Lab The Fourth
Markov Processes
Due Monday, April 9, at 5PM

A *Markov Process* is defined as a set of states together with a set of probabilities of moving from one state to the next at any given time. This movement is dependent only on which state an object is in at any given time (and not on the time itself, other states or their populations, prior states, or other factors). The idea is probably best expressed through....

1 A Basic Example

> Suppose that, in a certain community, in any given year, 90 percent of those that live in the suburbs remain in the suburbs, and the other 10 percent move to the city. Suppose also that 80 percent of those living in the city remain there, while the other 20 percent move to the suburbs.

Thus, the probability that any given resident moves is dependent only on whether they are currently living in the city or the suburbs. We seek a mathematical representation of this sociological phenomenon.

Problem 1 *Determine the* transition matrix *A, which has rows and columns indexed by the possible states, with the entry $A_{(i,j)}$ containing the probability of moving from state j to state i in any given year.*

You can set up this matrix in Maple by using the 'Matrix' option on your left toolbar. Maple will let you specify the number of rows and columns and will drop a generic matrix into your worksheet. From there, it is straightforward to put in your desired values. If you ask nicely (ie, by righclicking on your filled in matrix), Maple will even give you the LaTeX code to drop into your document for the matrices that you create. Beware: In this version of Maple, when multiplying matrices, you must use a period (.) rather than an asterisk to denote the multiplication. You can exponentiate a matrix just as you would a number.

Question 1 *What is the sum of each of the columns in your matrix? Why does this make sense?*

Any matrix with nonnegative entries that satisfies the property you discovered in Question 1 can serve as a transition matrix of a Markov Process. We can use our matrix to determine the percentages of people living in each locale after a given year, provided we have the initial conditions.

A *state vector* $\mathbf{x}$ is an $n \times 1$ matrix, where n is the number of states in our Markov Process. The ith entry in the matrix is the percentage of the population in state i. Your vector will have a similar property to that you discovered in Question 1.

We can use this state vector to determine the population from one year to the next as follows: If $\mathbf{x}_i$ is the state vector for the population in year i then $\mathbf{x}_{i+1} = A\mathbf{x}_i$, that is, we multiply the state vector by the transition matrix.

Problem 2 *Suppose that in the first year, half of the population lives in the suburbs and half in the city. Determine the state vector and the population in each after the second year. (Do this determination both by reasoning and by matrix multiplication, and verify that you're really doing the same calculation).*

Problem 3 *Determine, using the matrix multiplication and Maple, the percentage of the population living in the suburbs and the city after 2, 3, 5, and 10 years.*

Question 2 *Do you expect there to be a 'long term' population percentage in the suburbs and the city? If so, how would you calculate it?*

Problem 4 *Suppose that the initial population percentages are 20 % suburban and 80% city. Calculate the populations after 1, 2, 3, 5 and 10 years, and the long term projection for the populations. Do the same for an initial percentage of 80% suburban and 20% city.*

Question 3 *What do you notice about your long term population percentages in all 3 scenarios? Can you explain why the populations might turn out this way?*

2 Absorbing States and the Fundamental Matrix

The Markov Process described in the previous example was *regular*, as someone in any given state had a positive probability of moving to any other state. In this next example, this will not be the case. Some states will be *absorbing*, in the sense that once someone in the population enters that state, they cannot leave it.

> Suppose that there is an outbreak of a disease in a population. The disease will affect 15% of the well population in any given month. Of those sick with the disease, 50% will recover (and thereafter be immune to the disease), 30% will remain sick into the next month, and 20% will die from the disease.

Question 4 *How many states are there in this Markov Process? Which ones are absorbing? Why?*

Problem 5 *Set up the transition matrix for this Markov Process, and give a table of the percentages of people that are in each state after 1, 2, 6, and 12 months, assuming a population that is all well to begin with. Determine also the long-term prognosis for this population.*

Another question of interest is the number of months that one can expect to be afflicted with this disease. We walk through this problem in the next few steps.

We begin by rewriting the transition matrix as

$$\begin{bmatrix} I_k & R \\ 0 & Q \end{bmatrix},$$

where I_k is the $k \times k$ identity matrix, R is a $k \times (n-k)$ matrix, and Q is a $(n-k) \times (n-k)$ square matrix. We achieve this form by writing our k absorbing states first. We then calculate the matrix

$$F = (I - Q)^{-1}$$

where I here is the $(n-k) \times (n-k)$ identity matrix. [1]. Since inverses are icky to compute in general, we will enlist the help of Maple for this task.

Problem 6 *Rewrite your matrix for the outbreak problem in the desired form. Determine the matrices R, Q, and $(I - Q)^{-1}$.*

The matrix $(I-Q)^{-1}$ is known as the *fundamental matrix*. Its rows and columns are indexed by the non-absorbing states in our system, the entries give expected times spent in each state. Specifically, $(I - Q)^{-1}_{i,j}$ gives the amount of time one can expect to spend in state i given that they started in state j.

Question 5 *What is the expected time that a well person can expect to remain well? What is the expected time that a sick person can expect to remain sick (before either healing or dying). What does the 0 mean in the fundamental matrix in this case?*

Problem 7 *Verify that the long term behavior of the population that you determined from the system is explained by the matrix*

$$\begin{bmatrix} I_k & R(I - Q)^{-1} \\ 0 & 0 \end{bmatrix}$$

3 The Gambler

For this section, you'll apply the techniques from the previous sections to a problem involving a gambler. The gambler plays a simple game (blackjack, or better yet, casino war), in which their only possibilities are to win or lose each hand. Suppose, initially, that each hand plays for $5, and that there is a 50% chance of winning. The gambler will play until he has reached a predetermined amount (say, $20), or until they run out of money. Find out the probability of reaching $20 and

[1]Stay tuned in class for the algebraic legedermain that gives us this formula

the expected number of hands played if the starting stake is \$5, \$10, or \$15.

Now tweak the problem to include other feasible conditions (a larger starting stake, a higher probability of winning or of losing, a possibility of a tie or of doubling up). Compare the different effects of changing the starting conditions. In particular, what conditions on a game might enable a recreational gambler (such as myself) to maximize the length of time that they play in the casino? Give a prose account of the different conditions that you're placing on the game, and a mathematical exploration of how those conditions would actually play out.