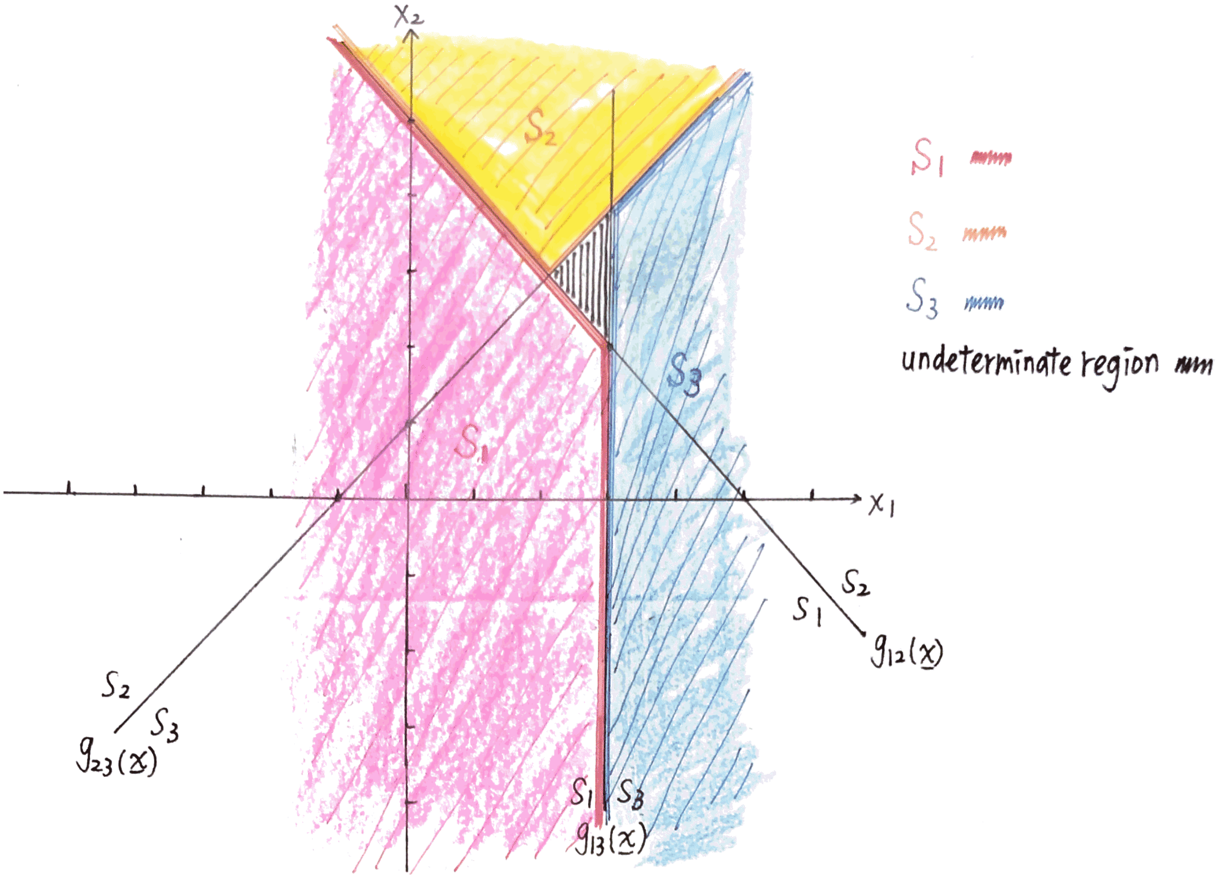
EE559 Week3 HW 1139093874 Pin-Hsuan Lee

1.



① $\underline{x} = (4,1) \in S_3$

$(1,5) \in S_2$ We can see from the above plot.

$(0,0) \in S_1$

② $\underline{x} = (2.5, 3) \in$ undeterminate region

prove: $g_{12}(2.5,3) = -2.5-3+5 = -0.5 < 0 \Rightarrow (2.5,3) \in S_2, \notin S_1$ —(a)

$g_{13}(2.5,3) = -2.5+3 = 0.5 > 0 \Rightarrow (2.5,3) \in S_1, \notin S_3$ —(b)

$g_{23}(2.5,3) = -2.5+3-1 = -0.5 < 0 \Rightarrow (2.5,3) \in S_3, \notin S_2$ — (c)

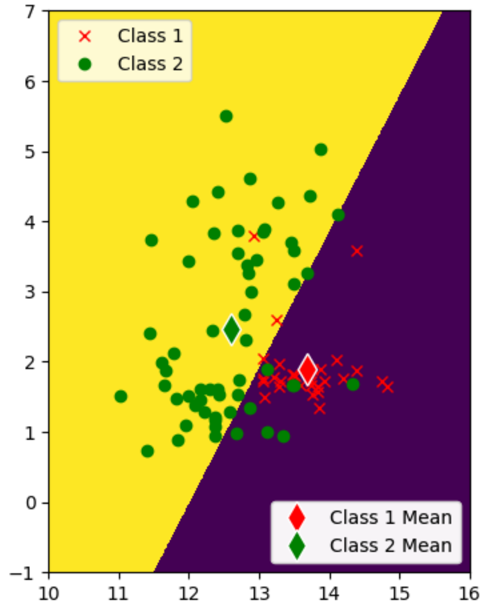
From (a)(b)(c), we cannot find an intersection region for $\underline{x} = (2.5,3)$

$\Rightarrow \underline{x} = (2.5,3)$ is in inderterminate region.

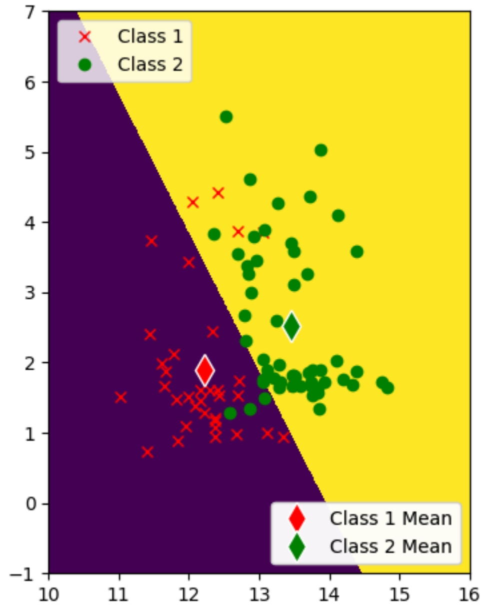
2.(a) The classification of 66 points in training data are correct and classification accuracy of training set is : 74.15%.

The classification of 63 points in testing data are correct and classification accuracy of testing set is : 70.79%.

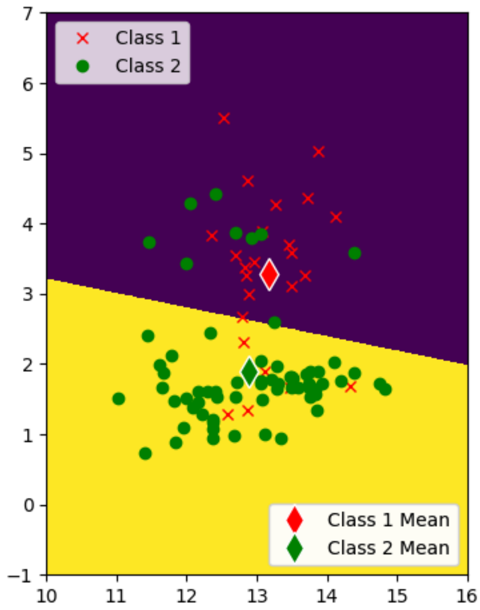
(b) S_1 (Class 1 in the plot) *vs.* S_1'(Class 2 in the plot)



S_2 (Class 1 in the plot) *vs.* S_2'(Class 2 in the plot)

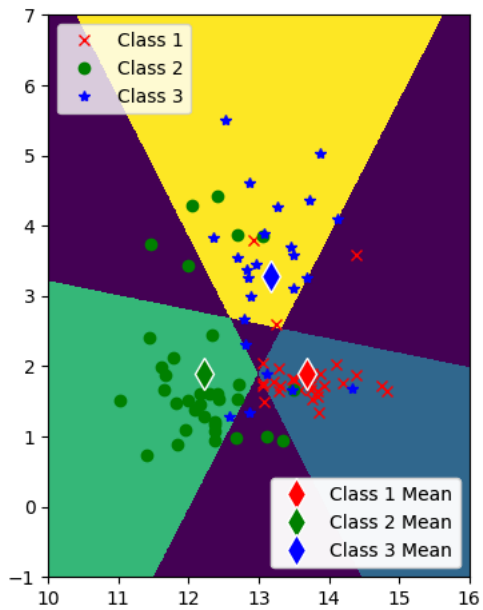


S_3 (Class 1 in the plot) *vs.* $S_3{}'$(Class 2 in the plot)



(c) In the plot below, $\Gamma_1, \Gamma_2, \Gamma_3$ is shown as class 1, class 2 and class 3.

And the purple areas are indeterminate regions.



3. a) A 2-class classifier is linear iff $g(\underline{x})$ can be expressed as a function of $\underline{x}$.

$g(\underline{x}) = w_0 + w_1x_1 + w_2x_2$. Assume there are 2 classes, the mean of class 1 is μ_1, the mean of class 2 is μ_2, and $x_1 \in S_1$.

$$\mu_1 = \begin{bmatrix} \mu_{11} \\ \mu_{12} \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} \mu_{21} \\ \mu_{22} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$\because x \in S_1$, based on Euclidean distance $\Rightarrow \sqrt{(x-\mu_1)^2} < \sqrt{(x-\mu_2)^2}$

$\Rightarrow (x_1-\mu_{11})^2 + (x_2-\mu_{12})^2 < (x_1-\mu_{21})^2 + (x_2-\mu_{22})^2$

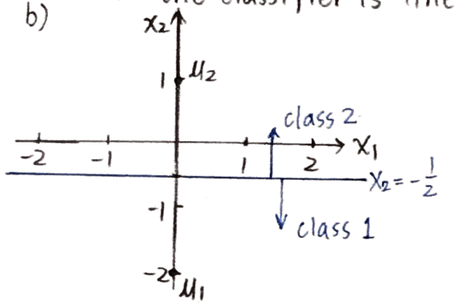
$\Rightarrow \cancel{x_1^2} - 2\mu_{11}x_1 + \mu_{11}^2 + \cancel{x_2^2} - 2\mu_{12}x_2 + \mu_{12}^2 < \cancel{x_1^2} - 2\mu_{21}x_1 + \mu_{21}^2 + \cancel{x_2^2} - 2\mu_{22}x_2 + \mu_{22}^2$

$\Rightarrow \underbrace{(\mu_{11}^2 + \mu_{12}^2 - \mu_{21}^2 - \mu_{22}^2)}_{w_0} + \underbrace{(2\mu_{21} - 2\mu_{11})}_{w_1}x_1 + \underbrace{(2\mu_{22} - 2\mu_{12})}_{w_2}x_2 < 0$

$\because$ we can use $g(x) = w_0 + w_1x_1 + w_2x_2$ to express the classifier,

$\therefore$ the classifier is linear.

b)



$\mu_1 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$, $\mu_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

we can put μ_1 and μ_2 into the function we get from (a)

$g(x) = (\mu_{11}^2 + \mu_{12}^2 - \mu_{21}^2 - \mu_{22}^2) + (2\mu_{21} - 2\mu_{11})x_1 + (2\mu_{22} - 2\mu_{12})x_2 < 0$

$\Rightarrow g(x) = 3 + 0x_1 + 6x_2 = 0$

$\therefore$ The function of the classifier is that $x_2 = -\frac{1}{2}$

(c) Assume $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mu_1 = \begin{bmatrix} \mu_{11} \\ \mu_{12} \end{bmatrix}$ $\mu_2 = \begin{bmatrix} \mu_{21} \\ \mu_{22} \end{bmatrix}$, $\mu_3 = \begin{bmatrix} \mu_{31} \\ \mu_{32} \end{bmatrix}$

$g_1(x) = \|x - \mu_1\|_2 = \sqrt{(x_1-\mu_{11})^2 + (x_2-\mu_{12})^2}$

$g_2(x) = \|x - \mu_2\|_2 = \sqrt{(x_1-\mu_{21})^2 + (x_2-\mu_{22})^2}$

$g_3(x) = \|x - \mu_3\|_2 = \sqrt{(x_1-\mu_{31})^2 + (x_2-\mu_{32})^2}$

$H_{12}: g_1(x) = g_2(x) \Rightarrow (x_1-\mu_{11})^2 + (x_2-\mu_{12})^2 = (x_1-\mu_{21})^2 + (x_2-\mu_{22})^2$

$\Rightarrow \cancel{x_1^2} - 2\mu_{11}x_1 + \mu_{11}^2 + \cancel{x_2^2} - 2\mu_{12}x_2 + \mu_{12}^2 = \cancel{x_1^2} - 2\mu_{21}x_1 + \mu_{21}^2 + \cancel{x_2^2} - 2\mu_{22}x_2 + \mu_{22}^2$

$\Rightarrow H_{12}: \underbrace{(\mu_{11}^2 + \mu_{12}^2 - \mu_{21}^2 - \mu_{22}^2)}_{w_0} + \underbrace{(2\mu_{21} - 2\mu_{11})}_{w_1}x_1 + \underbrace{(2\mu_{22} - 2\mu_{12})}_{w_2}x_2 = 0$

$\because$ we can use $g(x) = w_0 + w_1x_1 + w_2x_2$ to express $H_{12} \Rightarrow$ it's linear.

Also, use the same method, we can know H_{23} $(g_2(x) = g_3(x))$ and H_{13} $(g_1(x) = g_3(x))$ are also linear.

(d) $H_{12}: (\mu_{11}^2 + \mu_{12}^2 - \mu_{21}^2 - \mu_{22}^2) + (2\mu_{21} - 2\mu_{11})x_1 + (2\mu_{22} - 2\mu_{12})x_2 = 0$

$H_{23}: (\mu_{21}^2 + \mu_{22}^2 - \mu_{31}^2 - \mu_{32}^2) + (2\mu_{31} - 2\mu_{21})x_1 + (2\mu_{32} - 2\mu_{22})x_2 = 0$

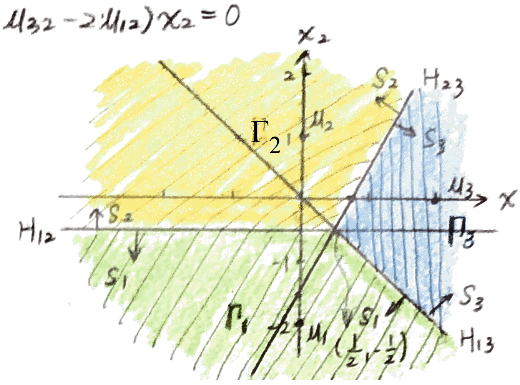
$H_{13}: (\mu_{11}^2 + \mu_{12}^2 - \mu_{31}^2 - \mu_{32}^2) + (2\mu_{31} - 2\mu_{11})x_1 + (2\mu_{32} - 2\mu_{12})x_2 = 0$

$\mu_1 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$, $\mu_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mu_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$\Rightarrow H_{12}: 3 + 6x_2 = 0 \Rightarrow x_2 = -\frac{1}{2}$

$H_{23}: (-3) + 4x_1 - 2x_2 = 0 \Rightarrow x_2 = 2x_1 - \frac{3}{2}$

$H_{13}: 4x_1 + 4x_2 = 0 \Rightarrow x_2 = -x_1$



4. $X = \sum_{i=1}^{n} \alpha_i x_i$, $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$

Assume there are 2 sets of vectors : $S_1 = \{x_1, x_2, \dots, x_n\}$ and $S_2 = \{y_1, y_2, \dots, y_n\}$
Also, S_1 are S_2 are linearly separable, there exists a discriminant function $g(x) = w_0 + w^T x$.

If $g(x) > 0$ means $x \in S_1$,
$g(x) < 0$ means $x \in S_2$
} We assume $g(S_1) > 0$ and $g(S_2) < 0$

Let's look at S_1 : $X = \sum_{i=1}^{n} \alpha_i x_i \Rightarrow g(X) > 0$

$$
\begin{aligned}
g(X) &= w_0 + w^T X \\
&= w_0 + w^T\left(\sum_{i=1}^{n} \alpha_i x_i\right) \\
&= \sum_{i=1}^{n} \alpha_i w_0 + \sum_{i=1}^{n} \alpha_i w^T x_i \quad \left(\because \sum_{i=1}^{n} \alpha_i = 1\right) \\
&= \sum_{i=1}^{n} \alpha_i \underbrace{(w_0 + w^T x_i)}_{g(x_i)} > 0 \Rightarrow g(x_i) > 0,\ x_i \in S_1,\ i = 1, 2, \dots, n
\end{aligned}
$$

$\Rightarrow$ when we put each x_i into $g(x)$, $g(x_i) > 0$

S_2: $Y = \sum_{i=1}^{n} \beta_i y_i \Rightarrow g(Y) < 0$

$$
\begin{aligned}
g(Y) &= w_0 + w^T Y \\
&= w_0 + w^T\left(\sum_{i=1}^{n} \beta_i y_i\right) \\
&= \sum_{i=1}^{n} \beta_i w_0 + \sum_{i=1}^{n} \beta_i \cdot w^T y_i \quad \left(\because \sum_{i=1}^{n} \beta_i = 1\right) \\
&= \sum_{i=1}^{n} \beta_i \underbrace{(w_0 + w^T y_i)}_{g(y_i)} < 0 \Rightarrow g(y_i) < 0,\ y_i \in S_2,\ i = 1, 2, \dots, n
\end{aligned}
$$

$\Rightarrow$ when we put each y_i into $g(x)$, $g(y_i) < 0$

Therefore, if the two sets of vectors are linearly separable, there's no intersection on their convex hulls. Conversely, if there's an intersection on their convex hulls, the two sets of vectors are not linear separable.