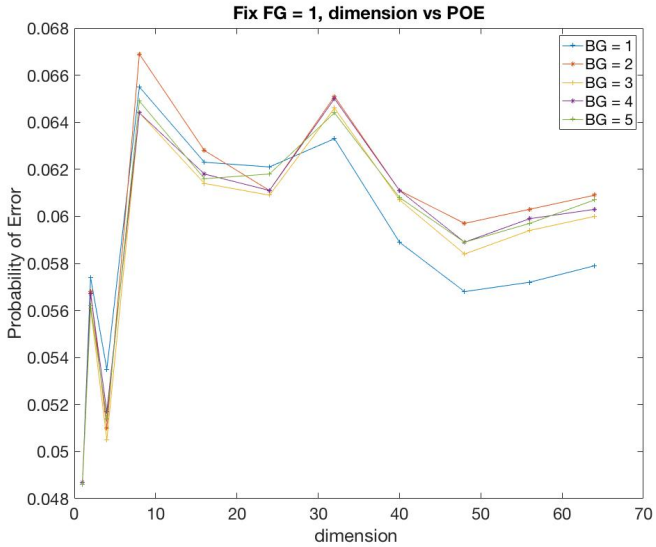
ECE 271A hw5

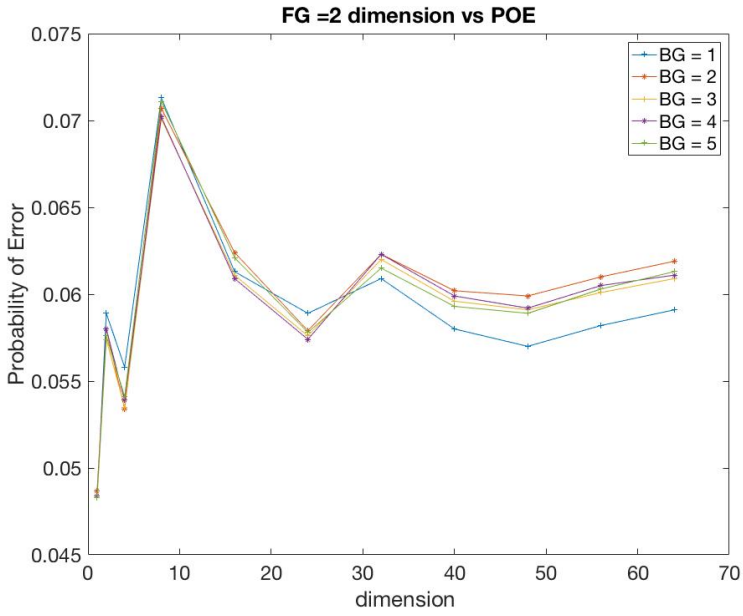
Question Part A:

For each class, learn 5 mixtures of C = 8 components, using a random initialization (recall that the mixture weights must add up to one). Plot the probability of error vs. dimension for each of the 25 classifiers obtained with all possible mixture pairs. Comment the dependence of the probability of error on the initialization.

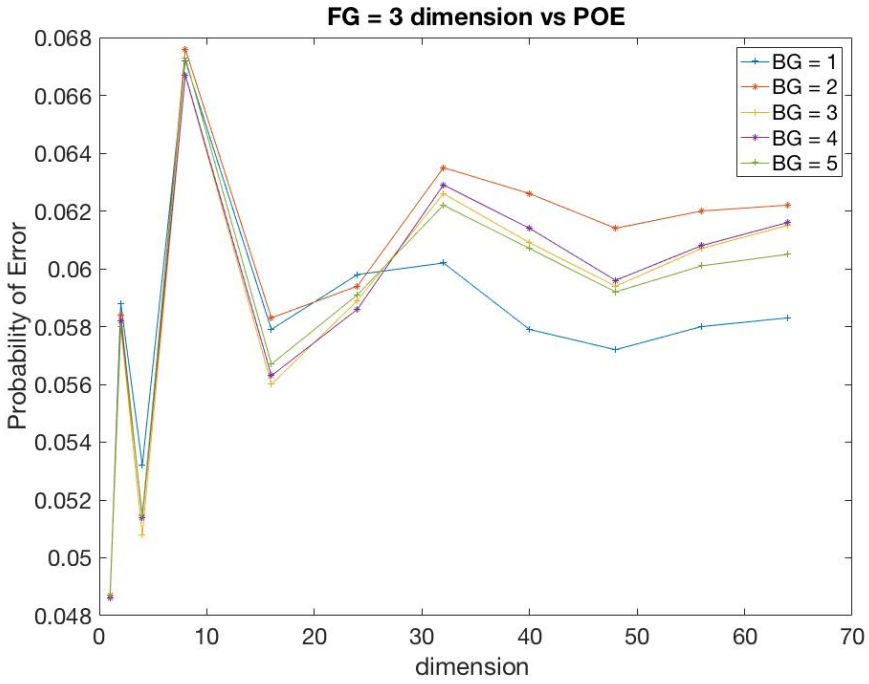
Solution: First I randomly generate 5 set of (mu_FG, sigma_FG, p_FG). I set the index in the set as {1, 2, 3, 4, 5}. Similarly, generate 5 set of (mu_BG, sigma_BG, p_BG). Then apply EM algorithm to find the estimate set of (mu_FG, sigma_FG, p_FG) and (mu_BG, sigma_BG, p_BG). Fix FG = 1, go BG from 1 : 5. The PoE vs dimension is as below:



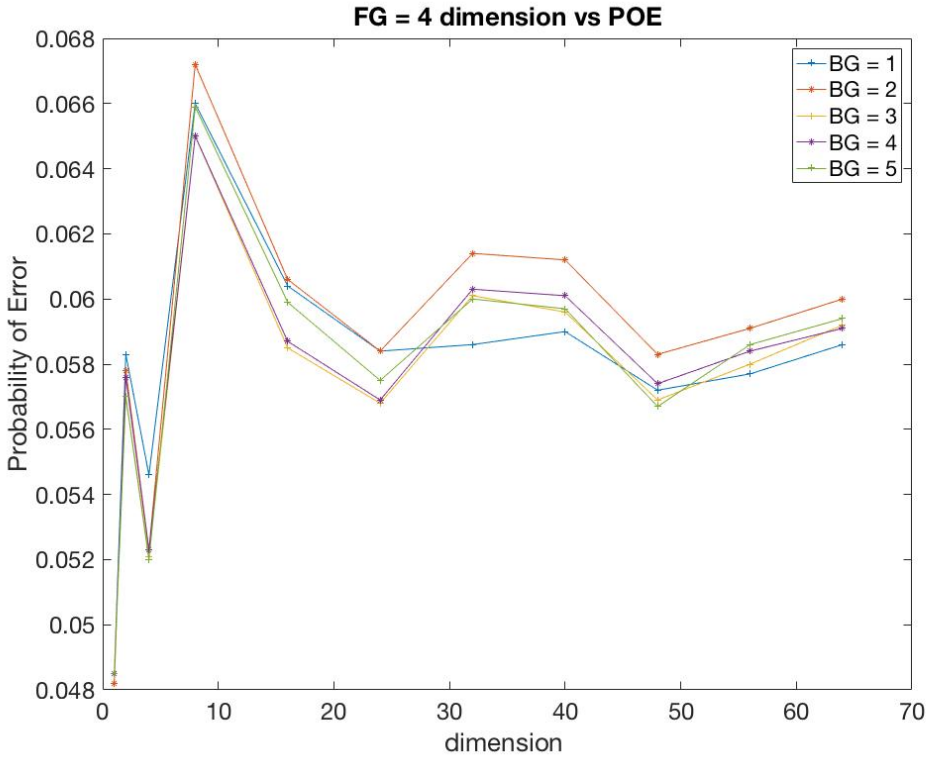
Fix FG = 2, go BG from 1 : 5. The PoE vs dimension is as below:



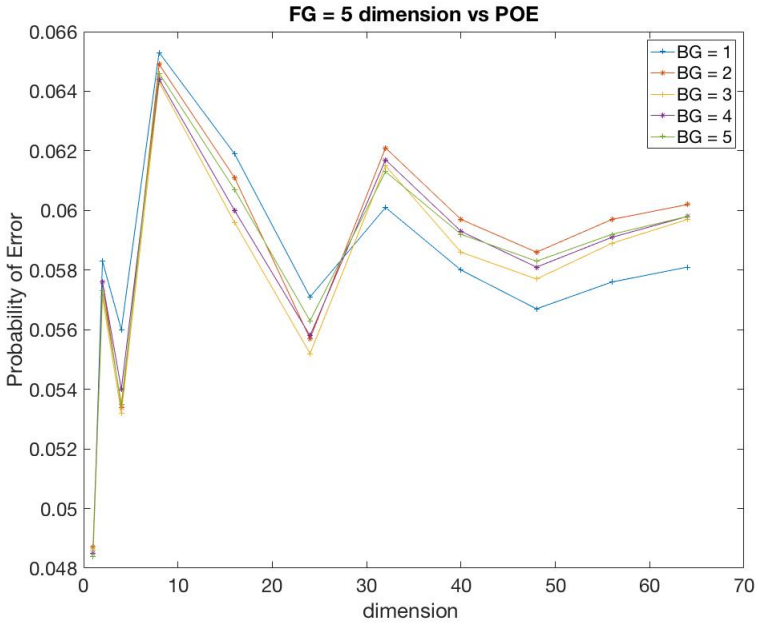
Fix FG = 3, go BG from 1 : 5. The PoE vs dimension is as below:



Fix FG = 4, go BG from 1 : 5. The PoE vs dimension is as below:



Fix FG = 5, go BG from 1 : 5. The PoE vs dimension is as below:

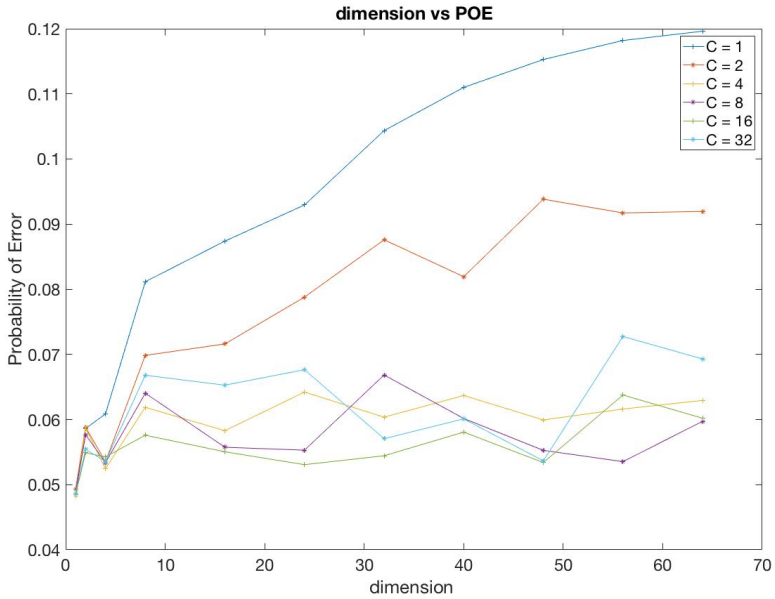


Comment: From the five picture we can generally say that when BG = 1, and FG = 5, the PoE performs best. And when the dim = 1, from all five picture, the PoE reach min, when dim = 2 the PoE increase, then when dim = 4 it decrease. After dim > 48 the PoE tends to be stable, which turns to a almost parallel line.

Lastly, attach on result of picture when we take FG = 1, BG = 1, dim = 1 as below:

Question Part B:
For each class, learn mixtures with C ∈{1, 2, 4, 8, 16, 32}. Plot the probability of error vs. dimension for each number of mixture components. What is the effect of the number of mixture components on the probability of error?
Solution: The following below picture is the EM algorithm by component C = 1 and feature = 1, the other picture is omitted.
the probability of error for EM method with parameter C= 1 and dim = 1 is: 0.049368

The probability of error vs. dimension for each number of mixture components:



From the picture, we can see: when the number of component is small. E.g. when C = {1, 2}. The POE is almost monotonically increasing. It is because, the first feature will make an great impact on picture. When the number of component is large. E.g. C = {4, 8, 16, 24, 32, 40, 48, 56, 64}. The POE will be varies small when the dimension increase. Overall, from the picture, we can see, the more components we have, the more stable our classifier is.