CSE 5522 Homework #4

Cheng Zhang
Dept. Computer Science and Engineering, The Ohio State University
zhang.7804@osu.edu

1 Question 1

For the first iteration:

E-step:

$$P(B = \{1, 2\}|Y, H) = \alpha P(Y|B) \cdot P(H|B) \cdot P(B)$$
$$= \alpha < 0.3, 0.7 > \cdot < 0.6, 0.4 > \cdot < 0.5, 0.5 > = < 0.39, 0.61 >$$

$$P(B = \{1, 2\}|Y, S) = \alpha P(Y|B) \cdot P(S|B) \cdot P(B)$$
$$= \alpha < 0.3, 0.7 > \cdot < 0.4, 0.6 > \cdot < 0.5, 0.5 > = < 0.23, 0.77 >$$

$$P(B = \{1, 2\}|R, H) = \alpha P(R|B) \cdot P(H|B) \cdot P(B)$$
$$= \alpha < 0.7, 0.3 > \cdot < 0.6, 0.4 > \cdot < 0.5, 0.5 > =< 0.77, 0.23 >$$

$$P(B = \{1, 2\}|R, S) = \alpha P(R|B) \cdot P(S|B) \cdot P(B)$$
$$= \alpha < 0.7, 0.3 > \cdot < 0.4, 0.6 > \cdot < 0.5, 0.5 > =< 0.61, 0.39 >$$

where:

$$P(B) = P(B|Y, H) \cdot P(Y, H) + P(B|Y, S) \cdot P(Y, S) + P(B|R, H) \cdot P(R, H) + P(B|R, S) \cdot P(R, S)$$
$$= < 0.39, 0.61 > \cdot 0.2 + < 0.23, 0.77 > \cdot 0.2 + < 0.77, 0.23 > \cdot 0.4 + < 0.61, 0.39 > \cdot 0.2$$
$$= < 0.554, 0.446 >$$

M-step:

$$P(C = \{Y, R\}|B = 1) = \alpha P(B = 1|C) \cdot P(C)$$
$$= \alpha < 0.39 * 2 + 0.23 * 2, 0.77 * 4 + 0.61 * 2 >$$
$$= \alpha < 1.24, 4.3 > = < 0.22, 0.78 >$$

$$P(C = \{Y, R\}|B = 2) = \alpha P(B = 2|C) \cdot P(C)$$
$$= \alpha < 0.61 * 2 + 0.77 * 2, 0.23 * 4 + 0.39 * 2 >$$
$$= \alpha < 2.76, 1.7 > = < 0.62, 0.38 >$$

$$P(S = \{H, S\}|B = 1) = \alpha P(B = 1|S) \cdot P(S)$$
$$= \alpha < 0.39 * 2 + 0.77 * 4, 0.23 * 2 + 0.61 * 2 >$$
$$= \alpha < 3.86, 1.68 > = < 0.70, 0.30 >$$

$P(S = \{H, S\}|B = 2) = \alpha P(B = 2|S) \cdot P(S)$

$= \alpha < 0.61 * 2 + 0.23 * 4, 0.77 * 2 + 0.39 * 2 >$

$= \alpha < 2.14, 2.32 > = < 0.48, 0.52 >$

Therefore:

$P(C)\_iter1 = < 0.3984, 0.6016 >$

$P(S)\_iter1 = < 0.6019, 0.3981 >$

For the second iteration:

E-step:

$P(B = \{1, 2\}|Y, H) = \alpha P(Y|B) \cdot P(H|B) \cdot P(B)$

$= \alpha < 0.22, 0.62 > \cdot < 0.7, 0.48 > \cdot < 0.554, 0.446 > = < 0.39, 0.61 >$

$P(B = \{1, 2\}|Y, S) = \alpha P(Y|B) \cdot P(S|B) \cdot P(B)$

$= \alpha < 0.22, 0.62 > \cdot < 0.3, 0.52 > \cdot < 0.554, 0.446 > = < 0.20, 0.80 >$

$P(B = \{1, 2\}|R, H) = \alpha P(R|B) \cdot P(H|B) \cdot P(B)$

$= \alpha < 0.78, 0.38 > \cdot < 0.7, 0.48 > \cdot < 0.554, 0.446 > = < 0.79, 0.21 >$

$P(B = \{1, 2\}|R, S) = \alpha P(R|B) \cdot P(S|B) \cdot P(B)$

$= \alpha < 0.78, 0.38 > \cdot < 0.3, 0.52 > \cdot < 0.554, 0.446 > = < 0.60, 0.40 >$

where:

$P(B) = P(B|Y, H) \cdot P(Y, H) + P(B|Y, S) \cdot P(Y, S) + P(B|R, H) \cdot P(R, H) + P(B|R, S) \cdot P(R, S)$

$= < 0.39, 0.61 > \cdot 0.2 + < 0.20, 0.80 > \cdot 0.2 + < 0.79, 0.21 > \cdot 0.4 + < 0.60, 0.40 > \cdot 0.2$

$= < 0.554, 0.446 >$

M-step:

$P(C = \{Y, R\}|B = 1) = \alpha P(B = 1|C) \cdot P(C)$

$= \alpha < 0.39 * 2 + 0.20 * 2, 0.79 * 4 + 0.60 * 2 >$

$= \alpha < 1.18, 4.36 > = < 0.21, 0.79 >$

$P(C = \{Y, R\}|B = 2) = \alpha P(B = 2|C) \cdot P(C)$

$= \alpha < 0.61 * 2 + 0.80 * 2, 0.21 * 4 + 0.40 * 2 >$

$= \alpha < 2.82, 1.64 > = < 0.63, 0.37 >$

$P(S = \{H, S\}|B = 1) = \alpha P(B = 1|S) \cdot P(S)$

$= \alpha < 0.39 * 2 + 0.79 * 4, 0.20 * 2 + 0.60 * 2 >$

$= \alpha < 3.94, 1.6 > = < 0.71, 0.29 >$

$P(S = \{H, S\}|B = 2) = \alpha P(B = 2|S) \cdot P(S)$

$= \alpha < 0.61 * 2 + 0.21 * 4, 0.80 * 2 + 0.40 * 2 >$

$= \alpha < 2.06, 2.4 > = < 0.46, 0.54 >$

Therefore:

P(C)_iter2 $= < 0.39732, 0.60268 >$

P (S)_iter2 $= < 0.5985, 0.4015 >$

For the third iteration:

E-step:

$P(B = \{1, 2\} | Y, H) = \alpha P(Y | B) \cdot P(H | B) \cdot P(B)$

$= \alpha < 0.21, 0.63 > \cdot < 0.71, 0.46 > \cdot < 0.554, 0.446 >$

$= < 0.39, 0.61 >$

$P(B = \{1, 2\} | Y, S) = \alpha P(Y | B) \cdot P(S | B) \cdot P(B)$

$= \alpha < 0.21, 0.63 > \cdot < 0.29, 0.54 > \cdot < 0.554, 0.446 >$

$= < 0.18, 0.82 >$

$P(B = \{1, 2\} | R, H) = \alpha P(R | B) \cdot P(H | B) \cdot P(B)$

$= \alpha < 0.79, 0.37 > \cdot < 0.71, 0.46 > \cdot < 0.554, 0.446 >$

$= < 0.80, 0.20 >$

$P(B = \{1, 2\} | R, S) = \alpha P(R | B) \cdot P(S | B) \cdot P(B)$

$= \alpha < 0.79, 0.37 > \cdot < 0.29, 0.54 > \cdot < 0.554, 0.446 >$

$= < 0.59, 0.41 >$

M-step:

$P(C = \{Y, R\} | B = 1) = \alpha P(B = 1 | C) \cdot P(C) = \alpha < 0.39 * 2 + 0.18 * 2, 0.80 * 4 + 0.59 * 2 >$

$= \alpha < 1.14, 4.38 > = < 0.20, 0.80 >$

$P(C = \{Y, R\} | B = 2) = \alpha P(B = 2 | C) \cdot P(C) = \alpha < 0.61 * 2 + 0.82 * 2, 0.20 * 4 + 0.41 * 2 >$

$= \alpha < 2.86, 1.62 > = < 0.64, 0.36 >$

$P(S = \{H, S\} | B = 1) = \alpha P(B = 1 | S) \cdot P(S) = \alpha < 0.39 * 2 + 0.80 * 4, 0.18 * 2 + 0.59 * 2 >$

$= \alpha < 3.98, 1.54 > = < 0.72, 0.28 >$

$P(S = \{H, S\} | B = 2) = \alpha P(B = 2 | S) \cdot P(S) = \alpha < 0.61 * 2 + 0.20 * 4, 0.82 * 2 + 0.41 * 2 >$

$= \alpha < 2.02, 2.46 > = < 0.45, 0.55 >$

Therefore:

P(C)_iter3 $= < 0.3971, 0.6029 >$

P (S)_iter3 $= < 0.5990, 0.4010 >$

During the iteration procedure, we can notice that the probabilities trend to be very close. However, we cannot say that EM algorithm achieves convergence after three iterations. Because we should run the algorithm more times to see whether the probabilities tend to be stable.

2 Question 2

2.1 Part A

In this section, we use a single Gaussian per vowel to create a classifier. Particularly , the Gaussian models have diagonal covariance. The average accuracy around 0.49.

Fig. 1 shows the classification accuracy across different subset: we can find that ah and ax fail to get 0 classification result.

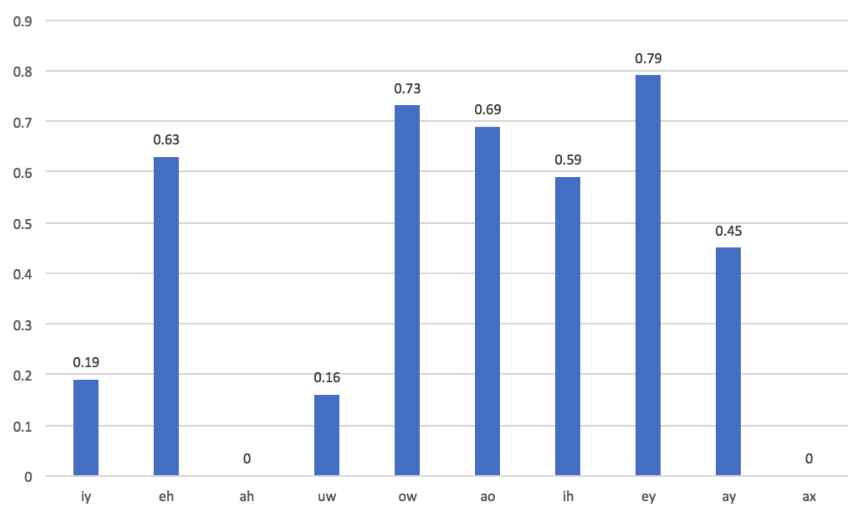

Figure 1

2.2 Part B

Now we will model P(Formant|Vowel) using a Mixture of Gaussians (MoG), which means introducing a vowel-dependent mixture class. Assume for the moment 2 classes per vowel. We run the EM algorithm for different iterations for 5 to 50 and get the average classification accuracy is as figure. 2.

Figure. 3 shows the classification accuracy across different subset under EM algorithm with iteration #50. We can find that ah also fails to get 0 classification result. But ax subset get improvement to 0.25 accuracy.

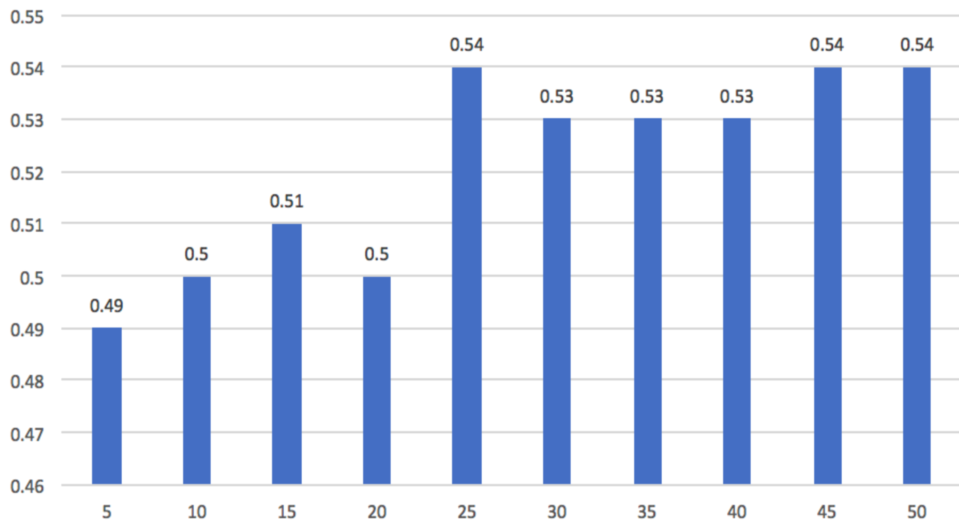

Figure 2

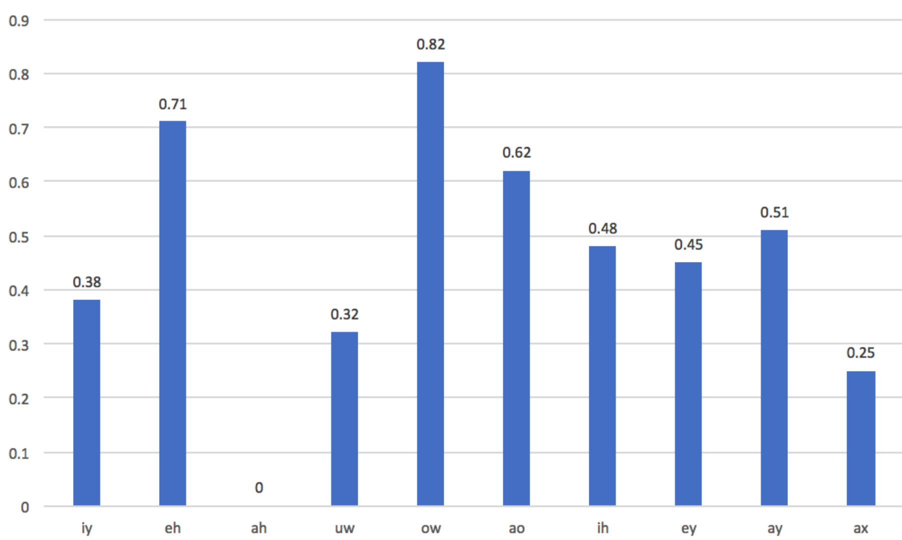

Figure 3

Extra Credit 1

In this section different Gaussian numbers. Figure. Shows the results. Note, in this problem, Gaussian is diagonal, which means more Gaussian number does not achieve better performance.

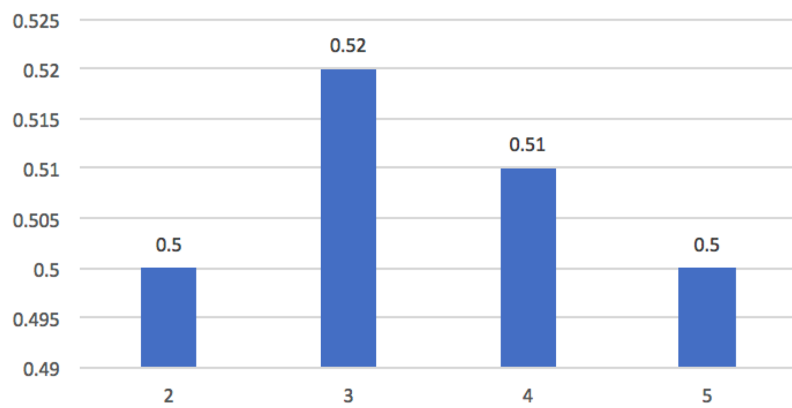

Figure 4

Extra Credit 2

In this section, we implement full covariance Gaussians for Part A and see whether the performance improve relative to Part A and Part B.

The average accuracy is 0.53 which is better than Part A and Part B.

Fig. 5 shows the classification accuracy across different subset: we can find that ah fails to get 0 classification result. Ax is also very low.

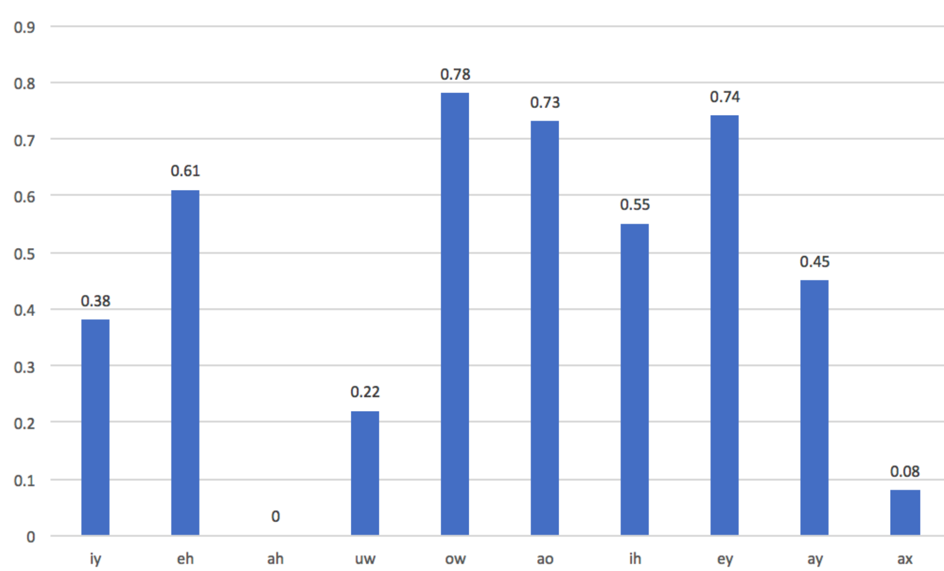

Figure 5

Extra Credit 3

In this section, we implement full covariance Gaussians for Part B (2 gaussian with convariance). Figure 6 shows the average accuracy under different iteration numbers with EM algorithm, which are better than Part A and Part B.

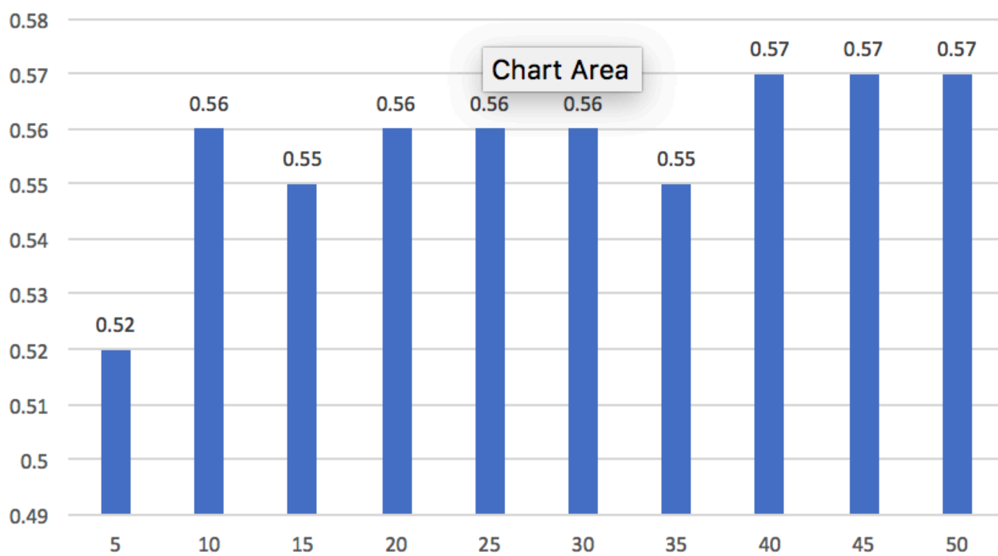
Chart Area
0.58
0.57
0.56
0.55
0.54
0.53
0.52
0.51
0.5
0.49
0.52
0.56
0.55
0.56
0.56
0.56
0.55
0.57
0.57
0.57
5
10
15
20
25
30
35
40
45
50

Figure 6