Discriminant Functions for Normal Density.

$$f(\underline{x}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})\right\}.$$

Take natural log on both the sides

$$\ln f(\underline{x}) = -\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1}(\underline{x}-\underline{\mu}) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma|$$

$$p(\underline{x}|w_i) \sim N(\mu_i, \Sigma_i).$$

$$p(\underline{X}, w_i) = p(\underline{x}|w_i)\, p(w_i)$$

$$\ln p(\underline{x}, w_i) = \ln p(\underline{x}|w_i) + \ln p(w_i).$$

Therefore:

$$\underbrace{\ln p(\underline{x}, w_i)}_{g_i(\underline{X})} = -\frac{1}{2}(\underline{X}-\underline{\mu}_i)^T \Sigma_i^{-1}(\underline{X}-\underline{\mu}_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma_i| + \ln p(w_i)$$

We will find $\ln p(\underline{x}, w_i)$ for each class and seek the maximum.

Let
$$W_i = -\frac{1}{2}\Sigma_i^{-1}$$
$$\underline{w}_i = \Sigma_i^{-1}\underline{\mu}_i$$
$$w_{i0} = -\frac{1}{2}\underline{\mu}_i^T \Sigma_i^{-1}\underline{\mu}_i - \frac{1}{2}\ln|\Sigma_i| + \ln p(w_i).$$

$$g_i(\underline{x}) = \underline{x}^T W_i \underline{X} + \underline{w}_i^T \underline{x} + w_{i0}$$

$\varepsilon,$

$$-\frac{1}{2}(\underline{x}-\underline{\mu}_i)^T\Sigma_i^{-1}(\underline{x}-\underline{\mu}_i)-\frac{1}{2}\ln|\Sigma_i|+\ln p(\omega_i).$$

$$=-\frac{1}{2}(\underline{x}^T-\underline{\mu}_i^T)\Sigma_i^{-1}(\underline{x}-\underline{\mu}_i)-\frac{1}{2}\ln|\Sigma_i|+\ln p(\omega_i).$$

$$=-\frac{1}{2}\underline{x}^T\Sigma_i^{-1}\underline{x}+\frac{1}{2}\underline{x}^T\Sigma_i^{-1}\underline{\mu}_i+\frac{1}{2}\underline{\mu}_i^T\Sigma_i^{-1}\underline{x}-\frac{1}{2}\underline{\mu}_i^T\Sigma_i^{-1}\mu_i$$
$$-\frac{1}{2}\ln|\Sigma_i|+\ln p(\omega_i).$$

$$=-\frac{1}{2}\underline{x}^T\Sigma_i^{-1}x+\underline{\mu}_i^T\Sigma_i^{-1}\underline{x}-\frac{1}{2}\underline{\mu}_i^T\Sigma_i^{-1}\underline{\mu}_i-\frac{1}{2}\ln|\Sigma_i|+\ln p(\omega_i)$$

$$=x^T\underbrace{\left(-\frac{1}{2}\Sigma_i^{-1}\right)}_{W_i}x+\underbrace{\left(\mu_i^T\Sigma_i^{-1}\right)}_{\underline{\omega}_i^T}\underline{x}\underbrace{-\frac{1}{2}\mu_i^T\Sigma_i^{-1}\mu_i-\frac{1}{2}\ln|\Sigma_i|+\ln p(\omega_i)}_{\omega_{0i}}$$

Ex: class 1: $\begin{pmatrix}2\\6\end{pmatrix}, \begin{pmatrix}4\\6\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}3\\8\end{pmatrix}$.

class 2: $\begin{pmatrix}3\\0\end{pmatrix}, \begin{pmatrix}1\\-2\end{pmatrix}, \begin{pmatrix}5\\-2\end{pmatrix}, \begin{pmatrix}3\\-4\end{pmatrix}$.

$$\mu_1 = \left[\begin{pmatrix}2\\6\end{pmatrix} + \begin{pmatrix}4\\6\end{pmatrix} + \begin{pmatrix}3\\4\end{pmatrix} + \begin{pmatrix}3\\8\end{pmatrix}\right]/4 = \begin{pmatrix}3\\6\end{pmatrix}.$$

$$\mu_2 = \left[\begin{pmatrix}3\\0\end{pmatrix} + \begin{pmatrix}1\\-2\end{pmatrix} + \begin{pmatrix}5\\-2\end{pmatrix} + \begin{pmatrix}3\\-4\end{pmatrix}\right]/4 = \begin{pmatrix}3\\-2\end{pmatrix}.$$

Subtract μ_1 from all samples of class 1.

$$\left.\begin{matrix}\begin{pmatrix}2\\6\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix} = \begin{pmatrix}-1\\0\end{pmatrix}, & \begin{pmatrix}4\\6\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} \\ \begin{pmatrix}3\\4\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix} = \begin{pmatrix}0\\-2\end{pmatrix}, & \begin{pmatrix}3\\8\end{pmatrix} - \begin{pmatrix}3\\6\end{pmatrix} = \begin{pmatrix}0\\2\end{pmatrix}\end{matrix}\right\} \begin{bmatrix}-1 & 1 & 0 & 0\\ 0 & 0 & -2 & 2\end{bmatrix} = X_1$$

Subtract μ_2 from all samples of class 2

$$\left.\begin{matrix}\begin{pmatrix}3\\0\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix} = \begin{pmatrix}0\\2\end{pmatrix}, & \begin{pmatrix}1\\-2\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix} = \begin{pmatrix}-2\\0\end{pmatrix} \\ \begin{pmatrix}5\\-2\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix} = \begin{pmatrix}2\\0\end{pmatrix} & \begin{pmatrix}3\\-4\end{pmatrix} - \begin{pmatrix}3\\-2\end{pmatrix} = \begin{pmatrix}0\\-2\end{pmatrix}\end{matrix}\right\}$$

$$\begin{bmatrix}0 & -2 & 2 & 0\\ 2 & 0 & 0 & -2\end{bmatrix} = X_2$$

Calculation of Σ_i

$$\Sigma_1 = \frac{1}{N} X_1 X_1^T = \frac{1}{4}\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -2 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -2 \\ 0 & 2 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\Sigma_2 = \frac{1}{N} \cancel{X_2} X_2^T = \frac{1}{4}\begin{bmatrix} 0 & -2 & 2 & 0 \\ 2 & 0 & 0 & -2 \end{bmatrix}\begin{bmatrix} 0 & 2 \\ -2 & 0 \\ 2 & 0 \\ 0 & -2 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\Sigma_1^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \qquad |\Sigma_1| = 1$$

$$\Sigma_2^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \qquad |\Sigma_2| = 4.$$

As both class have 4 ~~vectors~~ samples each,

$$p(\omega_1) = \frac{4}{8} = \frac{1}{2}. \qquad p(\omega_2) = \frac{4}{8} = \frac{1}{2}.$$

Let $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$

$$W_1 = -\frac{1}{2}\Sigma_1^{-1} = -\frac{1}{2}\begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1/4 \end{bmatrix}$$

$$W_2 = -\frac{1}{2}\Sigma_2^{-1} = -\frac{1}{2}\begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} -1/4 & 0 \\ 0 & -1/4 \end{bmatrix}$$

$$\underline{w}_1^{*} = \cancel{\mu_1}\,\Sigma_1^{-1}\,\underline{\mu}_1 = \begin{bmatrix}2 & 0\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}3\\ 6\end{bmatrix} = \begin{pmatrix}6\\ 3\end{pmatrix}.$$

$$\underline{w}_2^{*} = \Sigma_2^{-1}\,\underline{\mu}_2 = \begin{bmatrix}1/2 & 0\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}3\\ -2\end{bmatrix} = \begin{bmatrix}3/2\\ -1\end{bmatrix}$$

$$w_{01} = -\frac{1}{2}\,\underline{\mu}_1^{T}\,\underline{\underline{\Sigma_1^{-1}\underline{\mu}_1}} - \frac{1}{2}\ln|\Sigma_1| + \ln p(\omega_1)$$

$$\boxed{-\frac{1}{2}\begin{bmatrix}3 & 6\end{bmatrix}\begin{pmatrix}6\\ 3\end{pmatrix} = -\frac{1}{2}\times 36 = -18}$$

$$w_{01} = -18 - \frac{1}{2}\ln(1) + \ln\frac{1}{2} = -18 - 0.69315 = \cancel{18.3103}\ -18.69315$$

$$w_{02} = -\frac{1}{2}\,\underline{\mu}_2^{T}\,\Sigma_2^{-1}\,\underline{\mu}_2 - \frac{1}{2}\ln|\Sigma_2| + \ln p(\omega_2).$$

$$= -\frac{1}{2}\begin{bmatrix}3 & -2\end{bmatrix}\begin{bmatrix}3/2\\ -1\end{bmatrix} - \frac{1}{2}\ln 4 + \ln\left(\frac{1}{2}\right)$$

$$= -3.25 - 1.3863 - 0.69315 = -5.3294$$

$$g_1(\underline{x}) = \begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}-1 & 0\\ 0 & -1/4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + \begin{pmatrix}6 & 3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} - 18.69.315$$

$$= \begin{pmatrix}-x_1 & -x_2/4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + 6x_1 + 3x_2 - 18.69315$$

$$= -x_1^2 - \cancel{1}0.25\,x_2^2 + 6x_1 + 3x_2 - 18.69315$$

$$g_2(x) = \begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}-\frac{1}{4} & 0\\ 0 & -1/4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} + \begin{pmatrix}3/2 & \cancel{-1}\\ \cancel{-1/2} & \end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix} - 5.3294$$

$$= -\frac{1}{4}x_1^2 - \frac{1}{4}x_2^2 + \frac{3}{2}x_1 - \cancel{1}x_2 - 5.3294$$

Scanned with CamScanner

setting $g_1(\underline{x}) = g_2(\underline{x})$

$$-x_1^2 - \cancel{0.25x_2^2} + 6x_1 + 3x_2 - 18.69315 = -0.25x_1^2 - \cancel{0.25x_2^2} + \tfrac{3}{2}x_1 - 5x_2 - 5.3294$$

$$4x_2 = -0.25x_1^2 + x_1^2 + 1.5x_1 - 6x_1 + \tfrac{3}{2}x_1 - 5.3294 + 18.69315$$

$$4x_2 = 0.75x_1^2 - 4.5x_1 \cancel{- 2.5x_2} + 13.36375$$

$$\cancel{x_2 = 0.25x_1^2 - 1.8333x_1 - 0.8333x_2 + 4.4546}$$

$$x_2 = 0.1875x_1^2 - 1.125x_1 + 3.341$$

Scanned with CamScanner

Let Each class has N_i samples.

m = dimension

N_i = # of samples in class i

C = Total # of classes.

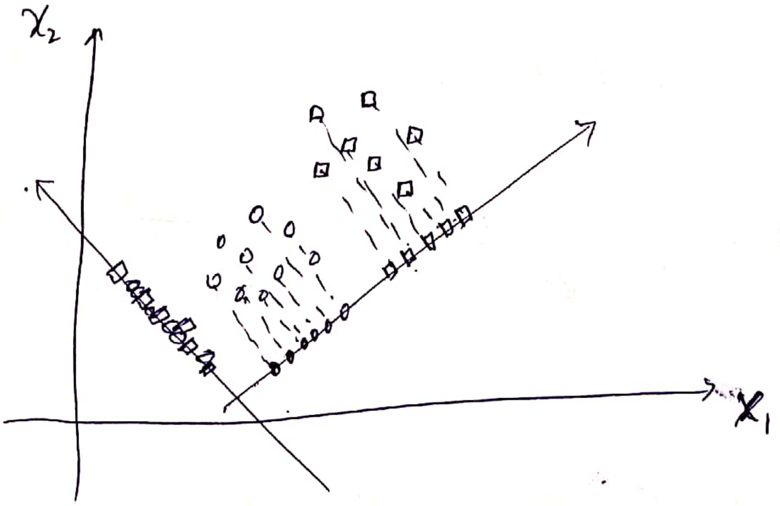
ω_i = class label for class i

$\{\underline{x}^1, \underline{x}^2, \ldots, \underline{x}^{N_i}\}$ are samples for class i.

<u>Data set</u>

X : Contains samples from all classes.
each sample is in a column.

<u>Objective</u>: Find a transformation of X to Y through projecting the samples in X onto a hyperplane with dimension $C-1$.



Case: Two classes

$$X = \{ \underline{x}^1, \underline{x}^2, \cdots, \underline{x}^N \}$$

N_1 of which belong to ω_1

N_2 " " " " ω_2

$$N = N_1 + N_2$$

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \qquad \text{Let } \underline{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}$$

$$y = \underline{w}^T \underline{x}$$

$\underline{w}$ → projection vectors used to project $\underline{x}$ to y.

- Find $\underline{\mu}_1$ and $\underline{\mu}_2$ (means in the $\underline{x}$-feature space)

$$\underline{\mu}_i = \frac{1}{N_i} \sum_{\underline{x} \in \omega_i} \underline{x}_i$$

- Find means in the projected plane.

$$\tilde{\mu}_i = \frac{1}{N_i} \sum_{y \in \omega_i} y = \frac{1}{N_i} \sum_{\underline{x} \in \omega_i} \underline{w}^T \underline{x}_i = \underline{w}^T \frac{1}{N_i} \sum_{x \in \omega_i} \underline{x}_i$$

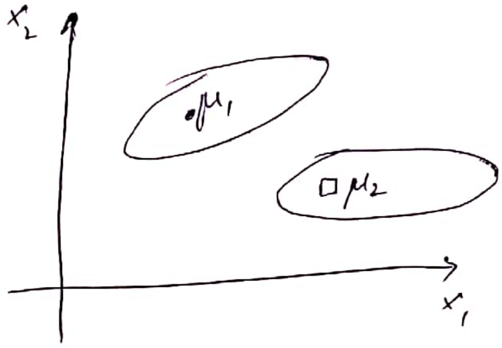
$$= \underline{w}^T \underline{\mu}_i$$

So we are looking for a projection ... each other

Distance b/w projected means

$$J(\underline{w}) = |\tilde{\mu}_1 - \tilde{\mu}_2| = |\underline{w}^T\underline{\mu}_1 - \underline{w}^T\underline{\mu}_2|$$

$$= |\underline{w}^T(\underline{\mu}_1 - \underline{\mu}_2)|$$



This is not a good measure. we should include standard deviation also.

<u>According to Fisher:</u> Maximize a function that represents the difference between the means, normalized by a measure of the within-class variability (scatter).

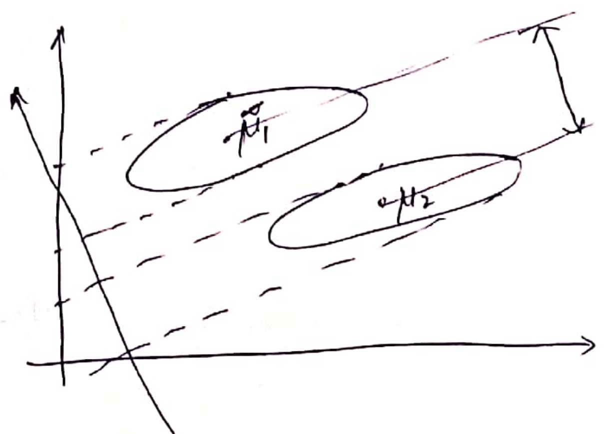
- Scatter of each class $\tilde{S}_i^2$

$$\tilde{S}_i^2 = \sum_{y \in w_i} (y - \tilde{\mu}_i)^2 \quad \text{[variability within class]}.$$

- Variability within two class $\tilde{S}_1^2 + \tilde{S}_2^2$ [within class scatter].

<u>Fisher linear Disriminant:</u>

$$J(\underline{w}) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{S}_1^2 + \tilde{S}_2^2}$$



So we are looking for a projection such that means are far apart but projected close to each other.

Let $\underline{w}^*$ is the optimal projection

Define scatter in $\underline{x}$-space

$$S_i = \sum_{x \in \omega_i} (\underline{x}-\underline{\mu}_i)(\underline{x}-\underline{\mu}_i)^T \quad \text{[covariance matrix of class } \omega_i\text{]}$$

$$S_w = S_1 + S_2 \quad \text{[within-class scatter matrix]}$$

$$\tilde{S}_i^2 = \sum_{y \in \omega_i} (y - \tilde{\mu}_i)^2 = \sum_{x \in \omega_i} (\underline{w}^T \underline{x} - \underline{w}^T \underline{\mu}_i)^2$$

$$= \sum_{x \in \omega_i} \underline{w}^T (\underline{x}-\underline{\mu}_i)(\underline{x}-\underline{\mu}_i)^T \underline{w}$$

$$= \underline{w}^T \left[\sum_{x \in \omega_i} (\underline{x}-\underline{\mu}_i)(\underline{x}-\underline{\mu}_i)^T\right] \underline{w} = \underline{w}^T S_i \underline{w}$$

$$\tilde{S}_1^2 + \tilde{S}_2^2 = \underline{w}^T S_1 \underline{w} + \underline{w}^T S_2 \underline{w} = \underline{w}^T (S_1 + S_2) \underline{w} = \underline{w}^T S_w \underline{w} = \tilde{S}_w$$

$$S_w = S_1 + S_2$$

$\tilde{S}_w$ = within-class scatter matrix of the projected samples y.

$$(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = (\underline{w}^T \underline{\mu}_1 - \underline{w}^T \underline{\mu}_2)^2$$

$$= \underline{w}^T \underbrace{(\underline{\mu}_1 - \underline{\mu}_2)(\underline{\mu}_1 - \underline{\mu}_2)^T}_{S_B} \underline{w} = \underline{w}^T S_B \underline{w} = \tilde{S}_B$$

S_B is called between-class scatter of x-space

$\tilde{S}_B$ " " " " " y-space

Therefore:

$$J(\underline{\omega}) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} = \frac{\underline{\omega}^T S_B \underline{\omega}}{\underline{\omega}^T S_W \underline{\omega}}$$

Difference b/w class means normalized by within class scatter.

Finding max $J(\omega)$.
differentiate it, set to zero

$$\frac{d}{d\underline{\omega}} J(\underline{\omega}) = \frac{d}{d\underline{\omega}} \left(\frac{\underline{\omega}^T S_B \underline{\omega}}{\underline{\omega}^T S_W \underline{\omega}} \right) = 0$$

$$\frac{(\underline{\omega}^T S_W \underline{\omega}) \frac{d}{d\underline{\omega}} (\underline{\omega}^T S_B \underline{\omega}) - (\underline{\omega}^T S_B \underline{\omega}) \frac{d}{d\underline{\omega}} (\underline{\omega}^T S_W \underline{\omega})}{(\underline{\omega}^T S_W \underline{\omega})^2} = 0$$

$$\underline{\omega}^T S_W \underline{\omega} \frac{d}{d\underline{\omega}} (\underline{\omega}^T S_B \underline{\omega}) - (\underline{\omega}^T S_B \underline{\omega}) \frac{d}{d\underline{\omega}} (\underline{\omega}^T S_W \underline{\omega}) = 0$$

$$(\underline{\omega}^T S_W \underline{\omega})(2 S_B \underline{\omega}) - (\underline{\omega}^T S_B \underline{\omega})(2 S_W \underline{\omega}) = 0$$

Divide by $2 \underline{\omega}^T S_W \underline{\omega}$

$$\underbrace{\frac{\underline{\omega}^T S_W \underline{\omega}}{\underline{\omega}^T S_W \underline{\omega}}}_{1} S_B \underline{\omega} - \underbrace{\frac{\underline{\omega}^T S_B \underline{\omega}}{\underline{\omega}^T S_W \underline{\omega}}}_{J(\omega)} S_W \underline{\omega} = 0$$

$$S_B \underline{\omega} - J(\underline{\omega}) S_W \underline{\omega} = 0 \qquad (J(\underline{\omega}) \text{ is a scalar})$$

$$J(\underline{\omega}) S_W \underline{\omega} = S_B \underline{\omega}$$

$$J(\underline{\omega}) \underline{\omega} = S_W^{-1} S_B \underline{\omega}$$

Genearlized eigen value problem

$$S_\omega^{-1} S_B \underline{\omega} = \lambda \underline{\omega} \quad \text{where } \lambda = J(\underline{\omega})$$

This yields

$$\omega^* = \arg\max_{\omega} J(\omega) = \arg\max_{\omega} \left(\frac{\underline{\omega}^T S_B \underline{\omega}}{\underline{\omega}^T S_\omega \underline{\omega}} \right)$$

$$= S_\omega^{-1} (\underline{\mu}_1 - \underline{\mu}_2)$$

This is known as Fisher Linear Discriminant.
[Not exactly a discriminant but a specific choice of direction for projection of data down to one dimension].

Ex :

class 1 : $\begin{pmatrix}4\\2\end{pmatrix}, \begin{pmatrix}2\\4\end{pmatrix}, \begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}3\\6\end{pmatrix}, \begin{pmatrix}4\\4\end{pmatrix}$ → X1=[4,2; 2,4; 2,3; 3,6; 4,4];

class 2 : $\begin{pmatrix}9\\10\end{pmatrix}, \begin{pmatrix}6\\8\end{pmatrix}, \begin{pmatrix}9\\5\end{pmatrix}, \begin{pmatrix}8\\7\end{pmatrix}, \begin{pmatrix}10\\8\end{pmatrix}$ → X2=[9,10; 6,8; 9,5; 8,7; 10,8];

$$\mu_1 = \frac{1}{5}\left[\begin{pmatrix}4\\2\end{pmatrix} + \begin{pmatrix}2\\4\end{pmatrix} + \begin{pmatrix}2\\3\end{pmatrix} + \begin{pmatrix}3\\6\end{pmatrix} + \begin{pmatrix}4\\4\end{pmatrix}\right] = \begin{bmatrix}3\\3.8\end{bmatrix}$$ → Mu1=mean(X1)';

$$\mu_2 = \frac{1}{5}\left[\begin{pmatrix}9\\10\end{pmatrix} + \begin{pmatrix}6\\8\end{pmatrix} + \begin{pmatrix}9\\5\end{pmatrix} + \begin{pmatrix}8\\7\end{pmatrix} + \begin{pmatrix}10\\8\end{pmatrix}\right] = \begin{pmatrix}8.4\\7.6\end{pmatrix}$$ → Mu2=mean(X2)';

$$S_1 = \sum_{x \in \omega_1} (\underline{x}-\underline{\mu}_1)(\underline{x}-\underline{\mu}_1)^T = \left[\begin{pmatrix}4\\2\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^2 + \left[\begin{pmatrix}2\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^2 +$$ → S1=cov(X1);

$$+ \left[\begin{pmatrix}2\\3\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^2 + \left[\begin{pmatrix}3\\6\end{pmatrix} - \begin{pmatrix}3\\3.6\end{pmatrix}\right]^2 + \left[\begin{pmatrix}4\\4\end{pmatrix} - \begin{pmatrix}3\\3.8\end{pmatrix}\right]^2 = \begin{bmatrix}1 & -0.25\\ -0.25 & 2.2\end{bmatrix}$$

$$S_2 = \sum_{x \in \omega_2} (\underline{x}-\underline{\mu}_2)(\underline{x}-\underline{\mu}_2)^T = \left[\begin{pmatrix}9\\10\end{pmatrix}-\begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 + \left[\begin{pmatrix}6\\8\end{pmatrix}-\begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 \rightarrow \boxed{S2 = cov(X2);}$$

$$+\left[\begin{pmatrix}9\\5\end{pmatrix}-\begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 + \left[\begin{pmatrix}8\\7\end{pmatrix}-\begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 + \left[\begin{pmatrix}10\\8\end{pmatrix}-\begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 = \begin{bmatrix}2.3 & -0.05\\ -0.05 & 3.3\end{bmatrix}$$

$$S_w = S_1 + S_2 = \begin{bmatrix}1 & -0.25\\ -0.25 & 2.2\end{bmatrix} + \begin{bmatrix}2.3 & -0.05\\ -0.05 & 3.3\end{bmatrix} = \begin{bmatrix}3.3 & -0.3\\ -0.3 & 5.5\end{bmatrix} \rightarrow \boxed{SW = S1 + S2}$$

$$S_B = (\underline{\mu}_1-\underline{\mu}_2)(\underline{\mu}_1-\underline{\mu}_2)^T = \left[\begin{pmatrix}3\\3.8\end{pmatrix}-\begin{pmatrix}8.4\\7.6\end{pmatrix}\right]\left[\begin{pmatrix}3\\3.8\end{pmatrix}-\begin{pmatrix}8.4\\7.6\end{pmatrix}\right]^2 \rightarrow \boxed{S_B = (Mu1-Mu2)(Mu1-Mu2)';}$$

$$= \begin{pmatrix}-5.4\\-3.8\end{pmatrix}\begin{pmatrix}-5.4 & -3.8\end{pmatrix} = \begin{pmatrix}29.16 & 20.52\\ 20.52 & 14.44\end{pmatrix}$$

$$S_w^{-1} S_B \underline{w} = \lambda \underline{w}$$

So eigen values are:

~~invSw = inv(Sw)~~

invSwbySB = inv(Sw) * SB;

[V, D] = eig(invSwbySB)

~~W = V(:,1)~~

$$\left| S_w^{-1} S_B - \lambda I \right| = 0$$

$$\left| \begin{pmatrix}3.3 & -0.3\\ -0.3 & 5.5\end{pmatrix}^{-1} \begin{pmatrix}29.16 & 20.52\\ 20.52 & 14.44\end{pmatrix} - \lambda \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \right| = 0$$

$$\left| \begin{pmatrix}0.3045 & 0.0166\\ 0.0166 & 0.1827\end{pmatrix} \begin{pmatrix}29.16 & 20.52\\ 20.52 & 14.44\end{pmatrix} - \lambda \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \right| = 0$$

$$\begin{vmatrix}9.2213-\lambda & 6.489\\ 4.2339 & 2.9794-\lambda\end{vmatrix} = 0$$

$$(9.2213-\lambda)(2.9794-\lambda) - 6.489 \times 4.2339 = 0$$

$$\lambda^2 - 12.2007\lambda = 0 \Rightarrow \lambda = 0, 12.2007$$

Hence

$$\begin{pmatrix} 9.2213 & 6.489 \\ 4.2339 & 2.9794 \end{pmatrix} \underline{w} = \underset{\lambda_1}{0} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \Rightarrow \underline{w} = \begin{pmatrix} -0.5755 \\ 0.8178 \end{pmatrix}$$

$$\begin{pmatrix} 9.2213 & 6.489 \\ 4.2339 & 2.9794 \end{pmatrix} \underline{w} = \underbrace{12.2007}_{\lambda_2} \underline{w} \Rightarrow \underline{w} = \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix} = w^*$$

or we can find directly:

$$w^* = S_w^{-1} (\underline{\mu}_1 - \underline{\mu}_2) = \begin{pmatrix} 3.3 & -0.3 \\ -0.3 & 5.5 \end{pmatrix}^{-1} \left[\begin{pmatrix} 3 \\ 3.8 \end{pmatrix} - \begin{pmatrix} 8.4 \\ 7.6 \end{pmatrix} \right] \rightarrow \boxed{w = V(\text{:}, 1)}$$

$$= \begin{pmatrix} 0.3045 & 0.0166 \\ 0.0166 & 0.1827 \end{pmatrix} \begin{pmatrix} -5.4 \\ -3.8 \end{pmatrix} = \begin{pmatrix} 0.9088 \\ 0.4173 \end{pmatrix}$$

<u>Classification Task</u>

```
y1 = X1 * w
y2 = X2 * w
m1 = mean(y1)
m2 = mean(y2)
V1 = var(y1)
V2 = var(y2)
% x is the input whose class is to be decided.
pd1 = 1/sqrt(2*pi*V1) * exp(-0.5*(x-m1)*(x-m1)/V1)
pd2 = 1/sqrt(2*pi*V2) * exp(-0.5*(x-m2)*(x-m2)/V2)
if pd1>pd2
    disp("Class 1")
else
    disp("Class 2")
end
```