Solutions to Exam 2 **ECE 534 Fall 2008**

Problem 1 (12 points) Suppose $a = (a_0, a_1, a_2, a_3)$ is a probability vector to be estimated by observing $Y = (Y_t : 1 \le t \le T)$. Assume $Y_1, \ldots, Y_T$ are independent, and $P\{Y_t = i\} = a_i$ for $1 \le t \le T$ and $i \in \{0, 1, 2, 3\}$.
(a)(6 points) Determine the maximum likelihood estimate, $\widehat{a}_{ML}(y)$, given a particular observation $y = (y_1, \ldots, y_T)$. Justify your answer.

(b)(6 points) Suppose in addition to the above that a has the form $a_i = \binom{3}{i} q^i (1-q)^{n-i}$, or in other words, that a is a binomial pmf with parameters (3,q). Determine the maximum likelihood estimate, $\widehat{q}_{ML}(y)$, given a particular observation $y = (y_1, \ldots, y_T)$. Justify your answer.

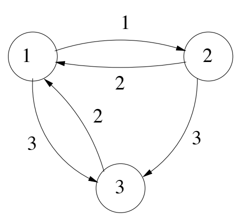
(a) The likelihood to be maximized with respect to a is $p(y|a) = a_{y_1} \cdots a_{y_T} = a_0^{n_0} a_1^{n_1} a_2^{n_2} a_3^{n_3}$, where $n_i = |\{t : y_t = i\}|$. The log likelihood is $\ln p(y|a) = n_0 \ln a_0 + n_1 \ln a_1 + n_2 \ln a_2 + n_3 \ln a_3$. By a lemma in the notes, this is maximized by the empirical distribution of the observations, namely $a_i = \frac{n_i}{T}$ for $\{0 \le i \le 3\}$. That is, $\widehat{a}_{ML} = (\frac{n_0}{T}, \frac{n_1}{T}, \frac{n_2}{T}, \frac{n_3}{T})$. (One proof of the lemma is as follows. The likelihood $\ln p(y|a)$ is concave in a and the constraint on a is linear. The maximizer must be a stationary point of the Lagrangian $L(y, a) = \ln p(y|a) + \lambda(a_0 + a_1 + a_2 + a_3 - 1)$. Setting $\frac{\partial L}{\partial y_i} = 0$ yields $\frac{n_i}{a_i} = \lambda$, for $0 \le i \le 3$, and since a must be a probability distribution, the maximizer is given as above.) Another approach to this problem is to note that Y is the observation vector for an HMM with a single state ($N_S = 1$) and the B matrix given by the row vector $B = (a_0, a_1, a_2, a_3)$. The estimator $\widehat{a}_{ML}$ is produced by one iteration of the Baum-Welch algorithm.)

(b) Under the additional assumption, the likelihood becomes

$$p(y|q) = \prod_{t=1}^{T} \left[\binom{3}{y_t} q^{y_t} (1-q)^{3-y_t} \right] = c q^s (1-q)^{3T-s},$$

where $s = y_1 + \cdots + y_T$, and c depends on y but not on q. The log likelihood is $\ln c + s \ln(q) + (3T - s) \ln(1-q)$. Maximizing over q yields $\hat{q}_{ML} = \frac{s}{3T}$. (An alternative way to think about this is to realize that each Y_t can be viewed as the sum of three independent Bernoulli(q) random variables, and s can be viewed as the observed sum of $3T$ independent Bernoulli(q) random variables.)

Problem 2 (12 points) Consider a continuous-time Markov process with the rate transition diagram shown:



(a) (2+4 points) Write down the rate transition matrix Q and find the equilibrium distribution π.

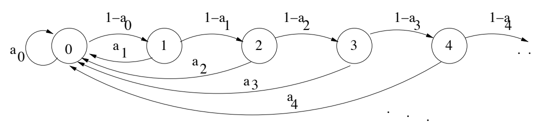
(b) (6 points) Find $P[X_t \ne 3 \text{ for } 0 \le t \le 5 | X_0 = 1]$. Be as explicit as possible. (Hint: for the particular rates given, this can be solved with very little calculation.)

(a)

$$Q = \begin{pmatrix} -4 & 1 & 3 \\ 2 & -5 & 3 \\ 2 & 0 & -2 \end{pmatrix}.$$

The condition $\pi Q = 0$ implies $4\pi_1 = 2(\pi_2 + \pi_3)$ and $\pi_1 = 5\pi_2$. Substituting the second equation into the first yields $20\pi_2 = 2(\pi_2 + \pi_3)$, or $\pi_3 = 9\pi_2$. Hence, π is proportional to $(5, 1, 9)$ and so $\pi = (\frac{5}{15}, \frac{1}{15}, \frac{9}{15}) = (\frac{1}{3}, \frac{1}{15}, \frac{3}{5})$.
(b) Since the jump rate to state 3 is 3 for both states one and two, if the intial state is either 1 or 2, the time τ that the process first reaches state 3 is exponentially distributed with parameter 3. Thus, $P[X_t \ne 3 \text{ for } 0 \le t \le 5 | X_0 = 1] = e^{-(3)(5)} = e^{-15}$.

Problem 3 (16 points) Consider a time-homogeneous Markov process with state space $\mathbb{Z}_+$ and the one-step transition probability diagram shown, where $0 < a_k < 1$ for all $k \geq 0$. Note that this is *not* a birth-death process.



(a)(4 points) Sketch the one-step transition probability matrix. (The matrix has infinitely many entries, so use dot dot dot "$\cdots$" notation.)

(b) (4 points) Under what condition on the a's is the process recurrent?

(c) (4 points) Under what condition on the a's is the process positive recurrent?

(d) (4 points) Find the equilibrium distribution π in terms of the a's, and in particular, indicate when an equilibrium distribution exists. Be as explicit as possible.

(a)

$$P = \begin{pmatrix} a_0 & 1-a_0 & 0 & 0 & 0 & \cdots \\ a_1 & 0 & 1-a_1 & 0 & 0 & \cdots \\ a_2 & 0 & 0 & 1-a_2 & 0 & \cdots \\ a_3 & 0 & 0 & 0 & 1-a_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

(b) Since the process is irreducible, all states are recurrent if and only if state 0 is recurrent. By definition, state 0 is recurrent if and only if $P[\tau_0 = +\infty | X_0 = 0] = 0$, where $\tau_0 = \min\{k \geq 1 : X_k = 0\}$. Observe that $P[\tau_0 > n | X_0 = 0] = (1-a_0)(1-a_1)\cdots(1-a_n)$. Taking the limit as $n \to \infty$ and using the continuity of probability yields $P[\tau_0 = +\infty | X_0 = 0] = (1-a_0)(1-a_1)(1-a_2)\cdots$. Hence, the process is recurrent if and only if $(1-a_0)(1-a_1)(1-a_2)\cdots = 0$. (Remark: By taking logs and using Taylor's theorem, this condition can be shown to be equivalent to the condition $\sum_k a_k = \infty$. This result can also be deduced from the Borel-Cantelli lemma.) (Remark: Since $P[\tau_0 < +\infty | X_0 = 0] = \sum_{k=1}^{\infty} P\{\tau = k\} = a_0 + \sum_{k=1}^{\infty}(1-a_0)\cdots(1-a_{k-1})a_k$, recurrence is also equivalent to this sum being equal to one.)

(c) Since the process is irreducible, all states are positive recurrent if and only if state 0 is positive recurrent, which, by definition, is true if and only if $M_0 < +\infty$, where

$$M_0 = E[\tau_0 | X_0 = 0] = \sum_{n=1}^{\infty} nP[\tau_0 = n | X_0 = 0] = \sum_{n=1}^{\infty} n(1-a_0)(1-a_1)\cdots(1-a_{n-1})a_n.$$

Alternatively,

$$M_0 = E[\tau_0 | X_0 = 0] = \sum_{n=1}^{\infty} P[\tau_0 \geq n | X_0 = 0] = 1 + (1-a_0) + (1-a_0)(1-a_1) + \cdots$$

(d) The equilbrium equation $\pi = \pi P$ is equivalent to the equations

$$\pi_0 = \pi_0 a_0 + \pi_1 a_1 + \cdots \quad (1)$$

$$\pi_k = \pi_{k-1}(1-a_{k-1}) \quad \text{for } k \geq 1 \quad (2)$$

By (2), any equilibrium distribution must satisfy $\pi_k = \pi_0(1-a_0)(1-a_1)\cdots(1-a_{k-1})$ for $k \geq 1$. Let $S = 1 + (1-a_0) + (1-a_0)(1-a_1) + \cdots$. On one hand, if $S = +\infty$ there is no way to select π_0 to make the π_k's sum to one, and thus no equilibrium distribution exists. On the other hand, if $S < \infty$, then the probability distribution π defined by $\pi_k = \frac{(1-a_0)(1-a_1)\cdots(1-a_{k-1})}{S}$ satisfies (2). Since π and πP both sum to one, π also satisfies the remaining equation, (1), and therefore π is an equilibrium distribution.

(Remark: Another way to do part (c) is to do part (d) first, and quote the fact that an irreducible, discrete-time Markov process is positive recurrent if and only if there exists an equilibrium distribution. Note that M_0 is the same as S. Or, similarly, the question of when there exists an equilibrium distribution, which is part of (d), could be answered using part (c).)