Assignment 1

Yann Debain
debain@kth.se

Harsha Holenarasipura Narasanna
harshahn@kth.se

April 16, 2018

1 Introduction

The task of this assignment is to code and verify MatLab methods to generate an output sequence of random real numbers $x = (x_1, ..., x_t, ..., x_T)$ from an HMM with scalar Gaussian output distributions. However, the code should be general enough to handle vector random variables as well.

The code is enclosed into the package **PattRecClasses-2**, you will find in this report the different tests implemented to verify that the code is correct.

2 Verify the Markov chain

The parameters of the Markov for the given system is :

$$q = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}; \quad A = \begin{pmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{pmatrix}$$

q represents the initial probabilities of the system $\Rightarrow q = \begin{pmatrix} P(S_0 = 1) = 0.75 \\ P(S_0 = 2) = 0.25 \end{pmatrix}$.
To compute $P(S_t = j), j \in \{1; 2\}$ we have to compute :

$$\begin{pmatrix} P(S_t = 1) \\ P(S_t = 2) \end{pmatrix} = \begin{pmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{pmatrix}^T * \begin{pmatrix} P(S_{t-1} = 1) \\ P(S_{t-1} = 2) \end{pmatrix}$$

By computing $\begin{pmatrix} P(S_1 = 1) \\ P(S_1 = 2) \end{pmatrix}$, we can see that $\begin{pmatrix} P(S_1 = 1) \\ P(S_1 = 2) \end{pmatrix} = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix} = q$.

We can deduce that the Markov chain is a stationary Markov chain and $\begin{pmatrix} P(S_t = 1) \\ P(S_t = 2) \end{pmatrix} = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$.

As the Markov chain is stationary we can verify the code by computing these probabilities by a frequency occurrence (for a large chain to be more precise).

For a Markov chain of length $T = 10000$, we have $\begin{cases} f_{S=1} = 0.7335 \simeq P(S_t = 1) \\ f_{S=2} = 0.2665 \simeq P(S_t = 2) \end{cases}$.

The results are in line with our expectations so we can assume that the code has been well implemented for a Markov chain.

3 Verify the HMM source

The parameters of the HMM source is :

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

where $b_i(x), i \in \{1; 2\}$ are Gaussian density functions.

3.1 Infinite Duration - Scalar Gaussian

3.1.1 Output parameters

To verify that the HMM source is well implemented for this problem we want to find to find $E[X_t]$ and $Var[X_t]$.

We have : $b_1(x) \sim \mathcal{N}(\mu_1 = 0; \sigma_1 = 1)$ and $b_2(x) \sim \mathcal{N}(\mu_2 = 3; \sigma_2 = 2)$

$$\begin{aligned} \mu_X = E[X] &= E_S[E_X[X|S]] \\ &= P(S=1) * E_X[X|S=1] + P(S=2) * E_X[X|S=2] \\ &= \sum_i P(S=i)\mu_i \end{aligned}$$

$$\begin{aligned} \sigma_X^2 = var[X] &= E_S[var_X[X|S]] + var_S[E_X[X|S]] \\ &= P(S=1) * var_X[X|S=1] + P(S=2) * var_X[X|S=2] \\ &+ P(S=1) * (E_X[X|S=1] - E_S[E_X[X|S]])^2 + P(S=2) * (E_X[X|S=2] - E_S[E_X[X|S]])^2 \\ &= P(S=1) * var_X[X|S=1] + P(S=2) * var_X[X|S=2] \\ &+ P(S=1) * (E_X[X|S=1] - \mu_X)^2 + P(S=2) * (E_X[X|S=2] - \mu_X)^2 \\ &= P(S=1) * \sigma_1^2 + P(S=2) * \sigma_2^2 + P(S=1) * (\mu_1 - \mu_X)^2 + P(S=2) * (\mu_2 - \mu_X)^2 \\ &= \sum_i P(S=i)[\sigma_i^2 + (\mu_i - \mu_X)^2] \end{aligned}$$

With our parameters, we have $\begin{cases} \mu_X = E[X] = 0.75 \\ \sigma_X^2 = var[X] = 3.4375 \end{cases}$

We can estimate these values via Matlab thanks to the function **mean(.)** and **var(.)**.
We have found $\begin{cases} \hat{\mu}_X = 0.7626 \\ \hat{\sigma}_X^2 = 3.4517 \end{cases}$ which are good estimates of the true values.

3.1.2 HMM behavior

To have an impression on how to behave the HMM source we have plotted the output result of the HMM source for a length $T = 500$

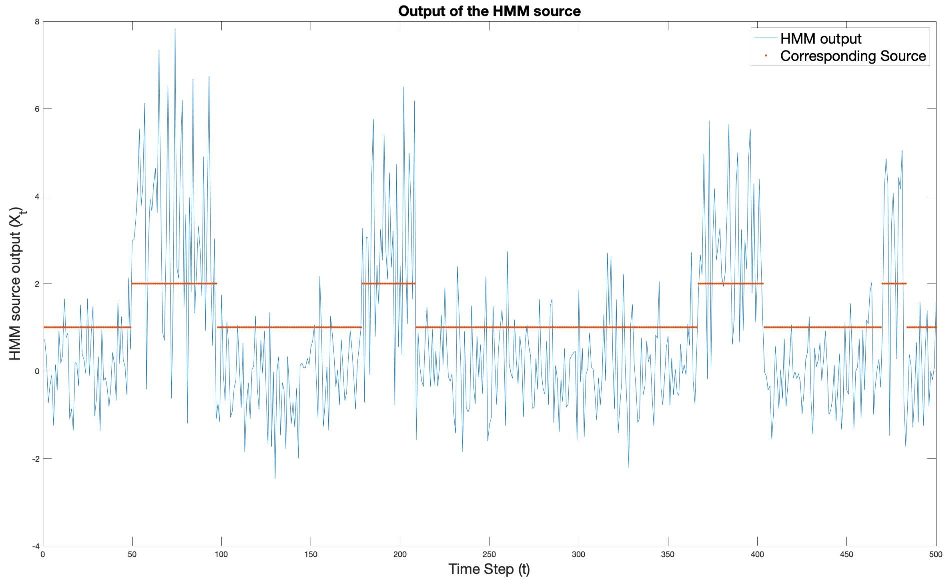

FIGURE 1 – Output of the HMM source for $\lambda = \{q, A, B\}$ and the corresponding sources

For this set of parameters, we can see that we jump from one state to the other. The jump between the states is depending on the transition matrix characterizing the Markov chain. Depending on the state the emission of the HMM source will have different statistics, we can see that when the emission is from S $= 1$, the emission is zero-mean and have low variance (compared to the emission from the other state). We can have the same remark for the emission from S $= 2$ (the mean is 3 and the emission sequence has a large variance compared to the emission from S $= 1$).

We can assume that the emission sequence follows the output distribution depending on its state. As $B = \begin{pmatrix} b_1(x) \sim \mathcal{N}(\mu_1 = 0; \sigma_1 = 1) \\ b_2(x) \sim \mathcal{N}(\mu_2 = 3; \sigma_2 = 2) \end{pmatrix}$, we can distinguish the states from the local statistics of the emission sequence.

To confirm this hypothesis, we change the statistics of the output distribution, we have now :
$b_1(x) \sim \mathcal{N}(\mu_1 = 0; \sigma_1 = 1)$ and $b_2(x) \sim \mathcal{N}(\mu_2 = 0; \sigma_2 = 2)$.

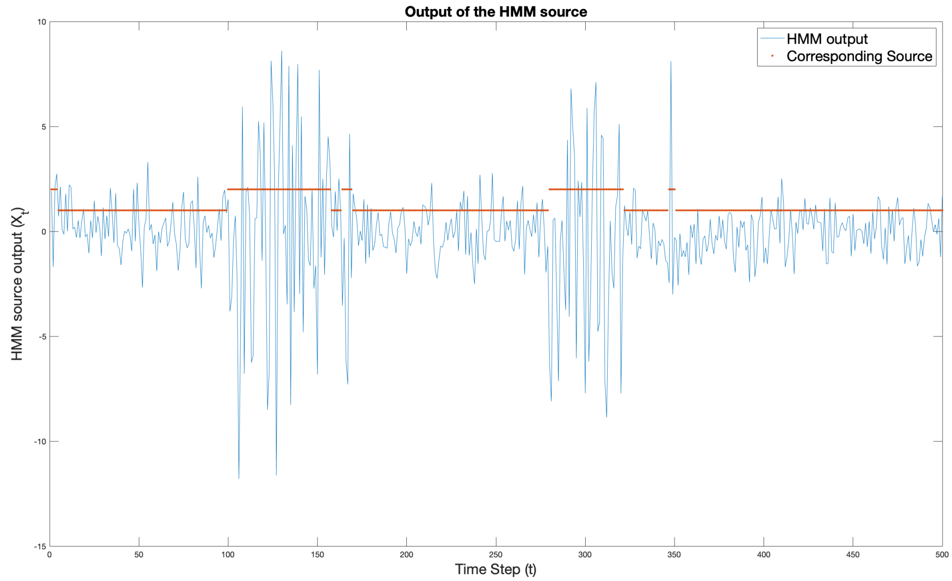

FIGURE 2 – Output of the HMM source for $\lambda = \{q, A, B\}$ and the corresponding sources

We can see that the distribution of the state stays the same as before. The change is in the local statistics of the emission, by putting $b_2(x) \sim \mathcal{N}(\mu_2 = 0; \sigma_2 = 5)$, the emissions from S = 2 have a zero-mean distribution.

To make a blind estimation of the state we have to look at the variance in a local region, we know that $\sigma_2^2 > \sigma_1^2$, so $\begin{cases} \text{in a high variance region} \Rightarrow S = 2 \\ \text{in a low variance region} \Rightarrow S = 1 \end{cases}$

3.2 Finite Duration - Scalar Gaussian

To test a finite HMM source, we have to create an **END** state.
The parameters for this HMM are :

$$q = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}; \quad A = \begin{pmatrix} 0.99 & 0.01 & 0 \\ 0.03 & 0.95 & 0.02 \end{pmatrix}; \quad B = \begin{pmatrix} b_1(x) \sim \mathcal{N}(\mu_1 = 0; \sigma_1 = 1) \\ b_2(x) \sim \mathcal{N}(\mu_2 = 3; \sigma_2 = 2) \end{pmatrix}$$

The transition probabilities to the **END** state for a given state is represented in the last column of A.

To verify the finite HMM implementation is correct we are checking if :

— The Markov chain stops at a time step t.
— There is no output after reaching the **END** state
— The last state is always S = 2 (for our parameters)

We can see in FIG.3, the output of the code for an HMM source of length $T = 500$.

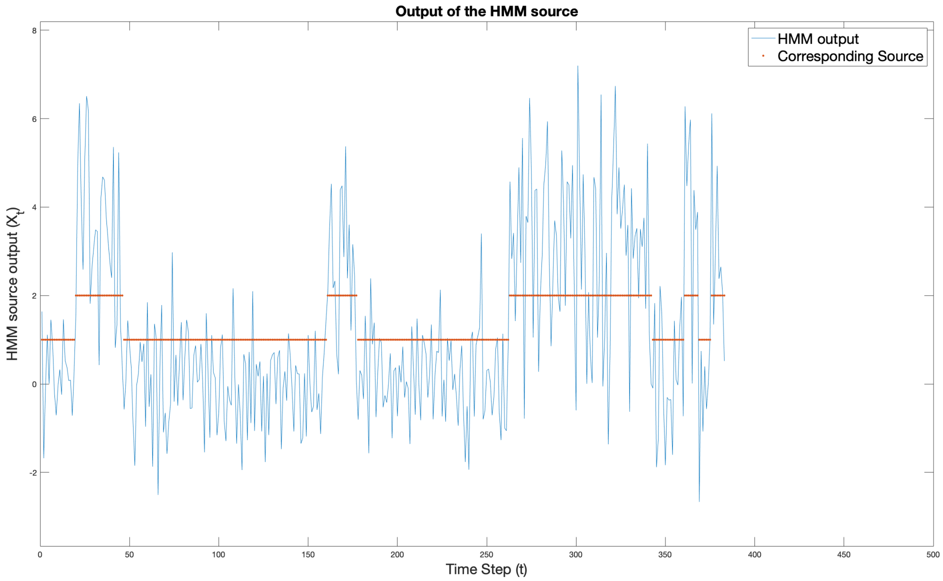

FIGURE 3 – Output of the HMM source for $\lambda = \{q, A, B\}$ and the corresponding sources

From this figure, we can see if our conditions are respected. We can see from FIG.3 :

— The sequence stopped before t = 500

— There is no sample after reaching the **END** state

— The last state is S = 2

As all our conditions are respected, we can assume that the HMM code has been well implemented for a finite sequence.

3.3 Gaussian Vector

3.3.1 Infinite duration

For an infinite duration the parameters for the HMM are :

$$sq = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}; \quad A = \begin{pmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{pmatrix}; \quad B = \begin{pmatrix} b_1(x) \sim \mathcal{N}(\mu_1 = \begin{pmatrix} -3 \\ -1 \end{pmatrix}; \Sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}) \\ b_2(x) \sim \mathcal{N}(\mu_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}; \Sigma_2 = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}) \end{pmatrix}$$

$$\begin{aligned} \mu_X = E[X] &= E_S[E_X[X|S]] \\ &= P(S=1) * E_X[X|S=1] + P(S=2) * E_X[X|S=2] \\ &= \sum_i P(S=i)\mu_i \end{aligned}$$

$$
\begin{aligned}
\Sigma_X = cov[X] &= E_S[cov_X[X|S]] + cov_S[E_X[X|S]] \\
&= P(S=1) * cov_X[X|S=1] + P(S=2) * cov_X[X|S=2] \\
&+ P(S=1) * (E_X[X|S=1] - E_S[E_X[X|S]]) * (E_X[X|S=1] - E_S[E_X[X|S]])^T \\
&+ P(S=2) * (E_X[X|S=2] - E_S[E_X[X|S]]) * (E_X[X|S=2] - E_S[E_X[X|S]])^T \\
&= P(S=1) * cov_X[X|S=1] + P(S=2) * cov_X[X|S=2] \\
&+ P(S=1) * (E_X[X|S=1] - \mu_X) * (E_X[X|S=1] - \mu_X)^T \\
&+ P(S=2) * (E_X[X|S=2] - \mu_X) * (E_X[X|S=2] - \mu_X)^T \\
&= P(S=1) * \Sigma_1 + P(S=2) * \Sigma_2 + P(S=1) * (\mu_1 - \mu_X) * (\mu_1 - \mu_X)^T + P(S=2) * (\mu_2 - \mu_X) * (\mu_2 - \mu_X)^T \\
&= \sum_i P(S=i)[\Sigma_i + (\mu_i - \mu_X) * (\mu_i - \mu_X)^T]
\end{aligned}
$$

With our parameters, we have $\begin{cases} \mu_X = E[X] = \begin{pmatrix} -1.5 \\ -0.5 \end{pmatrix} \\ \Sigma_X = cov[X] = \begin{pmatrix} 8 & 2.5 \\ 2.5 & 2.5 \end{pmatrix} \end{cases}$

Like for the Infinite Duration with Scalar Gaussian functions, we can verify that the sequence X_t gives the right values. We can estimate these values via Matlab thanks to the function **mean(.)** and **var(.)**.

We have found $\begin{cases} \hat{\mu}_X = \begin{pmatrix} -1.4976 \\ -0.4746 \end{pmatrix} \\ \hat{\Sigma}_X = \begin{pmatrix} 8.0249 & 2.5467 \\ 2.5467 & 2.5367 \end{pmatrix} \end{cases}$ which are good estimates of the true values.

We can see the HMM output for a length $T = 10000$.

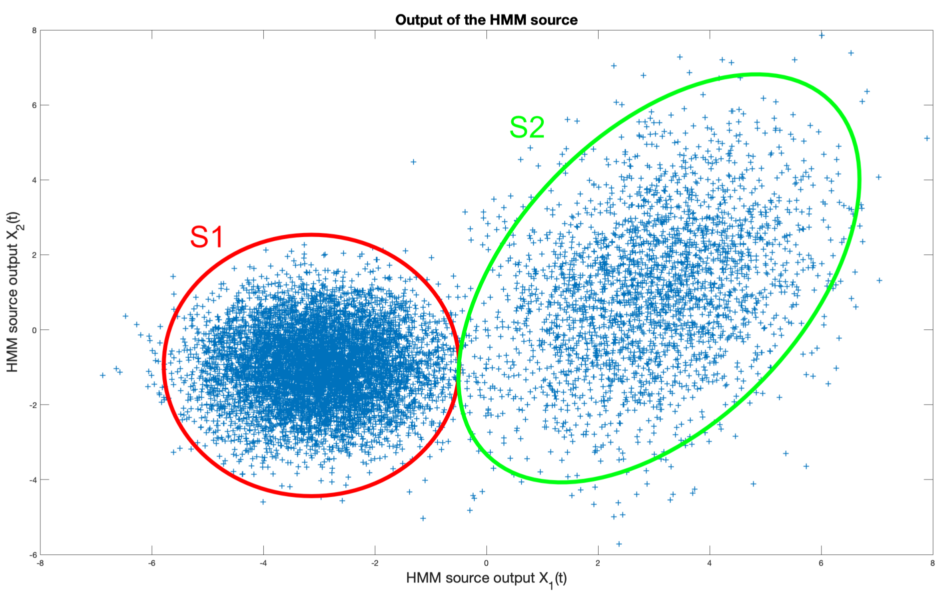

FIGURE 4 – Output of the infinite length HMM source for $\lambda = \{q, A, B\}$ and the corresponding sources

We can also verify from FIG.4 that the output sequence follows the statistics of the output distribution :

— The emission points from S = 1 seems to have a mass distribution around $\begin{pmatrix}-3\\-1\end{pmatrix}$

— The emission points from S = 2 seems to have a mass distribution around $\begin{pmatrix}3\\1\end{pmatrix}$

— the variance of the points from S = 1 is smaller than the variance of the points from S = 2.

3.3.2 Finite duration

In the case of finite length, HMM source we can verify that all the points for a Scalar Gaussian are respected.

$$A = \begin{pmatrix}0.99 & 0.01 & 0\\ 0.03 & 0.95 & 0.02\end{pmatrix}$$

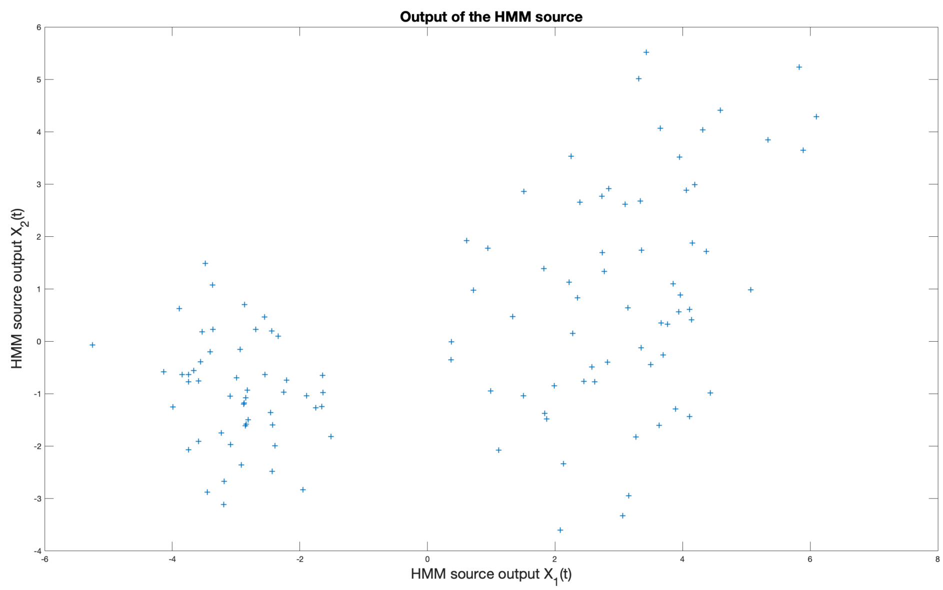

FIGURE 5 – Output of the finite length HMM source for $\lambda = \{q, A, B\}$ and the corresponding sources

We can see from FIG.5 that all the points from (3.2) are respected. (We have verified that the last point is from S = 2 with the code).

From these observations, we can conclude that the HMM code has been well implemented.

4 Conclusion

As a conclusion, we have seen how to implement functions to generate Markov chain sequences and HMM sequences.

Through observations and computations, we have verified if the codes have been well implemented. We have verified that all cases of HMM can be used (infinite/finite duration, vector/scalar output distributions).