Hidden Markov Models: Exercises

Xavier Didelot

Exercise 1

Let X be a random variable distributed as a δ_1, δ_2 mixture of two distributions with expectations μ_1, μ_2, and variances σ_1^2 and σ_2^2, respectively, where $\delta_1 + \delta_2 = 1$.

1. Show that $\mathrm{E}(x) = \delta_1\mu_1 + \delta_2\mu_2$
2. Show that $\mathrm{Var}(x) = \delta_1\sigma_1^2 + \delta_2\sigma_2^2 + \delta_1\delta_2(\mu_1 - \mu_2)^2$
3. Show that a mixture of two Poisson distributions is overdispersed, that is $\mathrm{Var}(x) > \mathrm{E}(x)$
4. Generalize to a mixture of $m \geq 2$ Poisson distributions

Exercise 2

Consider a stationary two-state Markov chain with transition probability matrix given by:

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$$

1. Show that the stationary distribution is:

$$(\delta_1, \delta_2) = \frac{1}{\gamma_{12} + \gamma_{21}}(\gamma_{21}, \gamma_{12})$$

2. Consider the case

$$\mathbf{\Gamma} = \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix}$$

   The following two sequences of observations are assumed to be generated by this Markov chain:

   Sequence A: 1 1 1 2 2 1
   Sequence B: 2 1 1 2 1 1

   Compute the probability of each of the sequences. Note that each sequence contains the same number of ones and twos. Why are these sequences not equally probable?

Exercise 3

Consider a stationary two-state Poisson-HMM with parameters

$$\mathbf{\Gamma} = \begin{pmatrix} 0.1 & 0.9 \\ 0.4 & 0.6 \end{pmatrix} \text{ and } \boldsymbol{\lambda} = (1, 3)$$

We observe the data $X = (X_1, X_2, X_3) = (0, 2, 1)$.

1. Compute the probability of the observation by considering all possible sequences of state of the Markov chain that could have occurred.

2. Compute the probability of the observation using the relevant matrix form of the HMM likelihood equation.

Exercise 4

Consider a stationary two-state Poisson-HMM with parameters

$$\boldsymbol{\Gamma} = \begin{pmatrix} 0.1 & 0.9 \\ 0.4 & 0.6 \end{pmatrix} \text{ and } \boldsymbol{\lambda} = (1, 3)$$

We observe the data $X = (X_1, X_3) = (0, 1)$. Note that X_2 is missing data. 1. Compute the probability of the observation by considering all possible sequences of state of the Markov chain that could have occurred. 2. Compute the probability of the observation using the relevant matrix form of the HMM likelihood equation.

Exercise 5

Show that the general expression for the likelihood of a stationary HMM simplifies into the expression for the likelihood of an independent mixture model if the transition matrix $\boldsymbol{\Gamma}$ is such that the current state does not depend on the previous one.

Exercise 6

Simplify the general expression for the likelihood of a stationary HMM in the case where for every state i and every observation x we have $p_i(x) = p(x)$.

Exercise 7

Consider a stationary m-state Poisson-HMM $\{X_t\}$ with transition probability matrix $\boldsymbol{\Gamma}$ and state-dependent means $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_m)$. Let $\boldsymbol{\delta}$ be the stationary distribution of the Markov chain. Let $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{\lambda})$.

1. Show that $\text{E}(X_t) = \boldsymbol{\delta}\boldsymbol{\lambda}'$
2. Show that $\text{E}(X_t^2) = \boldsymbol{\delta}\boldsymbol{\Lambda}\boldsymbol{\lambda}' + \boldsymbol{\delta}\boldsymbol{\lambda}'$
3. Show that $\text{Var}(X_t) = \boldsymbol{\delta}\boldsymbol{\Lambda}\boldsymbol{\lambda}' + \boldsymbol{\delta}\boldsymbol{\lambda}' - (\boldsymbol{\delta}\boldsymbol{\lambda}')^2$
4. Show that X_t is overdispersed, ie $\text{Var}(X_t) > \text{E}(X_t)$
5. Show that $\text{E}(X_t X_{t+k}) = \boldsymbol{\delta}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^k\boldsymbol{\lambda}'$
6. Show that $\text{Corr}(X_t, X_{t+k}) = \frac{\boldsymbol{\delta}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^k\boldsymbol{\lambda}' - (\boldsymbol{\delta}\boldsymbol{\lambda}')^2}{\boldsymbol{\delta}\boldsymbol{\Lambda}\boldsymbol{\lambda}' + \boldsymbol{\delta}\boldsymbol{\lambda}' - (\boldsymbol{\delta}\boldsymbol{\lambda}')^2}$

Exercise 8

Consider a stationary Bernouilli-HMM with transition probability matrix $\boldsymbol{\Gamma} = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}$, probability of emission of a one in state 1 equal to $p_1 = 1/4$, probability of emission of a one in state 2 equal to $p_2 = 3/4$. We observe the data $x_1 = 1$ and $x_2 = 1$.

1. Write down the emission probability matrix $\boldsymbol{P} = \boldsymbol{P}(x = 1)$
2. Find the stationary distribution $\boldsymbol{\delta}$
3. Calculate the forward vector $\boldsymbol{\alpha}_1$, then $\boldsymbol{\alpha}_2$ and deduce the likelihood L_T using the foward algorithm
4. Calculate the backward vector $\boldsymbol{\beta}_1'$ and deduce the likelihood L_T using the backward algorithm
5. Calculate L_T using $\boldsymbol{\alpha}_1$ and $\boldsymbol{\beta}_1$

Exercise 9

1. Show that state prediction can be performed for a given HMM model using the formula:

$$p(C_{T+h} = i|\boldsymbol{X}^{(T)} = \boldsymbol{x}^{(T)}) = \boldsymbol{\alpha}_T \boldsymbol{\Gamma}^h \boldsymbol{e}_i'/L_T$$

   where $\boldsymbol{e}_i$ is a vector of zeros except for a one at the i-th position.

2. Deduce that forecasting can be performed using the formula:

$$p(X_{T+h} = x|\boldsymbol{X}^{(T)} = \boldsymbol{x}^{(T)}) = \boldsymbol{\alpha}_T \boldsymbol{\Gamma}^h \boldsymbol{P}(x)\mathbf{1}'/L_T$$

3. Prove the forecasting formula again, using a ratio of likelihoods.

Exercise 10

Let $\{X_t\}$ be a second-order HMM, based on a stationary second-order Markov chain $\{C_t\}$ with m states. We use the following definitions:

$$u(i,j) = p(C_{t-1} = i, C_t = j)$$

$$p_i(x_t) = p(X_t = x_t|C_t = i)$$

$$\gamma(i,j,k) = p(C_t = k|C_{t-1} = j, C_{t-2} = i)$$

$$v_t(i,j|\boldsymbol{x}^{(t)}) = p(\boldsymbol{X}^{(t)} = \boldsymbol{x}^{(t)}, C_{t-1} = i, C_t = j)$$

1. Write down $v_2(j,k|\boldsymbol{x}^{(2)})$ in terms of the definitions above.

2. Show that for integers $t \geq 3$ we have:

$$v_t(j,k|\boldsymbol{x}^{(t)}) = \left(\sum_{i=1}^{m} v_{l-1}(i,j|\boldsymbol{x}^{(t-1)})\gamma(i,j,k)\right) p_k(x_t)$$

3. Show how the recursion above can be used to calculate the likelihood.

4. What is the computational effort required to calculate the likelihood?