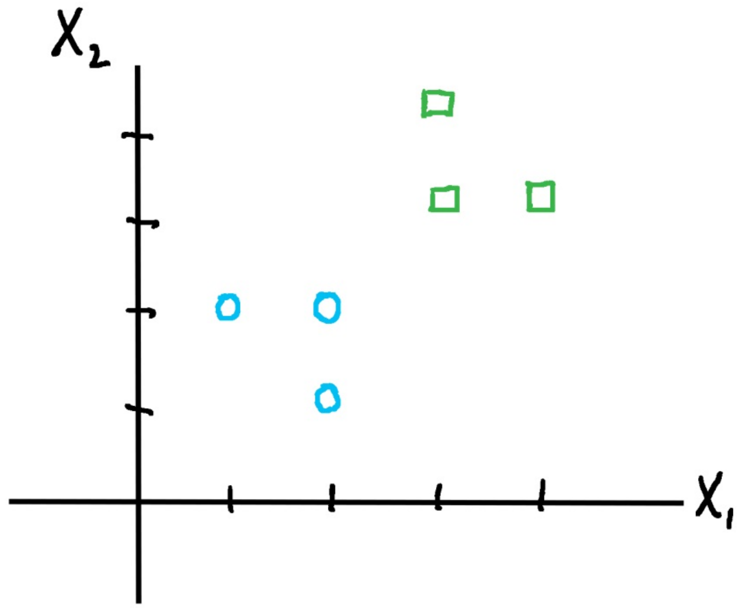
Solution Set: Linear Discriminant Analysis

1. Here is a graph of the data points:



The number of features p is 2, the number of classes K is 2, the total number of data points N is 6, the number N_1 of data points in class k_1 is 3, and the number N_2 of data points in class k_2 is 3.
First, we will find estimates for π_1 and π_2, the prior probabilities that $Y = k_1$ and $Y = k_2$, respectively.
Then, we will find estimates for μ_1 and μ_2, the class-specific mean vectors.
We can then calculate the estimate for the covariance matrix Σ.
Finally, using the estimates $\widehat{\pi_1}, \widehat{\pi_2}, \widehat{\mu_1}, \widehat{\mu_2}, \hat{\Sigma}$, we can find the estimates for the linear discriminant functions $\delta_1(x)$ and $\delta_2(x)$.

$$\widehat{\pi_1} = \frac{N_1}{N} = \frac{3}{6} = \frac{1}{2}$$
$$\widehat{\pi_2} = \frac{N_2}{N} = \frac{3}{6} = \frac{1}{2}$$
$$\widehat{\mu_1} = \frac{1}{N_1} \sum_{i:y_i=1} x_i = \frac{1}{3}[x_1 + x_2 + x_3] = \begin{bmatrix} \frac{5}{3} \\ \frac{5}{3} \end{bmatrix}$$

$$\widehat{\mu_2} = \frac{1}{N_2} \sum_{i:y_i=2} x_i = \frac{1}{3}[x_4 + x_5 + x_6] = \begin{bmatrix} \frac{10}{3} \\ \frac{10}{3} \end{bmatrix}$$

$$\hat{\Sigma} = \frac{1}{N-K} \sum_{k=1}^{K} \sum_{i:y_i=k} (x_i - \widehat{\mu_k})(x_i - \widehat{\mu_k})^T$$

$$= \frac{1}{6-2} \begin{bmatrix} 12/9 & -6/9 \\ -6/9 & 12/9 \end{bmatrix} = \begin{bmatrix} 1/3 & -1/6 \\ -1/6 & 1/3 \end{bmatrix}$$

$$\Longrightarrow \qquad \hat{\Sigma}^{-1} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\widehat{\delta_1}(x) = x^T \hat{\Sigma}^{-1} \widehat{\mu_1} - \frac{1}{2} \widehat{\mu_1}^T \hat{\Sigma}^{-1} \widehat{\mu_1} + \log \widehat{\pi_1}$$
$$= x^T \begin{bmatrix} 10 \\ 10 \end{bmatrix} - \frac{1}{2}\left(\frac{100}{3}\right) + \log \frac{1}{2}$$
$$= 10X_1 + 10X_2 - \frac{50}{3} + \log \frac{1}{2}$$

$$\widehat{\delta_2}(x) = x^T \hat{\Sigma}^{-1} \widehat{\mu_2} - \frac{1}{2} \widehat{\mu_2}^T \hat{\Sigma}^{-1} \widehat{\mu_2} + \log \widehat{\pi_2}$$
$$= x^T \begin{bmatrix} 20 \\ 20 \end{bmatrix} - \frac{1}{2}\left(\frac{400}{3}\right) + \log \frac{1}{2}$$
$$= 20X_1 + 20X_2 - \frac{200}{3} + \log \frac{1}{2}$$

Setting $\widehat{\delta_1}(x) = \widehat{\delta_2}(x)$

$\Longrightarrow \qquad 10X_1 + 10X_2 - \frac{50}{3} + \log \frac{1}{2} = 20X_1 + 20X_2 - \frac{200}{3} + \log \frac{1}{2}$

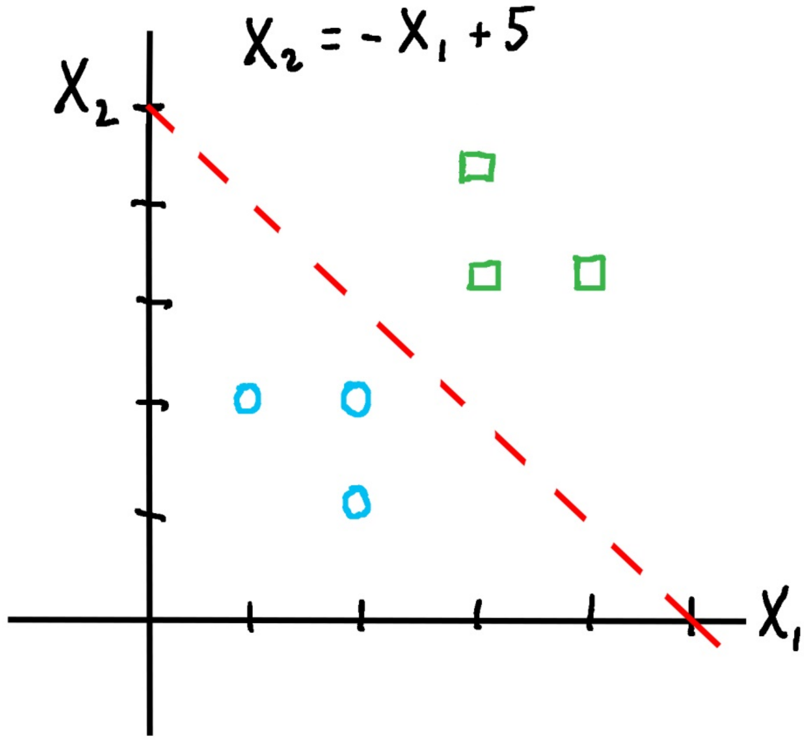
$\Longrightarrow \qquad \frac{150}{3} = 10X_1 + 10X_2$

$\Longrightarrow \qquad 50 = 10X_1 + 10X_2$

$\Longrightarrow \qquad 5 = X_1 + X_2$

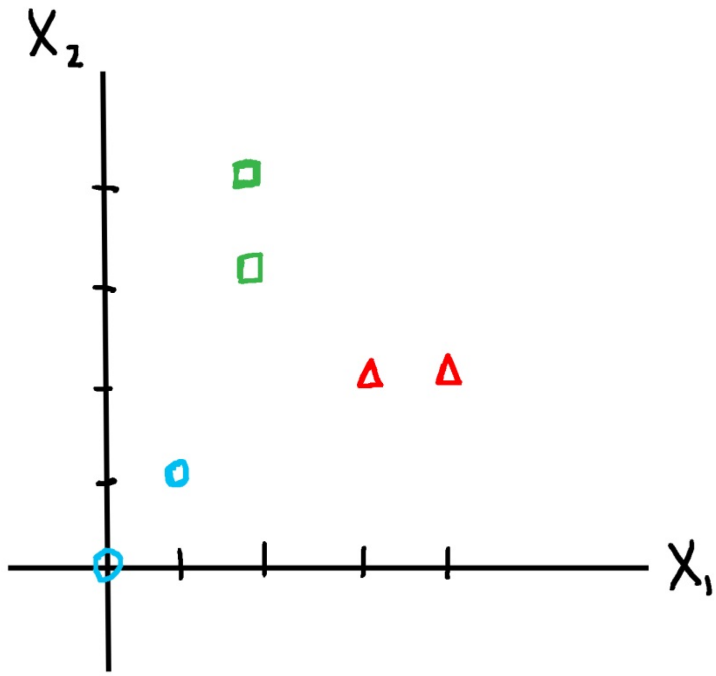
$\Longrightarrow \qquad -X_1 + 5 = X_2$

So, the line that decides between the two classes is given by $X_2 = -X_1 + 5$.
Here is a graph of the decision line:



If $\widehat{\delta_1}(x) > \widehat{\delta_2}(x)$, then we classify x as of class k_1.
So if x is below the line $X_2 = -X_1 + 5$, then we classify x as of class k_1.
Conversely, if $\widehat{\delta_1}(x) < \widehat{\delta_2}(x)$, then we classify x as of class k_2. This corresponds to x being above the line $X_2 = -X_1 + 5$.
The point $(4, 5)$ is above the line; so we classify it as of class k_2.

2. Here is a graph of the data points:



The number of features p is 2, the number of classes K is 3, the total number of data points N is 6, the number N_1 of data points in class k_1 is 2, the number N_2 of data points in class k_2 is 2, and the number N_3 of data points in class k_3 is 2.
First, we will find estimates for π_1, π_2, π_3, the prior probabilities that $Y = k_1, Y = k_2, Y = k_3$, respectively.
Then, we will find estimates for μ_1, μ_2, μ_3, the class-specific mean vectors.
We can then calculate the estimate for the covariance matrix Σ.
Finally, using the estimates $\widehat{\pi_1}, \widehat{\pi_2}, \widehat{\pi_3}, \widehat{\mu_1}, \widehat{\mu_2}, \widehat{\mu_3}, \hat{\Sigma}$, we can find the estimates for the linear discriminant functions $\delta_1(x)$, $\delta_2(x)$, and $\delta_3(x)$.

$$\widehat{\pi_1} = \frac{N_1}{N} = \frac{2}{6} = \frac{1}{3}$$
$$\widehat{\pi_2} = \frac{N_2}{N} = \frac{2}{6} = \frac{1}{3}$$
$$\widehat{\pi_3} = \frac{N_3}{N} = \frac{2}{6} = \frac{1}{3}$$

$$\widehat{\mu_1} = \frac{1}{N_1} \sum_{i:y_i=1} x_i = \frac{1}{2}[x_1 + x_2] = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$
$$\widehat{\mu_2} = \frac{1}{N_2} \sum_{i:y_i=2} x_i = \frac{1}{2}[x_3 + x_4] = \begin{bmatrix} 2 \\ 7/2 \end{bmatrix}$$

$$\widehat{\mu_3} = \frac{1}{N_3}\sum_{i:y_i=3} x_i = \frac{1}{2}[x_5 + x_6] = \begin{bmatrix} 7/2 \\ 2 \end{bmatrix}$$

$$\hat{\Sigma} = \frac{1}{N-K}\sum_{k=1}^{K}\sum_{i:y_i=k}(x_i - \widehat{\mu_k})(x_i - \widehat{\mu_k})^T$$

$$= \tfrac{1}{6-3}\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/6 \\ 1/6 & 1/3 \end{bmatrix}$$

$$\Longrightarrow \qquad \hat{\Sigma}^{-1} = \begin{bmatrix} 4 & -2 \\ -2 & 4 \end{bmatrix}$$

$$\begin{aligned}\widehat{\delta_1}(x) &= x^T\hat{\Sigma}^{-1}\widehat{\mu_1} - \frac{1}{2}\widehat{\mu_1}^T\hat{\Sigma}^{-1}\widehat{\mu_1} + \log\widehat{\pi_1} \\ &= x^T\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \tfrac{1}{2}(1) + \log\tfrac{1}{3} \\ &= X_1 + X_2 - \tfrac{1}{2} + \log\tfrac{1}{3}\end{aligned}$$

$$\begin{aligned}\widehat{\delta_2}(x) &= x^T\hat{\Sigma}^{-1}\widehat{\mu_2} - \frac{1}{2}\widehat{\mu_2}^T\hat{\Sigma}^{-1}\widehat{\mu_2} + \log\widehat{\pi_2} \\ &= x^T\begin{bmatrix} 1 \\ 10 \end{bmatrix} - \tfrac{1}{2}(37) + \log\tfrac{1}{3} \\ &= X_1 + 10X_2 - \tfrac{37}{2} + \log\tfrac{1}{3}\end{aligned}$$

$$\begin{aligned}\widehat{\delta_3}(x) &= x^T\hat{\Sigma}^{-1}\widehat{\mu_3} - \frac{1}{2}\widehat{\mu_3}^T\hat{\Sigma}^{-1}\widehat{\mu_3} + \log\widehat{\pi_3} \\ &= x^T\begin{bmatrix} 10 \\ 1 \end{bmatrix} - \tfrac{1}{2}(37) + \log\tfrac{1}{3} \\ &= 10X_1 + X_2 - \tfrac{37}{2} + \log\tfrac{1}{3}\end{aligned}$$

Setting $\widehat{\delta_1}(x) = \widehat{\delta_2}(x)$

$\Longrightarrow \qquad X_1 + X_2 - \tfrac{1}{2} + \log\tfrac{1}{3} = X_1 + 10X_2 - \tfrac{37}{2} + \log\tfrac{1}{3}$

$\Longrightarrow \qquad 18 = 9X_2$

$\Longrightarrow \qquad 2 = X_2$

So, the line that decides between classes k_1 and k_2 is given by $X_2 = 2$.

Setting $\widehat{\delta_1}(x) = \widehat{\delta_3}(x)$

$\Longrightarrow \qquad X_1 + X_2 - \tfrac{1}{2} + \log\tfrac{1}{3} = 10X_1 + X_2 - \tfrac{37}{2} + \log\tfrac{1}{3}$

$\Longrightarrow \qquad 18 = 9X_1$

$\Longrightarrow \qquad 2 = X_1$

So, the line that decides between classes k_1 and k_3 is given by $X_1 = 2$.

Setting $\widehat{\delta_2}(x) = \widehat{\delta_3}(x)$

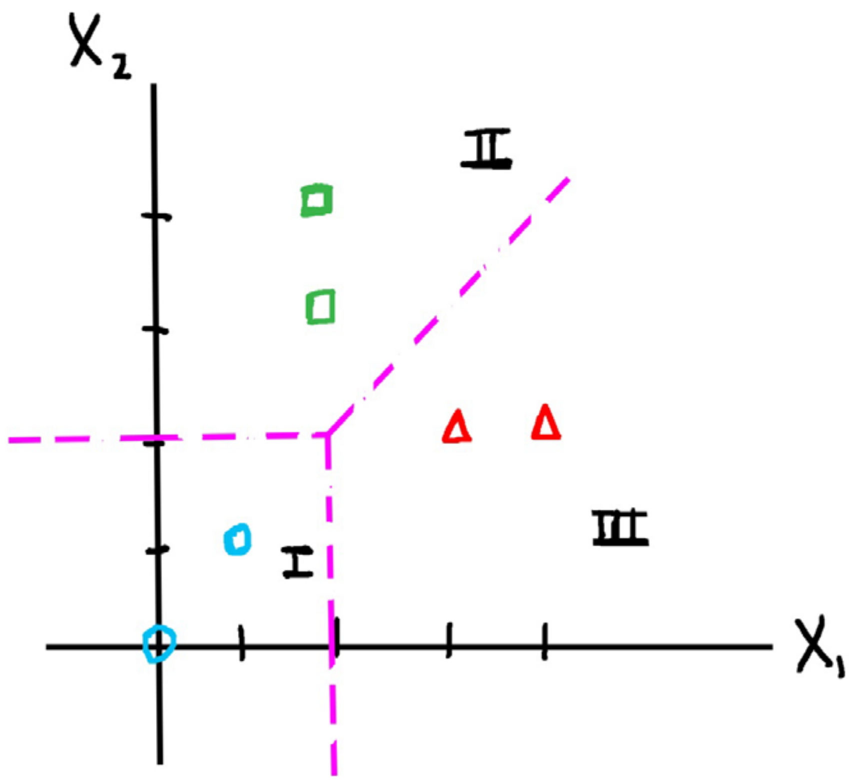
$\Longrightarrow \quad X_1 + 10X_2 - \frac{37}{2} + \log\frac{1}{3} = 10X_1 + X_2 - \frac{37}{2} + \log\frac{1}{3}$

$\Longrightarrow \quad 9X_2 = 9X_1$

$\Longrightarrow \quad X_2 = X_1$

So, the line that decides between classes k_2 and k_3 is given by $X_2 = X_1$.

Here is a graph of the decision lines:



The lines divide the plane into 3 regions.
If x is in region I, then we classify x as of class k_1. Similarly, points in region II get classified as of k_2, and points in region III get classified as of k_3.
The point $(3, 0)$ is in region III; so we classify it as of class k_3.