Pattern Recognition Assignment 1

Rendani Mbuvha, Huijie Wang
rendani@kth.se, huijiew@kth.se

September 2016

1. Theoretically we can verify this by calculating the stationary Distribution. This can be done by finding the vector $\mathbf{q}$ such that:

$$\mathbf{q}^T = \mathbf{q}^T\mathbf{A}$$

This implies that $\mathbf{q}$ is an eigenvector of transition matrix $\mathbf{A}$ corresponding to an eigenvalue 1.
In our given case : $\mathbf{q} = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} b_1(x) \\ b_2(x) \end{pmatrix}$.
where $b_1(x)$ is a scalar Gaussian density function with mean $\mu_1 = 0$ and standard deviation $\sigma_1 = 1$, and $b_2(x)$ is a gaussian distribution with mean $\mu_2 = 3$ and standard deviation $\sigma_2 = 2$.

$$\mathbf{q}^T\mathbf{A} = \begin{pmatrix} 0.75 & 0.25 \end{pmatrix} \begin{pmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{pmatrix} = \begin{pmatrix} 0.75 & 0.25 \end{pmatrix}$$

Therefore the given initial distribution is an eigenvector of the transition matrix with eigen value 1 meaning that $\mathbf{q}$ is stationary distribution of $\mathbf{A}$ and thus:

$$P(S_t = j) = \begin{cases} 0.75 & \text{if } j = 1, \\ 0.25 & j = 2. \end{cases}$$

2. By generating $T = 100000$ state integer numbers, we got 76.03% of occurrences of $S_t = 1$ and 23.97% of $S_t = 2$, which is approximately equal to $P(S_t = 1) = 0.75$ and $P(S_t = 1) = 0.25$, as stated in Question 1.

3. The theoretical calculations of $E[X_t]$ and $var[X_t]$ is shown as below:

$$\begin{aligned}
\mu_x &= E[X] = E_Z[E_X[X|Z]] \\
&= E_x[X|Z=1]p(Z=1) + E_x(X|Z=2)P(Z=2) \\
&= 0 \times 0.75 + 3 \times 0.25 \\
&= 0.75
\end{aligned}$$

$$\begin{aligned}
var[X] &= E_Z[var_X[X|Z]] + var_Z[E_X[X|Z]] \\
&= P(Z=1)var_X[X|Z=1] + P(Z=2)var_X[X|Z=2] \\
&+ P(Z=1)(E_X[X|Z=1]-\mu)^2 + P(Z=2)(E_X[X|Z=2]-\mu_X)^2 \\
&= (0.75 \times 1^2 + 0.25 \times 2^2) + [0.75 \times (0-0.75)^2 + 0.25 \times (3-0.75)^2] \\
&= 3.4375
\end{aligned}$$

As a result, by generating $T = 10000$ sequence in Matlab, we got $\mu = 0.7494$ and $var[X] = 3.3739$, which agrees approximately with the theoretical values.

4. The HMM emissions exhibit structured randomness -it can be seen the are jumps in the baseline as defined by the respective state conditional means of 0 and 3 respectively when the Markov Chain transitions between states. The variance/ volatility in the emissions is also seen to be higher when the MC in one state(2) than in the other(1).This pattern of structured randomness can be seen in figure 1.

5. The pattern of structured randomness is shown as figure 2. As can be seen, the pattern behaves similar to the former HMM in state transaction and variance; the only difference is that the emission mean of different states are equal. It surely makes it harder to estimate the state sequence S of the underlying Markov chain from the observed output, but it is still possible by estimating according to different variances.

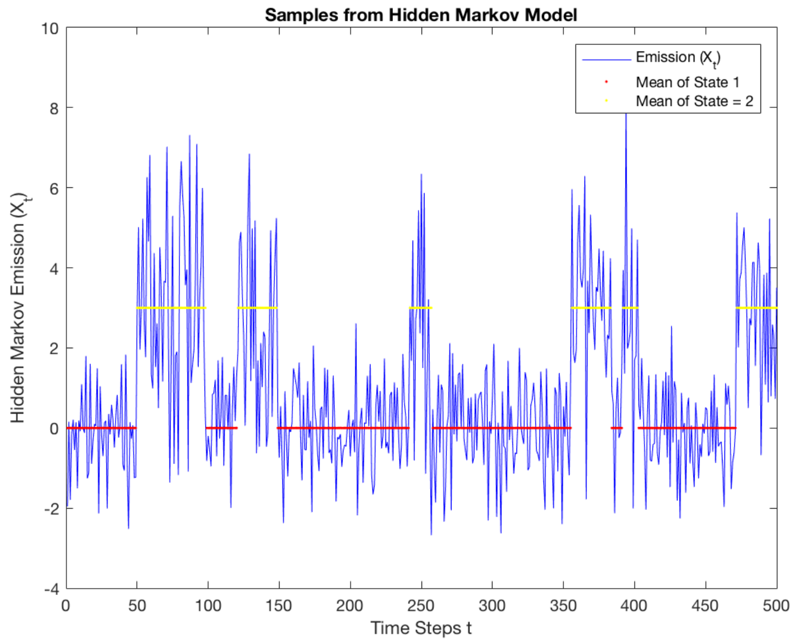

Figure 1: Samples from Hidden Markov Model with Different Conditional State Means

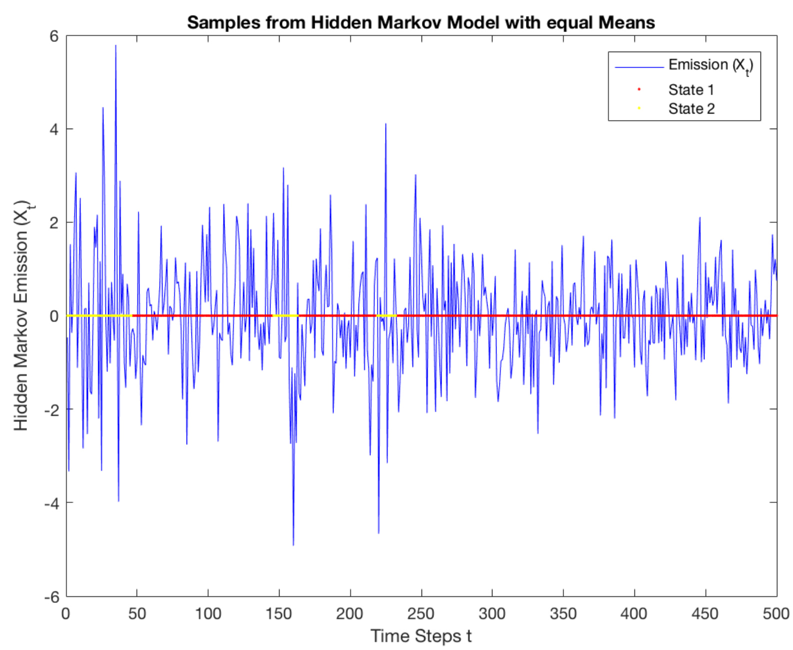

Figure 2: Samples from Hidden Markov Model with Equal Condtional State-Means

6. We defined a finite duration HMM with the following parameters Markov Chain Parameters: $\mathbf{q} = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$ and $\mathbf{A} = \begin{pmatrix} 0.99 & 0.01 & 0 \\ 0.02 & 0.97 & 0.01 \end{pmatrix}$.
The emission distributions where as given in question 1. Our test to evaluate if the code was working correctly was defined by checking:

   (a) If the output displayed a similar structured randomness as in question 4

   (b) If the MC stops when it reached the exit state

   (c) If when the MC is in the exit state there is no output.

   Figure 3 shows samples from a finite duration HMM which was set give 1000 time steps.

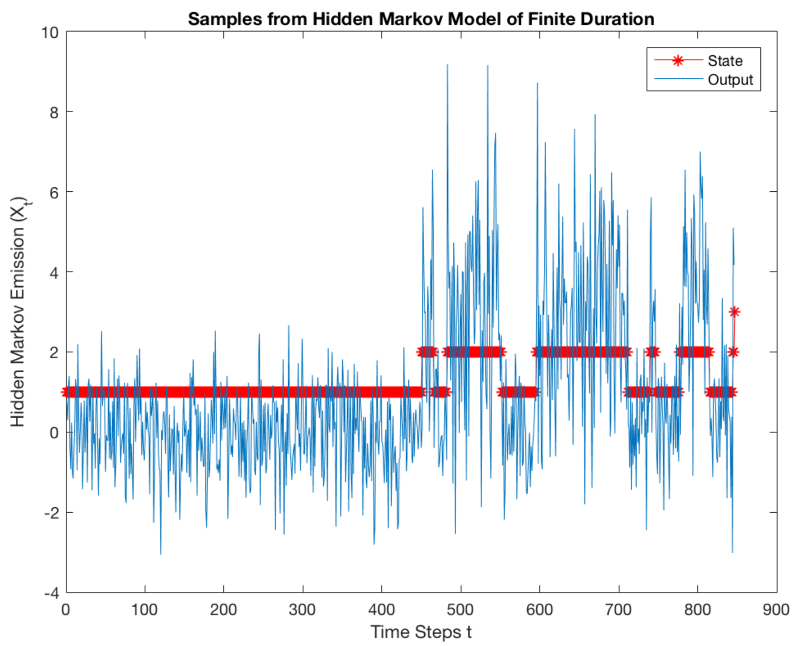

Figure 3: Samples from a Finite Duration Hidden Markov Model

From this out put we observe that:

   (a) The output still displays similar randomness to the HMM in question 4.

   (b) The MC stops after 847 steps when the exit state is reached (not the 1000) -thus indicating finite duration.It therefore implies that in this case it took 847 steps to reach the exit state.

(c) The output vector is only 846 in length - indicating that there was no emission when the chain was in the exit state $t = 847$

This confirms that this result meets our test for finite duration HMM.

7. We defined a new HMM where the outputs are two-dimensional Gaussian distributions with the following parameters:

$$\mathbf{q} = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} b_1(x) \\ b_2(x) \end{pmatrix}$$

, where $b_1(x)$ is a two-dimensional Gaussian distribution with $\mu_1 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ and $\Sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and $b_2(x)$ with $\mu_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\Sigma_2 = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$. The scatter plot of generated vector output is shown as 4. To make it clear, we plot the outputs with different underlying states using different colours. As can be seen, in the vector outputs of state 1, the scatter center is around $\begin{pmatrix} 4 \\ 5 \end{pmatrix}$, corresponding to μ_1. And as its covariance is diagonal, the output distribution is more circular for the two dimensions are independent. Similarly, the scatter center of the vector outputs of state2 is around μ_2, but the distribution is more elliptical as the two-dimensions are dependent.

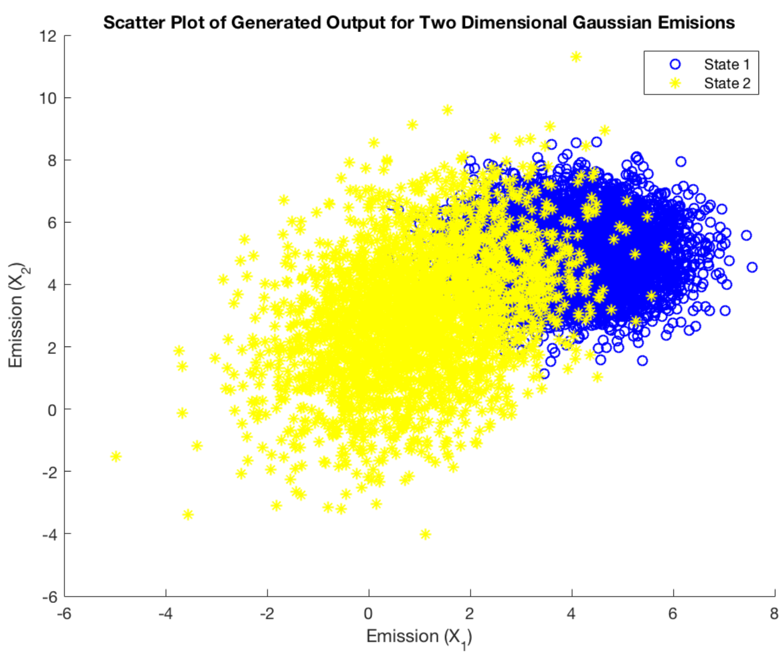

Figure 4: Generated Vector Output for Two Dimensional Gaussian Emissions