EQ2341 Pattern Recognition and Machine Learning
Project 1: HMM Signal Source

Letao Feng
letao@kth.se

Rao Xie
raox@kth.se

1 Questions

1.1

The infinite-duration HMM $\lambda = q, A, B$ is given as following

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

For $t = 1$, we can calculate $P(S_t = j)$, $j \in 1, 2$ as

$$P(S_1 = 1) = q(1) = 0.75$$

$$P(S_1 = 2) = q(2) = 0.25$$

For $t = 2$, we can calculate $P(S_t = j)$, $j \in 1, 2$ as

$$\begin{aligned} P(S_2 = 1) &= P(S_1 = 1)P(S_2 = 1|S_1 = 1) + P(S_1 = 2)P(S_2 = 1|S_1 = 2) \\ &= 0.75 * 0.99 + 0.25 * 0.03 = 0.75 \end{aligned}$$

$$\begin{aligned} P(S_2 = 2) &= P(S_1 = 1)P(S_2 = 2|S_1 = 1) + P(S_1 = 2)P(S_2 = 2|S_1 = 2) \\ &= 0.75 * 0.01 + 0.25 * 0.97 = 0.25 \end{aligned}$$

From above we can find that, the probabilities $P(S_t = 1)$ and $P(S_t = 2)$ remain the same for $t = 1, 2$. Similarly, we can calculate $P(S_t = j)$, $j \in 1, 2$ for any time t as

$$\begin{aligned} P(S_t = 1) &= P(S_{t-1} = 1)P(S_t = 1|S_{t-1} = 1) + P(S_{t-1} = 2)P(S_t = 1|S_{t-1} = 2) \\ &= 0.75 * 0.99 + 0.25 * 0.03 = 0.75 \end{aligned}$$

$$\begin{aligned} P(S_t = 2) &= P(S_{t-1} = 1)P(S_t = 2|S_{t-1} = 1) + P(S_{t-1} = 2)P(S_t = 2|S_{t-1} = 2) \\ &= 0.75 * 0.01 + 0.25 * 0.97 = 0.25 \end{aligned}$$

Therefore, we can verify that $P(S_t = j)$ is constant for all t.

1.2

Here we use our Markov chain rand function to generate a sequence of 10000 state integer numbers and calculate the relative frequency of occurrences of $S_t = 1$ and $S_t = 2$. The frequency of occurrences and their possibilities are shown in Table 1. From the table, we can see that the probability of $S_t = 1$ occurring in the generated state sequence is **0.7519**, which is approximately equal to $P(S_t = 1) = 0.75$. Similarly for the $S_t = 2$ condition, the probability is **0.2481**, which is close to $P(S_t = 2) = 0.25$. Therefore, we can conclude that the Markov chain rand function works well in this example.

Table 1: The Frequency of occurrences for different states

State	Frequency of Occurrences	Probabilities
$S_t = 1$	7519	0.7519
$S_t = 2$	2481	0.2481

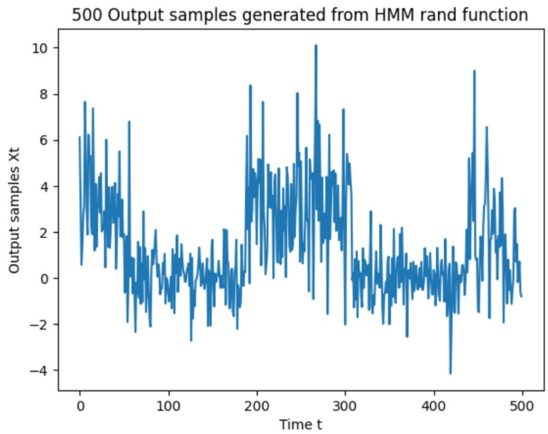

Figure 1: 500 Output Samples from A Infinite HMM Rand Function

1.3

The output probability distributions are given as $B = \begin{pmatrix} b_1(x) \\ b_2(x) \end{pmatrix}$, where $b_1(x)$ and $b_2(x)$ are the scalar Gaussian density functions.

$$b_1(x) \sim \mathcal{N}(\mu_1 = 0, \sigma_1^2 = 1^2); \qquad b_2(x) \sim \mathcal{N}(\mu_2 = 3, \sigma_2^2 = 2^2)$$

The output probability B depends on the state variable S and its theoretical mean $E[X_t]$ can be calculated as

$$\begin{aligned} E[X] &= E_S[E_X[X|S]] \\ &= P(S_t = 1)E_X[X|S_t = 1] + P(S_t = 2)E_X[X|S_t = 2] \\ &= 0.75 * 0 + 0.25 * 3 \\ &= 0.75 \end{aligned}$$

The theoretical variance $var[X_t]$ can be calculated as

$$\begin{aligned} var[X] &= E_S[var_X[X|S]] + var_S[E_X[X|S]] \\ &= P(S_t = 1)var_X[X|S_t = 1] + P(S_t = 2)var_X[X|S_t = 2] \\ &+ P(S_t = 1)var_S[E_X[X|S = 1]] + P(S_t = 2)var_S[E_X[X|S = 2]] \\ &= 0.75 * 1^2 + 0.25 * 2^2 + 0.75 * (0.75 - 0)^2 + 0.25 * (0.75 - 3)^2 \\ &= 3.4375 \end{aligned}$$

Now we use our HMM rand function to generate a sequence of 10000 output scalar random numbers x from the given HMM example. To verify our HMM rand method, we use the standard Numpy functions $np.mean()$ and $np.var()$ to calculate the mean and variance of the generated sequence x. The results are shown in Table 2 and it can be seen that the results obtained are in good agreement with those theoretically calculated.

Table 2: The mean and variance of the generated output sequence x

	Theoretical	Practical
Mean	0.75	0.7421
Variance	3.4375	3.4027

1.4

We use our HMM rand function to generate a series of 500 contiguous samples X_t from the above-defined HMM and plot them as a function of time t. As we can see from the results in Figure 1, we can easily identify which output sample comes from which state. This is because the given distribution $b_1(x)$ has a mean equal to 0 and $b_2(x)$ has a mean equal to 3. From the results, we can see that samples from state 1 are near the line $y = 0$ and samples from state 2 are near the line $y = 3$, which is consistent with the characteristics of the given output distribution. In addition, the standard deviation of $b_2(x)$ is larger than that of $b_1(x)$. It is also clear from the results that the samples with higher positions are more widely distributed and more variable than those with lower positions.

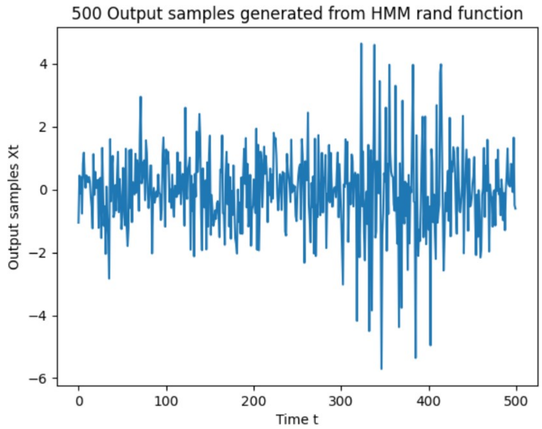

Figure 2: 500 Output Samples from A New Infinite HMM Rand Function

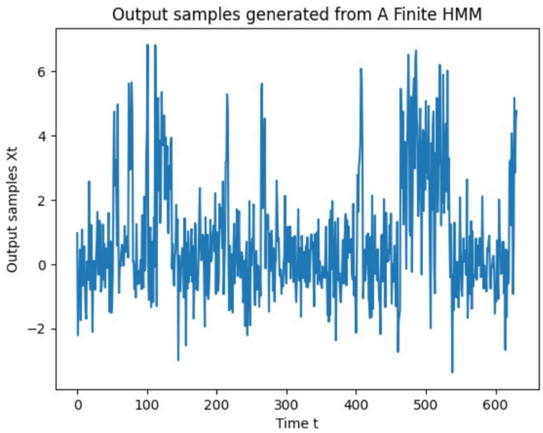

Figure 3: Output length less than 1000 samples

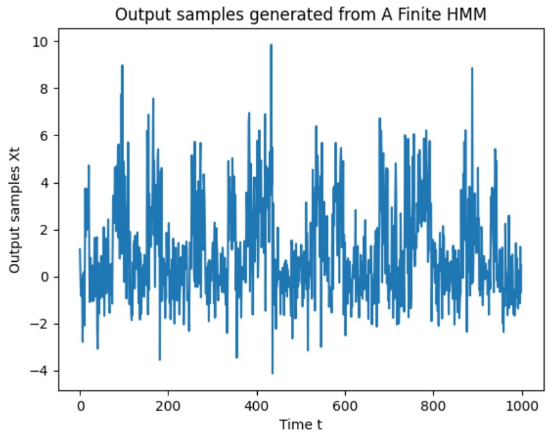

Figure 4: Output length equal to 1000 samples

1.5

A new HMM identical to the above one except for the mean of the second output distributions becomes $\mu_2 = 0$ is created.

$$b_1(x) \sim \mathcal{N}(\mu_1 = \mathbf{0}, \sigma_1^2 = 1^2); \qquad b_2(x) \sim \mathcal{N}(\mu_2 = \mathbf{0}, \sigma_2^2 = 2^2)$$

To compare the behaviour of these two HMM models, we also generate a series of 500 contiguous samples X_t from the new HMM and plot them as a function of time t, as shown in Figure 2. We can see that we cannot identify state transitions as easily in this graph as in Figure 1. Because the mean of both output distributions is 0 and all the samples are close to 0, we can only identify the state transitions from the variation amplitude of the samples. The standard deviation of $b_2(x)$ is larger than that of $b_1(x)$, which shows a large fluctuation in the results.

1.6

Here, we define a **finite-duration** HMM as follows,

$$q = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}; \quad A = \begin{pmatrix} 0.98 & 0.019 & 0.001 \\ 0.039 & 0.96 & 0.001 \end{pmatrix}; \quad B = \begin{pmatrix} b_1(x) \\ b_2(x) \end{pmatrix}$$

where

$$b_1(x) \sim \mathcal{N}(\mu_1 = 0, \sigma_1^2 = 1^2); \qquad b_2(x) \sim \mathcal{N}(\mu_2 = 3, \sigma_2^2 = 2^2)$$

To check whether our rand function is also applicable to this finite-duration HMM. We try to generate a series of 1000 samples from the HMM to check if the length of the output samples is reasonable. We ran it many times and found that the length of the output samples was always less than or equal to the set value of 1000 samples. This is consistent with the properties of the finite-duration HMM. The output length of less than 1000 samples is shown in Figure 3 and the length of equal to 1000 samples is shown in Figure 4.

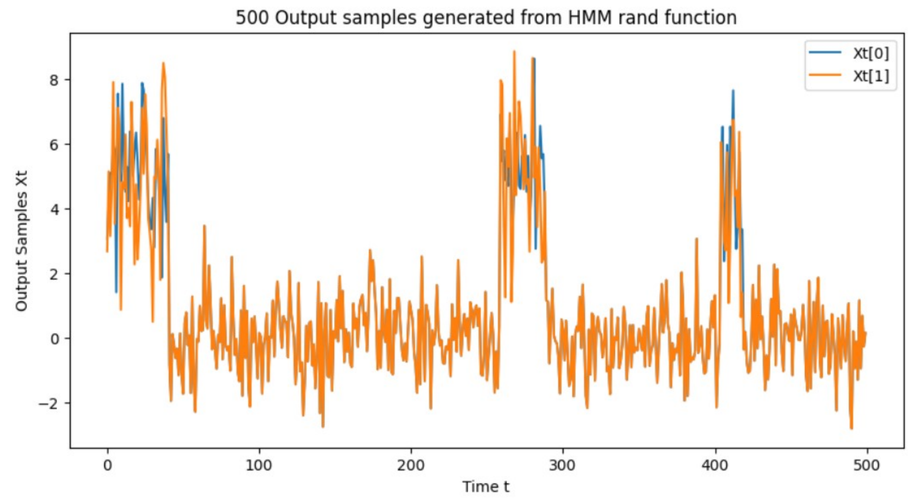

Figure 5: 500 Output Samples from A New Test HMM where Outputs are Gaussian Vector Distributions

1.7

To check if our rand function works when the state-conditional output distributions generate random vectors, we define a new text HMM as follows,

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

where

$$b_1(x) \sim \mathcal{N}(\mu_1 = 0, \mu_2 = 0, \Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}); \qquad b_2(x) \sim \mathcal{N}(\mu_1 = 5, \mu_2 = 5, \Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix})$$

We use the HMM rand function to generate a series of 500 output samples to see if it works well when the outputs are Gaussian vector distributions. The results are shown in Figure 5 and we can see that it successfully generated 500 output samples for each.