EQ2340 Pattern Recognition Project

ÖZER OZKARHAMAN, MATIAS MÜLLER

Stockholm 2016

Assignment One

Özer Ozkarhaman, Matias Müller

September 24, 2016

1 Verifying the `MarkovChain` and `HMM` Sources

1.1 $P(S_t = j)$

Using both the sum and product rules:

$$\begin{aligned}P(S_t = j) &= \sum_{i=\{1,2\}} P(S_t = j \mid S_{t-1} = i)P(S_{t-1} = i)\\ &= P(S_t = j \mid S_{t-1} = 1)P(S_{t-1} = 1) + P(S_t = j \mid S_{t-1} = 2)P(S_{t-1} = 2).\end{aligned}$$

Trivial case of initial probabilities are given as the q vector. S_0 for $t = 0$

$$P(S_0 = 1) = q_1 = 0.75$$
$$P(S_0 = 2) = q_2 = 0.25.$$

For $j = 1$ and $t = 1$

$$\begin{aligned}P(S_1 = 1) &= P(S_1 = 1 \mid S_0 = 1)P(S_0 = 1) + P(S_1 = 1 \mid S_0 = 2)P(S_0 = 2)\\ &= a_{11}q_1 + a_{21}q_2\\ &= 0.99 \times 0.75 + 0.03 \times 0.25\\ &= 0.75 = q_1. \end{aligned} \tag{1}$$

For $j = 2$ and $t = 1$,

$$\begin{aligned}P(S_1 = 2) &= P(S_1 = 2 \mid S_0 = 1)P(S_0 = 1) + P(S_1 = 2 \mid S_0 = 2)P(S_0 = 2)\\ &= a_{12}q_1 + a_{22}q_2\\ &= 0.01 \times 0.75 + 0.97 \times 0.25\\ &= 0.25 = q_2. \end{aligned} \tag{2}$$

Since the unconditional probabilities for both states at $t = 1$ turned out to be the same as the initial probabilities, the following steps will have the same results due to the Markovian assumption. Thus,

$$P(S_t = j) = q_j.$$

The practical application of the same Markov chain with $T = 10000$ generated 7525 1's and 2475 2's which correspond to 0.7525 and 0.2475 of frequency, proving that the unconditional probabilities calculated here are correct.

1.2 Mean and variance

Mean In this setup, the random variable is X_t such that,

$$X_t = \mathcal{N}(0,1)P(S_t = 1) + \mathcal{N}(3,4)P(S_t = 2).$$

I think is N(3,2)s

Following this, the mean would be

$$\begin{aligned}
E[X_t] &= \sum_t^{\infty} \mathcal{N}(0,1)P(S_t = 1) + \mathcal{N}(3,4)P(S_t = 2) \\
&= \sum_t^{\infty} \mathcal{N}(0,1) \times 0.75 + \mathcal{N}(3,2) \times 0.25 \\
&= E[\mathcal{N}(0,1)] \times 0.75 + E[\mathcal{N}(3,2)]\ 0.25 \\
&= 0 \times 0.75 + 3 \times 0.25 \\
&= 0.75. \qquad (3)
\end{aligned}$$

A Markov chain with $T = 10000$ have produced means between 0.6 and 0.85 on our tests, which are sufficiently close to the theoretical mean.

Variance Since our random variable X_t is a mixture of two uncorrelated normal distributions, we can use the formula

$$\mathrm{Var}(P_A\mathcal{N}_A + P_B\mathcal{N}_B) = P_A\sigma_A^2 + P_B\sigma_B^2 + [P_A\mu_A^2 + P_B\mu_B^2 - (P_A\mu_A + P_B\mu_B)^2],$$

where $\mathcal{N}_A$ and $\mathcal{N}_B$ are two Normal distributions with means μ_A and μ_B and variances σ_A^2 and σ_B^2. Applying this formula to our problem, we have that

$$\begin{aligned}
\mathrm{Var}(X_t) &= 0.75 \times 1 + 0.25 \times 4 + [0.75 \times 0 + 0.25 \times 9 - (0.75 \times 0 + 0.25 \times 3)^2] \\
&= 3.4375.
\end{aligned}$$

Our implementation resulted in variances between 2.8 and 3.9, which agree with the theoretical result.

1.3 Signal Generation

Infinite HMM with different output distributions As it can be seen on Figure 1 the HMM with different output means and variances produces a noisy two level result that changes its level depending on $P(S_{t+1} \mid S_t)$. Each state can be easily distinguished from the graph. The exact point of time where the state changes is harder to see.

Infinite HMM with similar output distributions As seen on Figure 2, it is still possible to distinguish the state sequence from the output, albeit harder and less certain. The difference in the variances is what allows this.

Finite HMM with different output distributions We create an extended 2-states Markov Chain by adding an extra exit state. The MC itself is defined as

$$q = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \quad A = \begin{pmatrix} 0.69 & 0.3 & 0.01 \\ 0.1 & 0.89 & 0.01 \end{pmatrix}, \quad B = \begin{pmatrix} \mathcal{N}(0,1) \\ \mathcal{N}(10,1) \end{pmatrix}.$$

Figure 3 shows the result: at $T = 274$ the Markov Chain reaches the exit state ($S_{274} = 3$) and no more samples are then generated.

Why do you think this is a reasonable length to end the sequence?How long did y

Infinite HMM with vector output distributions In the case of a vector as state outputs, it is harder to visually distinguish the states, but it is still possible. As can be seen on Figure 4 the variance of the data allows us to distinguish the states over time. Colours are added for visualizations sake. The following HMM definition was used.

$$q = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}, \quad A = \begin{pmatrix} 0.99 & 0.01 \\ 0.03 & 0.97 \end{pmatrix}, \quad B = \begin{pmatrix} \mathcal{N}(\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}) \\ \mathcal{N}(\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma = \begin{bmatrix} 10 & 3 \\ 3 & 7 \end{bmatrix}) \end{pmatrix}.$$

Can you explain the results a bit more? Or maybe provide an anlytical way to ve

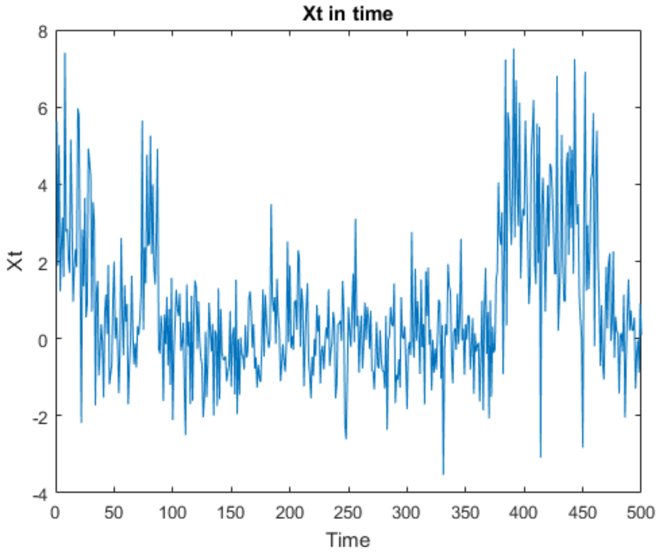

Figure 1: First HMM with different means.

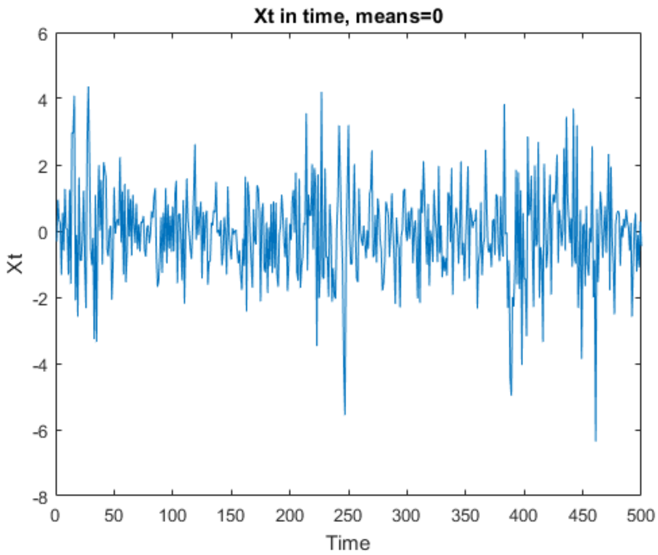

Figure 2: Second HMM with 0 means, different variances.

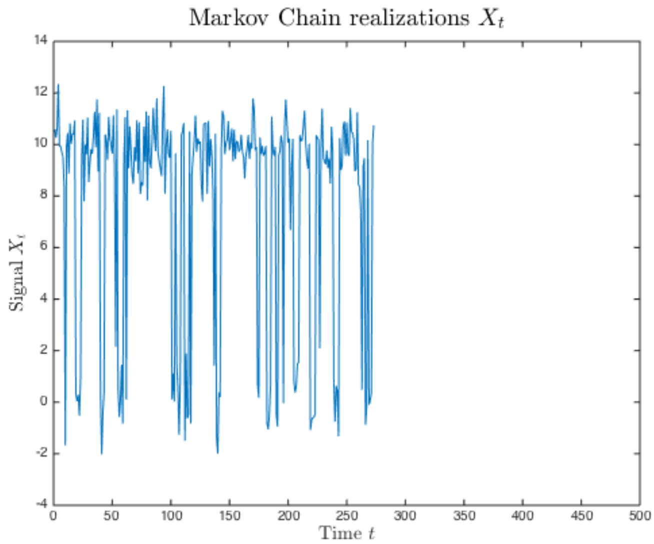

Figure 3: Finite 2-states scalar Markov Chain.

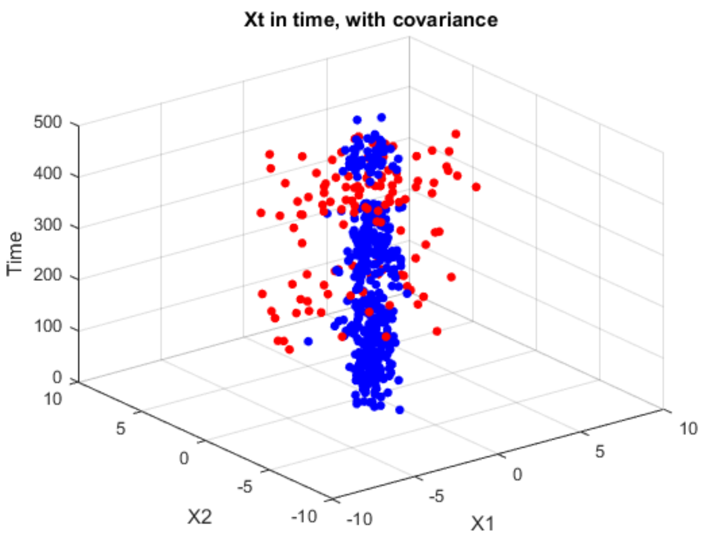

Figure 4: HMM with a vector output distribution with correlation between output elements X_n