In the Bayesian decision rule, we use the axiom of maximum posterior probability for the problem of classification. Hence, we use a discriminant function to classify (separate) the feature space into different decision regions. The property of a discriminant function is that it must be monotonically increasing w.r.t maximum posterior probability function.

These different regions are separated by decision boundaries. The shape & characteristics of this boundary depend upon various factors like variance, the covariance of the distributions of the concerned classes. Hence we can think of 3 possible cases:

1. Where covariance of classes follows $\sum_i = \sigma^2 I$: This means that features of all classes have equal variance with non-diagonal elements 0. (Statistically independent)
2. Where covariance of classes follows $\sum_i = \sum$: This means that all classes have equal covariance matrices
3. Where covariance of classes follows $\sum_i = \sum_i$: This means that all classes have different covariance matrices

Solving the discriminant function for the above three cases gives us different kinds of decision boundaries. Linear Discriminant Analysis for cases (1 & 2) will give us a **Linear** boundary & LDA for case(3) will give us a **Quadratic** boundary.

These different boundaries occur due to the fact that we solve the discriminant functions of concerned classes using the above assumptions for each case.

Where general discriminant function equation (considering Gaussian distribution) is given by the equation:

$$g_i(x) = log\, p(x|w_i) + logP(w_i)$$

$$\Downarrow$$

$$g_i(x) = -\tfrac{1}{2}(x - \mu_i)' \sum_i^{-1} (x - \mu_i) - d/2 * log(2\pi) - 1/2 * log(|\sum_i|) + log(P(w_i))$$

We solve $g_i(x)$ for all classes by considering the assumptions & ignoring the values that don’t depend on any class. Hence we get different boundaries (i.e. Linear or Quadratic).