Problem Set 9 (due Tues May 6)
Markov processes and random walks

April 28, 2003

1. Consider a discrete-time Markov process with transition matrix Q. By definition, this matrix satisfies the constraints $Q_{ij} \geq 0$ and $\sum_i Q_{ij} = 1$. Any right eigenvector v of Q satisfies the equation

$$\sum_j Q_{ij} v_j = \lambda v_i$$

   where λ is an eigenvalue. Prove that $\lambda \leq 1$. Hint: use the triangle inequality, $|x + y| \leq |x| + |y|$.

2. *Cyclic three-state Markov model in continuous time.* Imagine a hypothetical channel with three states, numbered 1, 2, and 3. Only three transitions are possible: $1 \to 2$ with rate constant α, $2 \to 3$ with rate constant β, and $3 \to 1$ with rate constant γ.

   (a) Let $p = (p_1, p_2, p_3)$ be the vector of probabilities for being in each of the three states. They are governed by dynamical equations of the form

$$\frac{dp_i}{dt} = \sum_j R_{ij} p_j$$

   Write down the matrix R.

   (b) Find the stationary distribution, defined as the probability vector that is a steady state of these differential equations.

3. *Two-state Markov process in continuous time.* Imagine a hypothetical ion channel with two states, called "C" for closed and "O" for open. The rate constant of the C→ O transition is α, while for the O→C transition it's β. Consider a population of N channels. Suppose that all are closed at time $t = 0$.

   (a) Derive a formula for the probability that the channel is open as a function of time. (Hint: the probability starts out zero, and then approaches a steady state exponentially.)

   (b) The MATLAB function `rand` gives a uniformly distributed random number between zero and one. Prove that `-mu*log(rand)` yields an exponentially distributed random variable with mean μ.

   (c) Simulate the opening and closing of a single channel by alternately drawing time intervals from two exponential distributions, one with mean $1/\alpha$, and the other with mean $1/\beta$. Run your simulation from $t = 0$ to $t = 30$, assuming that the channel starts out closed at time $t = 0$, and that $\alpha = 0.2$ and $\beta = 0.1$. Graph the state of the channel as a function of time.

   (d) To simulate a population of N two-state channels, run the single channel simulation N times, each time starting with the channel closed at $t = 0$. Using these results, compute the total number of open channels in the population as a function of time, again from $t = 0$ to $t = 30$. Do this for $N = 1$, 10, 100, and 1000, and plot the results. For comparison, also graph on the same plot the analytical result for the average number of open channels, which you calculated previously. The agreement between the stochastic simulation and the analytical result should become good for large N.

4. Simulation of diffusion in one dimension. Simulate a population of 1000 random walkers in MATLAB using the following code:

```
n=1000;
x=zeros(1,n);
for t=1:100;
  x=x+sign(randn(1,n));
  hist(x,-50:50);
  axis([-50 50 0 1000]);
  drawnow;
end
```

The n-dimensional vector x in memory stores the positions of the random walkers. They're initialized at $x = 0$ at time $t = 0$. At each time step, ± 1 is added at random to each position. The extra argument to the `hist` command specifies the position of the bin centers, and makes the histogram look nice. If you run the code, you should see that the histogram becomes wider as time increases.

(a) Change the code to keep track of the variance of x as a function of time. Make a graph of this relationship. It should be linear. What is the proportionality constant?

(b) The probability distribution of the positions can be well-approximated by a Gaussian. Write down a formula for the approximation as a function of x and t, and modify the above code to plot the Gaussian so that it lines up with the bars of the histogram. Note that you will have to convert between probabilities, which are normalized, and raw counts.