① Apply LDA on the following Dataset with 4 pts

where $x \in R^2$ and $y \in \{-1, +1\}$

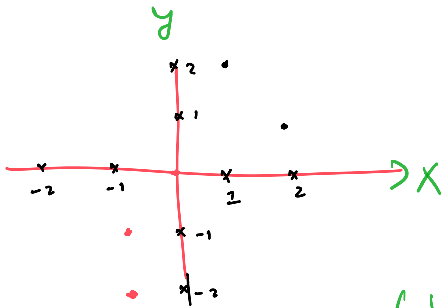
* [illegible] ✓
* [illegible] the projected dataset ✓ $\hat{w}^T x$
* [illegible] the classifier ✓

$$D = \{ (\underbrace{(1,2)}_{1}, +1), (\underbrace{(2,1)}_{2}, +1), (\underbrace{(-1,-1)}_{3}, -1), (\underbrace{(-1,-2)}_{4}, -1) \}$$

$x_1 = (1,2)$

$y_1 = +1$

D in Raw space



LDA algo

① class means

$$m_{+1} = \frac{(1,2) + (2,1)}{2}$$

$$= \frac{(3,3)}{2} = \text{[illegible]})$$

$(1,2) - (1.5, 1.5)$

$= (-0.5, +0.5)$

$$m_{-1} = \frac{(-1,-1) + (-1,-2)}{2} = \frac{(-2,-3)}{2} = \text{[illegible]}$$

② within class Variance $S_w = \sum_{y \in C} S^2$ [illegible]

[illegible] = [illegible] + [illegible]

$$S_{+1}^2 = \frac{1}{n_{+1}} \sum_i \left[(x_i - m_i)(x_i - m_i)^T \right]$$ [illegible]

matrix

$$= \frac{1}{2} \left[\begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix} [-0.5 \quad 0.5] + \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix} [0.5 \quad -0.5] \right]$$

$0.25 + 0.25 =$

$= 0.5$

Inner product

$$= \frac{1}{2} \left[\begin{bmatrix} 0.25 & -0.25 \\ -0.25 & 0.25 \end{bmatrix} + \begin{bmatrix} 0.25 & -0.25 \\ -0.25 & 0.25 \end{bmatrix} \right]$$

$$= \begin{bmatrix} 0.25 & -0.25 \\ -0.25 & 0.25 \end{bmatrix}$$

For point $(-1,-2)$

$$S_{-1}^{2} = \frac{1}{2}\left(\begin{bmatrix} 0 \\ 0.5 \end{bmatrix} \begin{bmatrix} 0 & 0.5 \end{bmatrix} + \begin{bmatrix} \\ -0.5 \end{bmatrix} \right)$$

$$= \frac{1}{2}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0.25 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0.25 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0.25 \end{bmatrix}$$

$$\vec{w} = S_W^{-1}(m_2 - m_1)$$

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ Invert sign, swap

$$|A| = ad - bc$$

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \quad \frac{1}{|A|}$$

$$S_W^{-1}(m_{+1} - m_{-1})$$

$$S_W = \begin{bmatrix} 0.25 & -0.25 \\ -0.25 & 0.5 \end{bmatrix}$$

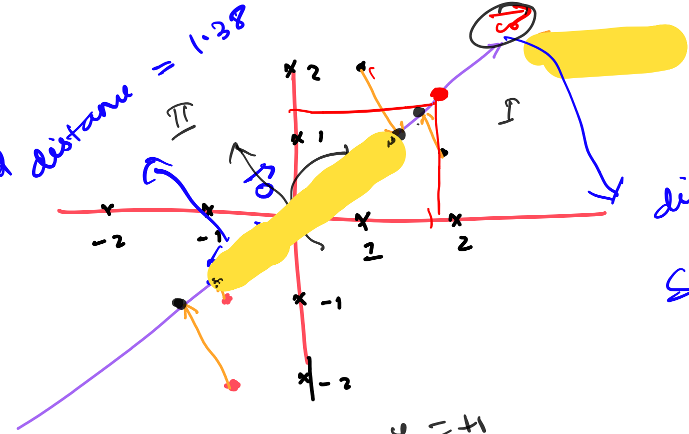
$$S_W^{-1} = k \begin{bmatrix} 0.5 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}$$

$$w \propto S_W^{-1} \left[m_{+1} - m_{-1} \right]$$

$$= \begin{bmatrix} 0.5 & 0.25 \\ 0.25 & 0.25 \end{bmatrix} \begin{bmatrix} 2.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 1.25 + 0.75 \\ 0.25 * 5.5 \end{bmatrix} = \begin{bmatrix} \quad \end{bmatrix}$$

$$\vec{w} = \begin{bmatrix} 2 & 1.375 \end{bmatrix}$$

projected distance = 1.38



direction of $\vec{w}$ is

Scale of $\vec{w}$ is Immaterial

If $\hat{w}^T x \geq 0$ $y = +1$

else $y = -1$

when we compute the projection of $[-1, -1]$ we need to normalize $\vec{w}$

Normalization

$$\hat{w} = \frac{[2, 1.375]}{\sqrt{4 + 1.89}} = \frac{[2, 1.375]}{2.42}$$

$$||\vec{w}|| = \sqrt{2^2 + 1.375^2}$$

$$= [0.82, 0.56] = \text{Imprecise}$$

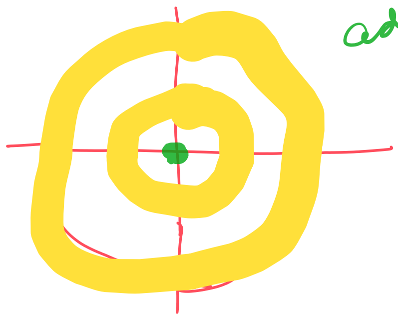
projection $= \hat{w}^T x = [0.82, 0.56]^T [-1, -1]$

$$\boxed{= -0.82 - 0.56 = -1.38}$$

classifier $w^T x >$ any pt $\in [\max_{y \in -1} w^T x, \min_{y \in +1} w^T x]$

② mention the Impact of LDA on the following

Concentric circle dataset



adversarial d.c. for LDA

LDA will have no Impact because

$m_{+1} = m_{-1}$ and $(m_{+1} - m_{-1}) = 0$ $\qquad w \propto S_w^{-1} \underline{(m_2 - m_1)} = 0$

$\boxed{w = \vec{0}}$

and $\vec{w} = 0$

The projected dataset is $\longrightarrow \vec{w} = \vec{0}$

0

③ Assume that features are obtained from 5 dimensions

(ii) $x \in \mathbb{R}^5$ and $y \in \{-1, +1\}$. what will be the dimensions of the projected dataset

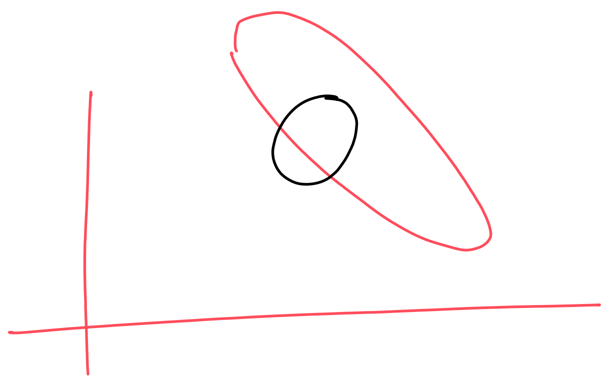
$$w \propto S_w^{-1} (m_{+1} - m_{-1})$$

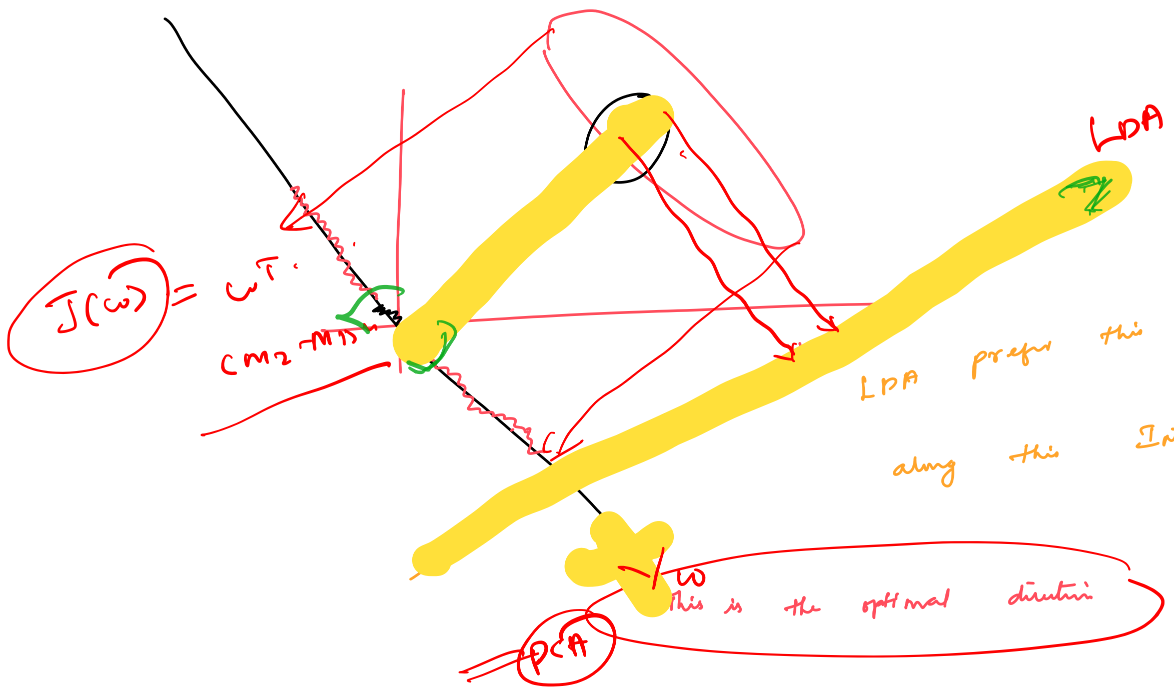
[illegible] [illegible]

$\Rightarrow w \in$ [illegible]

The projected pts are [illegible] $\in$ [illegible]

does LDA work for the following dataset





LDA prefers this direction because along this Inter class mean is maximized

④ Apply LDA on the following degenerate dataset

$\{ ((-2\ \ ), -1), ((-1\ \ ), -1), ((1\ \ ), +1), ((2\ \ ), +1) \}$

$w = (1, 0)$

You will see that

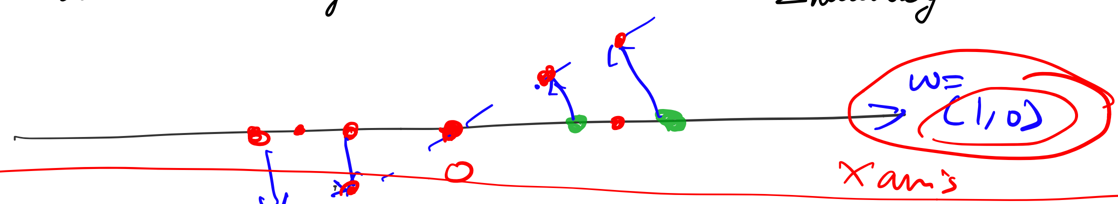
$S_w^{-1}(M_2 - M_1)$

$(1, 0.75)$

$J(\omega)$

For different $\vec{\omega}$

however we can argue the solution Intuitively



⑤ Given a 2-class dataset $D = \{ (x_i, y_i) \}$

we applied FLD and obtained the direction (say) $\vec{w}$

Now form the projected dataset $\tilde{D} = \{ (w^T x_i, y_i)$

$\forall (x_i, y_i) \in D \}$

Now apply FLD on $\tilde{D}$ and obtain $\tilde{w}$

$w^T x + \cancel{b}$