

Problem Set 13

Spring 2019

Issued: Monday May 6, 2019

Due: N/A

1. Higher-Order Markov Chains

Let k be a fixed positive integer. A stochastic process $(X_n)_{n \in \mathbb{N}}$ taking values in a discrete state space \mathcal{X} is called a **k th order (time homogeneous) Markov chain** if for all $n \in \mathbb{N}$ and all feasible sequences $x_0, x_1, \dots, x_{n+k} \in \mathcal{X}$,

$$\begin{aligned}\mathbb{P}(X_{n+k} = x_{n+k} \mid X_0 = x_0, X_1 = x_1, \dots, X_{n+k-1} = x_{n+k-1}) \\ &= \mathbb{P}(X_{n+k} = x_{n+k} \mid X_n = x_n, \dots, X_{n+k-1} = x_{n+k-1}) \\ &= P_k(x_{n+k} \mid x_n, \dots, x_{n+k-1}).\end{aligned}$$

In other words, the transition to the next state depends only on the previous k states. For example, if X_n represents the position of a particle moving with constant velocity at time n , then the system is a second-order Markov chain because the previous two position measurements are needed to infer the particle's velocity.

Show that we can “embed” $(X_n)_{n \in \mathbb{N}}$ into a *first-order* Markov chain $(Z_n)_{n \in \mathbb{N}}$ with an augmented state space, in the sense that X_n can be recovered from Z_n . This allows us to apply algorithms such as the Viterbi algorithm to systems with higher orders of dependence.

2. Hidden Markov Models

A hidden Markov model (HMM) is a Markov chain $\{X_n\}_{n=0}^\infty$ in which the states are considered “hidden” or “latent”. In other words, we do not directly observe $\{X_n\}_{n=0}^\infty$. Instead, we observe $\{Y_n\}_{n=0}^\infty$, where $Q(x, y)$ is the probability that state x will emit observation y . π_0 is the initial distribution for the Markov chain, and P is the transition matrix.

- (a) What is $\mathbb{P}(X_0 = x_0, Y_0 = y_0, \dots, X_n = x_n, Y_n = y_n)$, where n is a positive integer, x_0, \dots, x_n are hidden states, and y_0, \dots, y_n are observations?
- (b) What is $\mathbb{P}(X_0 = x_0 \mid Y_0 = y_0)$?
- (c) We observe (y_0, \dots, y_n) and we would like to find the most likely sequence of hidden states (x_0, \dots, x_n) which gave rise to the observations. Let

$$U(x_m, m) = \max_{x_{m+1}, \dots, x_n \in \mathcal{X}} \mathbb{P}(X_m = x_m, X_{m+1:n} = x_{m+1:n}, Y_{0:n} = y_{0:n})$$

denote the largest probability for a sequence of hidden states beginning at state x_m at time $m \in \mathbb{N}$, along with the observations (y_0, \dots, y_n) . Develop a recursion for $U(x_m, m)$ in terms of $U(x_{m+1}, m+1)$, $x_{m+1} \in \mathcal{X}$.

3. Most Likely Sequence of States

In this problem, we give an example of an HMM and a sequence of observations which demonstrates that the most likely sequence of hidden states (i.e., the output of the Viterbi algorithm) is *not* the same as computing the most likely state at each time. Your task is to verify that the following example works:

Consider a HMM with two states $\{0, 1\}$ and the hidden state is observed through a BSC with error probability $1/3$. The hidden state transitions according to $P(0, 0) = P(1, 1) = 3/4$. Assume that the initial state is equally likely to be 0 or 1. We see the observation 0 at time 0 and 1 at time 1.

4. Cheapest Fare using HMM

Companies A and B run identical buses in Berkeley, where company A has higher fares. The number of people in buses run by company A and B is a Poisson random variable with rate 10 and 20, respectively. You are counting the number of people on the buses at a bus-stop where only one bus comes each hour; let X_k be the bus company and N_k be the number of people in the k -th hour, respectively. A Markov chain with transition probabilities $P(X_{k+1} = B|X_k = A) = 0.7$ and $P(X_{k+1} = A|X_k = B) = 0.8$ determines the company of buses arriving.

- (a) Let the initial state be $X_1 = A$. What is $P(X_2 = A|N_2 = 13)$?
- (b) Assuming that the initial state was $X_1 = A$, say you observed the number of people $N_1 = 7$, $N_2 = 21$ and $N_3 = 9$ in the first three hours. What is your MLSE estimate for the sequence of first three buses?
- (c) You board the bus in the fourth hour. Assuming that the most likely sequence is true, what is the probability that you board the bus with a cheaper fare?