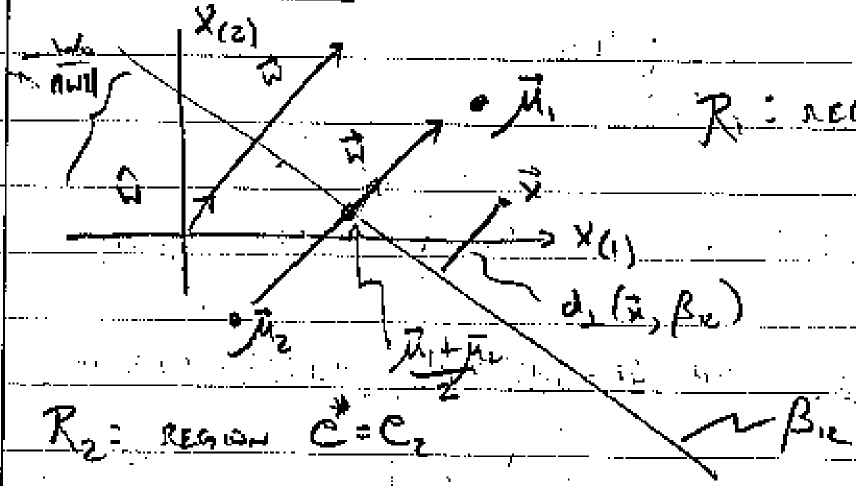
LINEAR DISCRIMINANTS

1. MINIMUM DISTANCE
2. GAUSSIAN CLASSES
3. BINARY OBSERVATIONS

MINIMUM-DISTANCE CLASSIFICATION:

CHOOSE $C^* = C_k$ IFF $|\vec{x} - \vec{\mu}_k| < |\vec{x} - \vec{\mu}_\ell| \quad \forall \ell \neq k$

TWO CLASSES



$$R_1 \equiv \{ x : |\vec{x} - \vec{\mu}_1| < |\vec{x} - \vec{\mu}_2| \}$$

$$= \{ x : |\vec{x} - \vec{\mu}_1|^2 < |\vec{x} - \vec{\mu}_2|^2 \}$$

$$= \{ x : \cancel{\vec{x}^T\vec{x}} - 2\vec{x}^T\vec{\mu}_1 + \vec{\mu}_1^T\vec{\mu}_1 < \cancel{\vec{x}^T\vec{x}} - 2\vec{x}^T\vec{\mu}_2 + \vec{\mu}_2^T\vec{\mu}_2 \}$$

$$= \{ x : \vec{w}^T\vec{x} + w_0 > 0 \}$$

$$\vec{w} = 2(\vec{\mu}_1 - \vec{\mu}_2) \qquad w_0 = \vec{\mu}_2^T\vec{\mu}_2 - \vec{\mu}_1^T\vec{\mu}_1 = -\vec{w}^T\left(\frac{\vec{\mu}_1 + \vec{\mu}_2}{2}\right)$$

UNIT NORMAL VECTOR : $\hat{w} = \frac{\vec{w}}{\|\vec{w}\|_2}$

SEPARATRIX IS THE HYPERPLANE $\vec{w}^T\vec{x} + w_0 = 0$

$\frac{w_0}{\|\vec{w}\|} = -\hat{w}^T\left(\frac{\vec{\mu}_1 + \vec{\mu}_2}{2}\right)$

$$\mathcal{B}_{12} = \left\{ x : \frac{\vec{w}^T\vec{x}}{\|\vec{w}\|} + \frac{w_0}{\|\vec{w}\|} = 0 \right\}$$

DISTANCE FROM x TO SEPARATRIX

$$d_\perp(\vec{x}, \mathcal{B}_{12}) \doteq \min_{\vec{y} \in \mathcal{B}_{12}} \|\vec{x} - \vec{y}\| = \frac{\vec{w}^T\vec{x} + w_0}{\|\vec{w}\|}$$

$$= \hat{w}^T\vec{x} + \frac{w_0}{\|\vec{w}\|}$$

AUGMENTED VECTORS

$$\tilde{x} = \begin{bmatrix} 1 \\ \vec{x} \end{bmatrix} \quad \tilde{w} = \begin{bmatrix} w_0 \\ \vec{w} \end{bmatrix}$$

$$\mathcal{R}_1 = \{ \vec{x} : \tilde{w}^T\tilde{x} > 0 \} \Leftarrow \text{NO OFFSET}$$

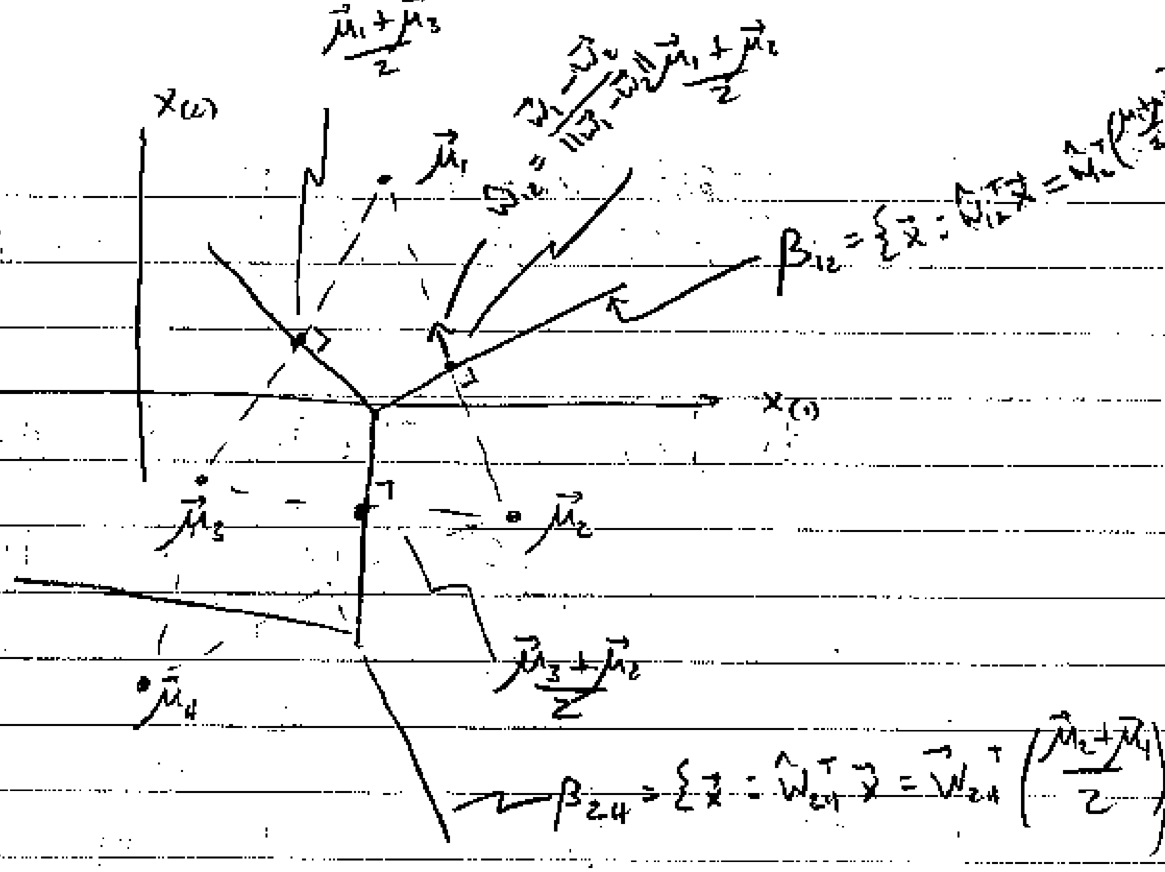
MULTI-CLASS MINIMUM-DISTANCE CLASSIFICATION

$$\mathcal{R}_k = \{ \vec{x} : |\vec{x} - \vec{\mu}_k| < |\vec{x} - \vec{\mu}_\ell| \quad \forall \ell \neq k \}$$

$$= \{ \vec{x} : \cancel{\vec{x}^T\vec{x}} - 2\vec{x}^T\vec{\mu}_k + \vec{\mu}_k^T\vec{\mu}_k < \cancel{\vec{x}^T\vec{x}} - 2\vec{x}^T\vec{\mu}_\ell + \vec{\mu}_\ell^T\vec{\mu}_\ell \}$$

$$\boxed{= \{ \vec{x} : \vec{w}_k^T\vec{x} + w_{k0} > \vec{w}_\ell^T\vec{x} + w_{\ell 0} \}}$$

$$\vec{w}_k = 2\vec{\mu}_k \quad , \quad w_{k0} = -\vec{\mu}_k^T\vec{\mu}_k$$



GAUSSIAN CLASSES WITH SAME COVARIANCE

$$p(\vec{x}\mid C_1)=\mathcal{N}(\vec{x};\vec{\mu}_1,\Sigma) \qquad \pi_1=P(C_1)$$
$$p(\vec{x}\mid C_2)=\mathcal{N}(\vec{x};\vec{\mu}_2,\Sigma) \qquad \pi_2=P(C_2)$$

$$p(C_1\mid\vec{x})=\frac{\pi_1\,p(\vec{x}\mid C_1)}{\pi_1\,p(\vec{x}\mid C_1)+\pi_2\,p(\vec{x}\mid C_2)}$$
$$=\frac{\pi_1 e^{-\frac{1}{2}(\vec{x}-\vec{\mu}_1)^T\Sigma^{-1}(\vec{x}-\vec{\mu}_1)}}{\pi_1 e^{-\frac{1}{2}(\vec{x}-\vec{\mu}_1)^T\Sigma^{-1}(\vec{x}-\vec{\mu}_1)}+\pi_2 e^{-\frac{1}{2}(\vec{x}-\vec{\mu}_2)^T\Sigma^{-1}(\vec{x}-\vec{\mu}_2)}}$$

$$(\vec{x}-\vec{\mu}_1)^T\Sigma^{-1}(\vec{x}-\vec{\mu}_1)-(\vec{x}-\vec{\mu}_2)^T\Sigma^{-1}(\vec{x}-\vec{\mu}_2)+\ln\left(\frac{\pi_2}{\pi_1}\right)$$
$$=\cancel{\vec{x}^T\Sigma^{-1}\vec{x}}-2\vec{x}^T\Sigma^{-1}\vec{\mu}_1+\vec{\mu}_1^T\Sigma^{-1}\vec{\mu}_1$$
$$-\cancel{\vec{x}^T\Sigma^{-1}\vec{x}}+2\vec{x}^T\Sigma^{-1}\vec{\mu}_2-\vec{\mu}_2^T\Sigma^{-1}\vec{\mu}_2+\ln\left(\frac{\pi_2}{\pi_1}\right)$$
$$=-\vec{w}^T\vec{x}-w_0$$

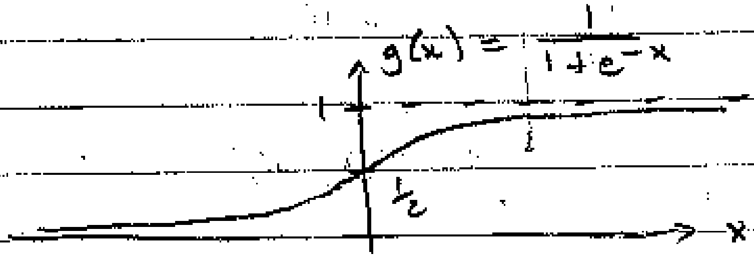
$$\vec{w} = 2\Sigma^{-1}(\vec{\mu}_1 - \vec{\mu}_2)$$

$$w_0 = \ln\left(\frac{\pi_2}{\pi_1}\right) + \vec{\mu}_2^T \Sigma^{-1} \vec{\mu}_2 - \vec{\mu}_1^T \Sigma^{-1} \vec{\mu}_1 = -\vec{w}^T\left(\frac{\vec{\mu}_1 + \vec{\mu}_2}{2}\right) + \ln\left(\frac{\pi_2}{\pi_1}\right)$$

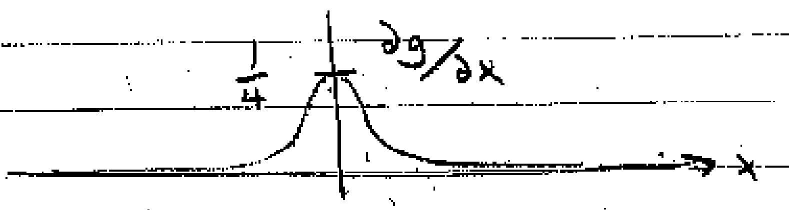
so

$$p(C_1 | \vec{x}) = \frac{1}{1 + e^{-(\vec{w}^T\vec{x} + w_0)}} = \underbrace{g(\vec{w}^T\vec{x} + w_0)}_{\boxed{\text{LOGISTIC}} \text{ or } \boxed{\text{SIGMOID}}}$$

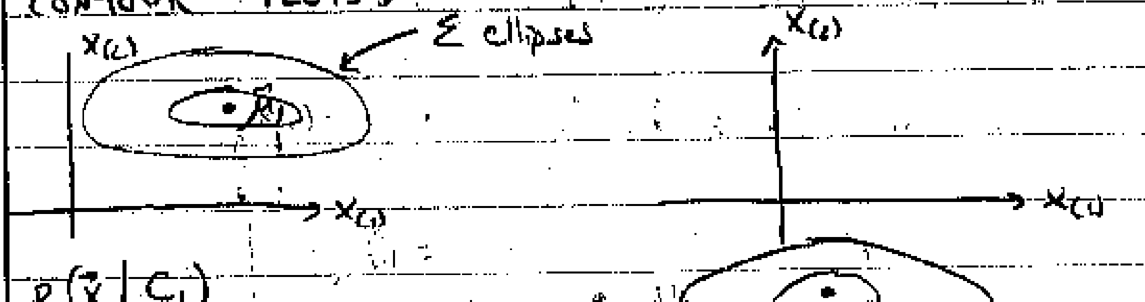
$$g(x) = \frac{1}{1+e^{-x}}$$

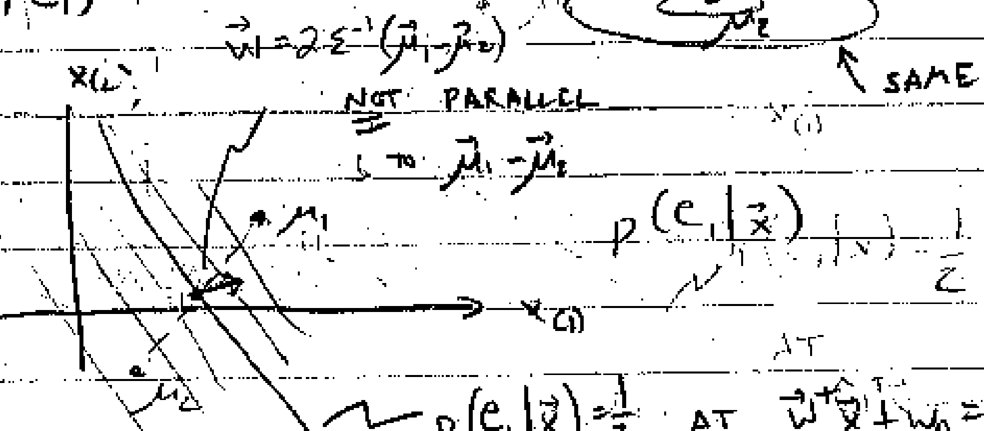


$$\frac{\partial g(x)}{\partial x} = \frac{\partial}{\partial x}\left(\frac{1}{1+e^{-x}}\right) = +\frac{e^{-x}}{(1+e^{-x})^2} = g(x)(1-g(x))$$



CONTOUR PLOTS:





$p(C_1 | \vec{x}) = \frac{1}{2}$ AT $\vec{w}^T\vec{x} + w_0 = 0$

BINARY OBSERVATIONS

$$Pr\{x_{(d)} \mid C_k\} \equiv \begin{cases} P_{dk} & x_{(d)} = 1 \\ 1 - P_{dk} & x_{(d)} = 0 \end{cases} = P_{dk}^{x_{(d)}} (1-P_{dk})^{(1-x_{(d)})}$$

$$p(\vec{x} \mid C_k) = \prod_{d=1}^{D} P_{dk}^{x_{(d)}} (1-P_{dk})^{(1-x_{(d)})}$$

$$\ln p(\vec{x} \mid C_k) = \sum_{d=1}^{D} \left[\ln(1-P_{dk}) + x_{(d)} \left(\ln P_{dk} - \ln(1-P_{dk}) \right) \right]$$

$$\ln \pi_k + \ln p(\vec{x} \mid C_k) = W_{k0} + \vec{W}_k^T \vec{x}$$

$$W_{dk} = \ln P_{dk} - \ln(1-P_{dk})$$
$$W_{0k} = \ln \pi_k + \sum_{d=1}^{D} \ln(1-P_{dk})$$

CHOOSE $C^* = C_k$ IFF

$$\ln \pi_k + \ln p(\vec{x} \mid C_k) > \ln \pi_\ell + \ln p(\vec{x} \mid C_\ell)$$

FOR ALL $\ell \neq k$

$$\Updownarrow$$

$$W_{k0} + \vec{W}_k^T \vec{x} > W_{\ell 0} + \vec{W}_\ell^T \vec{x} \quad \forall\, \ell \neq k$$

TWO-CLASS

$$p(C_1 \mid \vec{x}) = \frac{1}{1 + e^{-\vec{w}^T \vec{x} + w_0}}$$

$$w_d = \ln\left(\frac{P_{d2}}{P_{d1}}\right) - \ln\left(\frac{1-P_{d2}}{1-P_{d1}}\right), \quad w_0 = \ln\left(\frac{\pi_2}{\pi_1}\right) + \sum_{d=1}^{D} \ln\left(\frac{1-P_{d2}}{1-P_{d1}}\right)$$