Extra exercises MSA101/MVE187, autumn 2017

Petter Mostad

October 12, 2017

1. Consider a hidden Markov model as depicted in Figure 1. We assume the X_i are discrete variables with three different values: 1, 2, and 3. We assume
$$\Pr(X_0 = 1) = \Pr(X_0 = 2) = \Pr(X_0 = 3) = 1/3$$
and for $i = 1, \dots, T$,
$$\Pr(X_i = j \mid X_{i-1}) = \begin{cases} 0.8 & \text{when } j = X_{i-1} \\ 0.1 & \text{when } j \neq X_{i-1} \end{cases}$$
We assume the Y_i are non-negative whole numbers, with
$$Y_i \mid X_i \sim \text{Poisson}(\lambda_{X_i}),$$
where λ_1, λ_2, and λ_3 are parameters. We assume the values of $Y_0, \dots, Y_T$ are observed as $y_0, \dots, y_T$.

   Define, for $i = 0, \dots, T$ and $j = 1, 2, 3$,
$$b_{ij} = \Pr(x_0, \dots, x_i, y_0, \dots, y_i)$$
where $x_0, \dots, x_i$ is the sequence of values for $X_0, \dots, X_i$ maximizing the probability, where we demand that $x_i = j$. Let c_{ij} be the value of x_{i-1} in this sequence.

   (a) Assuming all b_{ij} and c_{ij} have been computed for $i \leq k$. Find formulas for $b_{k+1,j}$ and $c_{k+1,j}$.

   (b) Assuming all the b_{ij} and c_{ij} have been computed, describe an algorithm that produces the sequence $x_0, \dots, x_T$ maximizing the probability of the data $y_0, \dots, y_T$ in this model. Give the name of the algorithm.

2. Consider a hidden Markov model like the one depicted in Figure 1. We assume the variables are Normally distributed, with
$$X_0 \sim \text{Normal}(0, \frac{1}{\tau_0}),$$

for $i = 1, \ldots, T$

$$X_i \mid X_{i-1} \sim \text{Normal}(X_{i-1}, \frac{1}{\tau_x}),$$

and for $i = 0, \ldots, T$

$$Y_i \mid X_i \sim \text{Normal}(X_i, \frac{1}{\tau_y}).$$

We assume the precision parameters τ_0, τ_x, τ_y are given, as well as values $y_0, \ldots, y_T$ for the variables $Y_0, \ldots, Y_T$.

(a) Assume we have shown, for some $0 \leq i < T$, that

$$X_i \mid y_0, \ldots, y_i \sim \text{Normal}(a_i, \frac{1}{t_i})$$

for some parameters a_i and t_i. Prove that $X_{i+1} \mid y_0 \ldots, y_{i+1}$ is Normally distributed, and find formulas for the parameters of this distribution in terms of $a_i, t_i, \tau_0, \tau_x, \tau_y, y_0, \ldots, y_i$.

(b) Assume you have shown, for some $0 < i \leq T$, that the likelihood

$$\pi(y_{i+1}, \ldots, y_T \mid x_i)$$

as a function of x_i is proportional to a normal distribution with expectation b_i and precision s_i. Show that the likelihood

$$\pi(y_i, y_{i+1}, \ldots, y_T \mid x_i)$$

as a function of x_i is also is proportional to a normal distribution, and find the parameters of this distribution.

(c) Assume you have shown that the likelihood

$$\pi(y_i, \ldots, y_T \mid x_i)$$

as a function of x_i is proportional to a normal likelihood with parameters c_i och prercision w_i. Show that the likelihood

$$\pi(y_i \ldots, y_T \mid x_{i-1})$$

is also proportional to a normal likelihod, and find its parameters.

(d) Given the above definitions and computations, show that, for all $i = 0, \ldots, T$,

$$X_i \mid y_0 \ldots, y_T$$

is Normally distributed, and find the parameters of this distribution in term of previously computed variables.

(e) The above computations can also be adapted into an algorithm simulating $X_0, \ldots, X_T$ in the distribution

$$\pi(X_0, \ldots, X_T \mid y_0, \ldots, y_T)$$

Explain how.

3. Reconsider exercise 1 above, but now assume that the parameters $\theta = (\lambda_1, \lambda_2,$ and $\lambda_3)$ are unknown. Find a maximum likelihood estimate for these parameters as follows:

   (a) Assuming you view $X_0, \ldots, X_T$ as augmented data, find a formula for the full data likelihood, i.e.,

   $$\pi(X_0, \ldots, X_T, Y_0, \ldots, Y_T \mid \lambda_1, \lambda_2, \lambda_3)$$

   Use a notation where you use exponents with $I(X_i = j)$, where I is the indicator function.

   (b) Fixing $\theta' = (\lambda_1', \lambda_2', \lambda_3')$, describe how one may compute the marginal distribution of each X_i.

   (c) Compute the function $Q(\theta, \mid \theta')$ of the EM-algorithm.

   (d) Describe how you may maximize $Q(\theta \mid \theta')$ as a function of $\theta = (\lambda_1, \lambda_2,$ and $\lambda_3)$.

   (e) Describe how you may put together your answers above to obtain an algorithm answering the original question.

4. Reconsider exercise 2 above, now assuming the parameters $\theta = (\tau_0, \tau_x, \tau_y)$ are unknown. Find a maximum likelihood estimate for these parameters as follows:

   (a) Assuming you view $X_0, \ldots, X_T$ as augmented data, find a formula for the full data likelihood

   $$\pi(X_0, \ldots, X_T, Y_0, \ldots, Y_T \mid \tau_1, \tau_2, \tau_3)$$

   (b) Fixing $\theta' = (\tau_1', \tau_2', \tau_3')$, describe how one may compute the marginal distribution of each X_i.

   (c) Compute the function $Q(\theta \mid \theta')$ of the EM-algorithm.

   (d) Describe how you may maximize $Q(\theta \mid \theta')$ as a function of $\theta = (\tau_1, \tau_2,$ and $\tau_3)$.

   (e) Describe how you may put together your answers above to obtain an algorithm answering the original question.

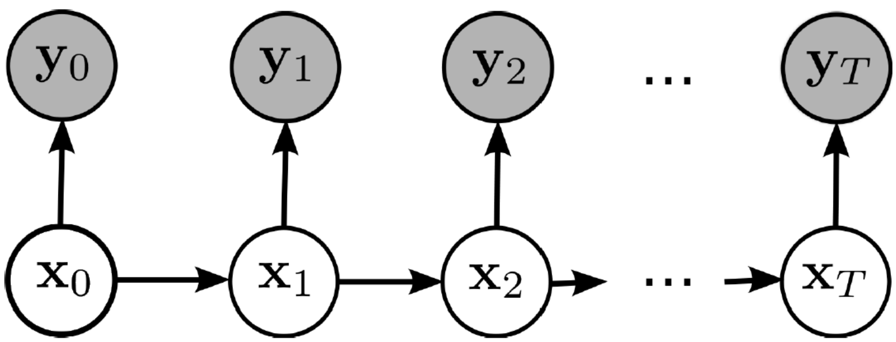

Figure 1: A hidden Markov model