HOMEWORK 2

STA 624.01, Applied Stochastic Processes
Spring Semester, 2009

Due: Tuesday, February 3, 2009

Readings: Chapter 1 of text, Tutorial 2 on MATLAB

Note: the computer problems require simulation and the use of a computer. You are allowed (encouraged, even) to use a computer in solving the other problems as well.

When giving numerical answers, please give results to four siginificant figures unless they are integer answers. So $1/2 = .5000$ for example. Also box your numerical answers.

Regular Problems

1 In a simple model of the weather, each day is classified as 'sunny' or 'cloudy' and the sequence of weathers is modelled as a Markov chain with transition probability matrix

$$P = \begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix}$$

Here the first row and column correspond to state 'sunny'.

(a) What is the probability that a cloudy day is followed by sunny one?

(b) What is the probability that a cloudy day is followed by two sunny ones in a row?

(c) If Friday is sunny, what is the probability that the following Sunday is sunny too?

(d) Compute P^2, P^4, P^8, P^{16} and P^{32} by repeatedly squaring the matrices you compute. What is your conclusion? What is the practical implication regarding the weather?

2 Consider the telecommunications Markov chain with transition matrix:

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

where $0 \le p, q \le 1$.

Suppose you flip a fair coin to start the chain in the 0 or 1 position.

a) What is the probability vector for X_0?

b) What is the distribution of X_1 (the chain after one step)?

c) What is the distribution of X_2 (the chain after two steps)?

d) Compute $P(X_{10} = 0 | X_0 = 0)$ if $p = 1/4$ and $q = 1/3$.

e) Compute $P(X_{10} = 1, X_{15} = 0 | X_0 = 1)$, if $p = 1/4$ and $q = 1/2$.

3 A simple model of DNA is that the sequence of the bases A, C, G and T forms a Markov chain. Assume this chain has transition probability matrix

$$P = \begin{pmatrix} 0.3 & 0.22 & 0.21 & 0.27 \\ 0.28 & 0.22 & 0.30 & 0.20 \\ 0.23 & 0.32 & 0.23 & 0.22 \\ 0.18 & 0.22 & 0.30 & 0.30 \end{pmatrix}.$$

here the first row and column corresponds to A, etc. A subsequence that has been analysed reads AT-GxxCGT, where 'xx' means that one is uncertain about these two bases. Biological considerations however suggest that these two symbols are either AC ot TG. Which of these two alternatives is the most probable?

4 Consider two urns A and B containing N balls. An experiment is performed in which one of the N balls is selected with probability depending on the urn contents (i.e., if A currently has k balls, a ball is chosen from A with probability k/N or from B with probability $(N-k)/N$). Then, an urn is selected and then depositing the selected ball in the selected urn. Urn A is chosen with probability k/N or urn B is chosen with probability $(N-k)/N$. Determine the transition matrix of the Markov chain with states represented by the contents of A.

5 A matrix $P = \{P_{i,j}\}$ is called *stochastic* if $P_{i,j} \geq 0$ for all i, j and $\sum_j P_{i,j} = 1$ for all i. A matrix P is called *doubly stochastic* if in addition to the above, $\sum_i P_{i,j} = 1$ for all j. Prove that if a finite irreducible Markov chain has doubly stochastic transition probability matrix, the stationary probabilities π_i for all i exist and are equal.

Computer Problems

For this problem, please print out all code used and all results.

This Markov chain is called simple random walk with reflecting boundaries. The state space is $\{1, 2, \ldots, n\}$. It is defined as follows:

$$\begin{aligned}
P(X_{t+1} = i+1 | X_t = i) &= p, \ \forall i \in \{2, \ldots, n-1\} \\
P(X_{t+1} = i-1 | X_t = i) &= 1-p, \ \forall i \in \{2, \ldots, n-1\} \\
P(X_{t+1} = n-1 | X_t = n) &= 1-p \\
P(X_{t+1} = n | X_t = n) &= p \\
P(X_{t+1} = 1 | X_t = 1) &= 1-p \\
P(X_{t+1} = 2 | X_t = 1) &= p.
\end{aligned}$$

a) Write code for simulating this Markov chain.

b) Find the limiting distribution when $n = 10$ and $n = 15$ for $p = 0.2$, 0.5, 0.6 by simulating the chain multiple times starting from $X_0 = 1$.

c) For $n = 5$ with $p = 0.2$, 0.5, 0.6, estimate the expected number of steps needed to return to i starting at i for all $i \in \{1, 2, 3, 4, 5\}$.