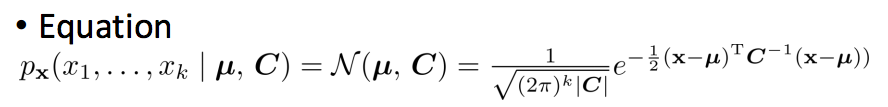
Report for mid-point assessment

In this coursework, we need to build a classifier which can classifying images. One of the process is to establish Gaussian mixture model(GMM), to build this, we need to use multivariate Gaussian Distribution which describes the probabilities for given data around the mean for each classes.

For this mid-point assessment, we are supposed to find the parameters of Gaussian Distribution from given data. According to:



We need to write function that compute mean µ and covariance matrix C. To achieve this, First, I wrote the getFeatures function to obtain the absolute value of N of the Fourier coefficients. By looping over the train folder with getFeatures, I get the feature vector, and the feature vector is the data matrix that used to compute mean and covariance matrix. Based on this, I write calcCov and calcMean to compute the parameters of Gaussian Distribution.

For a training image, the use of chain code will trace the boundary of the image shape in binary which results in a jagged boundary. The noise of the boundary is usually in high frequency, to avoid them, we need Fourier transform and apply top hat filter. This is how getFeatures work.

The way I compute mean is straightforward, get the sum and divided by its length. To compute the covariance matrix C, I refered to the equation:

**C = x'x ( 1 / n )**

Where

**x = X - 11'X ( 1 / n )**

11’ is an n\*n identity matrix, X is the data matrix and n is the number of the data element.

The test result of above functions can be observed in matlab script named coursework1.m.