QDA

The process of constructing a QDA classifier is similar to that of constructing an LDA classifier.

- Using Bayes' rule we can express the QDA formula as:

$$f_i = -\frac{1}{2}\boldsymbol{x}\boldsymbol{C}_k^{-1}\boldsymbol{x}^T + \boldsymbol{x}\boldsymbol{C}_k^{-1}\boldsymbol{\mu}_k^T - \frac{1}{2}\boldsymbol{\mu}_k\boldsymbol{C}_k^{-1}\boldsymbol{\mu}_k^T - \log\left|\boldsymbol{C}_k^{-1}\right| + \log(\pi_k)$$

  Note: This is equivalent to equation (4.23) on page 149 in our textbook.

  $\mu_{\mathbf{k}}$ is the vector of means for class k. This is a row vector
  $\mathbf{C_k^{-1}}$ is the inverse of the covariance matrix for class k. Its dimension is the number of classes we have.
  $\mathbf{x_k^T}$ is transpose of the vector of data (predictors) for class k.
  $\log(\pi_k)$ is the log of the prior probabilities (number of samples of each class)

- We assign object k to group k that has maximum f_k.

- We have a set of data from two classes. Since the data from each sample point is organized in a row. We will use row vectors instead of column vectors.

  - Let our class 1 samples be represented in the following matrix:

$$\mathbf{X_1} = \begin{bmatrix} 1 & 9 \\ 4 & 9 \\ 7 & 8 \\ 9 & 8 \\ 8 & 6 \\ 10 & 12 \\ 11 & 11 \\ 3 & 11 \\ 1 & 10 \\ 10 & 4 \\ 9 & 2 \\ 13 & 8 \end{bmatrix}$$

    with

$$\mu_{\mathbf{1}} = \begin{bmatrix} \frac{43}{6} & \frac{49}{6} \end{bmatrix}$$

– Let our class 2 samples be represented in the following matrix:

$$\mathbf{X_2} = \begin{bmatrix} 2 & 6 \\ 3 & 7 \\ 5 & 6 \\ 3 & 3 \\ 4 & 1 \\ 7 & 4 \\ 7 & 1 \\ 4 & 4 \\ 2 & 3 \\ 1 & 1 \\ 2 & 2 \\ 5 & 4 \end{bmatrix}$$

with

$$\mu_{\mathbf{2}} = \begin{bmatrix} \frac{15}{4} & \frac{7}{2} \end{bmatrix}$$

– Our overall mean is

$$\begin{bmatrix} \frac{131}{24} & \frac{35}{6} \end{bmatrix}$$

– We normalize everything to have a mean of 0 by subtracting the mean of each class.

$$\mathbf{X_1^0} = \begin{bmatrix} -\frac{107}{24} & \frac{19}{6} \\ -\frac{35}{24} & \frac{19}{6} \\ \frac{37}{24} & \frac{13}{6} \\ \frac{85}{24} & \frac{13}{6} \\ \frac{61}{24} & \frac{1}{6} \\ \frac{109}{24} & -\frac{11}{6} \\ \frac{85}{24} & -\frac{23}{6} \end{bmatrix}$$

$$\mathbf{X_2^0} = \begin{bmatrix} -\frac{83}{24} & \frac{1}{6} \\ -\frac{59}{24} & \frac{7}{6} \\ -\frac{11}{24} & \frac{1}{6} \\ -\frac{59}{24} & -\frac{17}{6} \\ -\frac{35}{24} & -\frac{29}{6} \\ \frac{37}{24} & -\frac{11}{6} \\ \frac{37}{24} & -\frac{29}{6} \end{bmatrix}$$

- The covariance matrix for each class is:

$$\boldsymbol{C}_i = \frac{\left(\boldsymbol{X}_i^0\right)^T \boldsymbol{X}_i^0}{n_i}$$

- The covariance matrix for class 1 is

$$\boldsymbol{C}_1 = \begin{bmatrix} \frac{44095}{4032} & -\frac{4205}{1008} \\ -\frac{4205}{1008} & \frac{1711}{252} \end{bmatrix}$$

- The covariance matrix for class 2 is

$$\boldsymbol{C}_2 = \begin{bmatrix} \frac{17935}{4032} & \frac{31}{1008} \\ \frac{31}{1008} & \frac{2143}{252} \end{bmatrix}$$

- The overall Covariance is given as

$$C(r,s) = \frac{1}{n}\sum_{i=1}^{g} n_i \cdot c_i(r,s)$$

  Note: We do not use the overall Covariance matrix for QDA.

  The inverse of the covariance matrices:

$$\boldsymbol{C}_1^{-1} = \begin{bmatrix} \frac{3304}{27665} & \frac{406}{5533} \\ \frac{406}{5533} & \frac{61733}{320914} \end{bmatrix}$$

$$\boldsymbol{C}_2^{-1} = \begin{bmatrix} \frac{60004}{266901} & -\frac{217}{266901} \\ -\frac{217}{266901} & \frac{125545}{1067604} \end{bmatrix}$$

- Our p matrix is:

$$p = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$$

- The model is now trained. Lets look at some examples

$$x_1 = \begin{bmatrix} 1 & 7 \end{bmatrix}$$
$$f_1 = \begin{bmatrix} -1.60256465536242665 \end{bmatrix}$$
$$f_2 = \begin{bmatrix} -0.454833849878940377 \end{bmatrix}$$

This is class 2

$$x_2 = \begin{bmatrix} 4 & 6 \end{bmatrix}$$

$$f_1 = \begin{bmatrix} -0.226711311475901 \end{bmatrix}$$
$$f_2 = \begin{bmatrix} 0.749350473147201 \end{bmatrix}$$

This is class 2

$$x_3 = \begin{bmatrix} 4 & 5 \end{bmatrix}$$

$$f_1 = \begin{bmatrix} -0.972051184463682 \end{bmatrix}$$

$$f_2 = \begin{bmatrix} 0.984337415871282 \end{bmatrix}$$

This is class 2

$$x_4 = \begin{bmatrix} 8 & 1 \end{bmatrix}$$

$$f_1 = \begin{bmatrix} -3.21623747736458 \end{bmatrix}$$

$$f_2 = \begin{bmatrix} -1.28315315179238 \end{bmatrix}$$

This is class 2

$$x_5 = \begin{bmatrix} 8 & 4 \end{bmatrix}$$

$$f_1 = \begin{bmatrix} -0.129457218479025 \end{bmatrix}$$

$$f_2 = \begin{bmatrix} -0.920001646177942 \end{bmatrix}$$

This is class 1

$$x_6 = \begin{bmatrix} 5 & 7 \end{bmatrix}$$

$$f_1 = \begin{bmatrix} 0.730346965819571 \end{bmatrix}$$

$$f_2 = \begin{bmatrix} 0.231000972770656 \end{bmatrix}$$

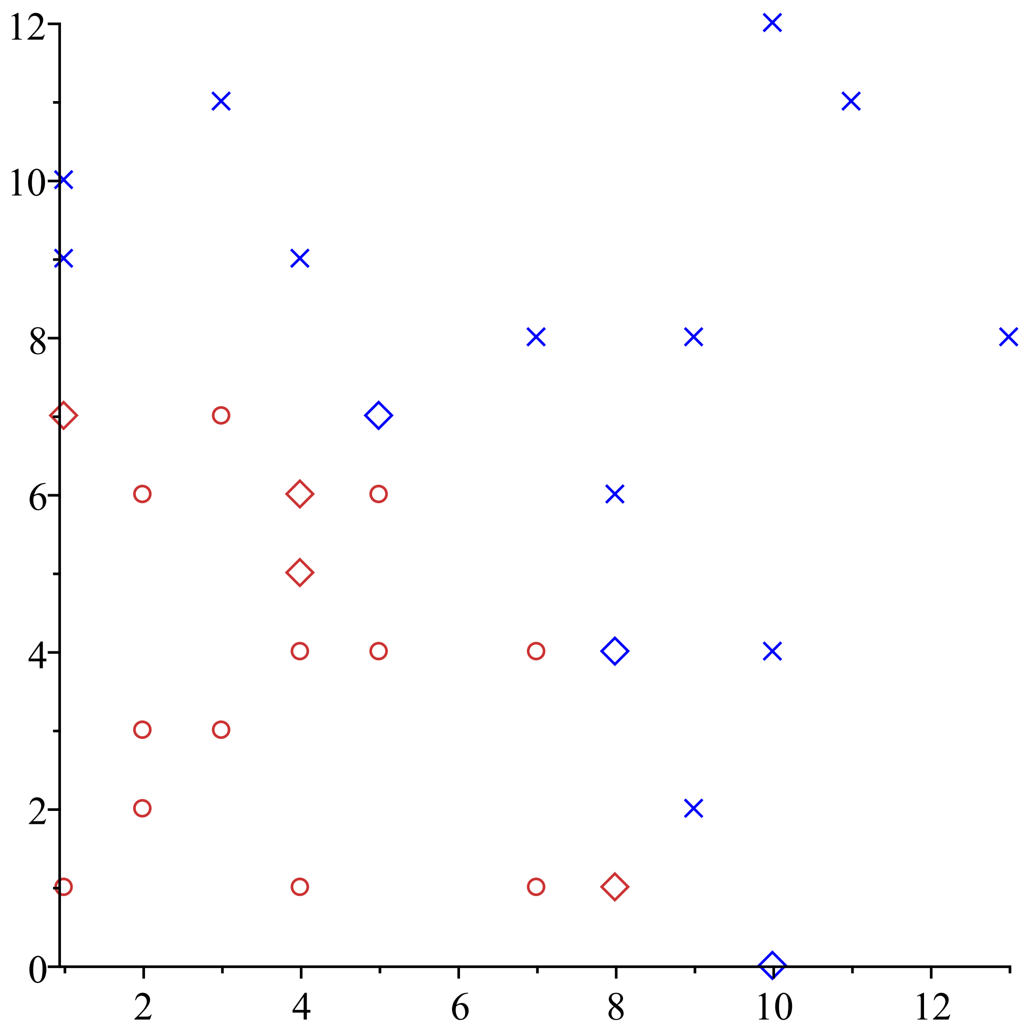
This is class 1

$$x_7 = \begin{bmatrix} 10 & 0 \end{bmatrix}$$

$$f_1 = \begin{bmatrix} -3.86929655238156 \end{bmatrix}$$

$$f_2 = \begin{bmatrix} -4.00566824165716 \end{bmatrix}$$

This is class 1



12
10
8
6
4
2
0
2
4
6
8
10
12