Markov Matrices problem set 2011

You're studying a system that can be in one of N discrete states. At each timestep, the probability of jumping from state j to state i is m_{ij}. We can think of the numbers m_{ij} as being the elements of a square matrix M. (Some of the m_{ij} could be zero; for example if there is no way to jump from state 2 to state 3, then $m_{32}=0$.) Note that because the numbers m_{ij} are probabilities, they also satisfy

$$1 \geq m_{ij} \geq 0$$

- We know for certain that after a jump, the system will be in one of the N states. Show that, consequently, the sum along each column of M (i.e., sum over i) equals 1.

A matrix M whose elements satisfy the conditions above is called a *Markov matrix.*

Consider a vector p with elements p_i , such that .

$$\sum_i p_i = 1 \quad ; \quad p_i \geq 0$$

We're going to call a vector with such properties a *probability vector.* We think of p as a vector whose i-th element represents the probability that the system is in state i. The fact that the elements of p add up to 1 represents the idea that the N states are *all* the possible states the system could be in-- therefore the system *has* to be in one of the N states.

Suppose p is a probability vector that represents the distribution over states for the system at some time. Then Mp will result in a vector whose i-th element represents the probability that the system is in state i after one timestep; M^2p will result in a vector whose i-th element represents the probability that the system is in state i after 2 timesteps; and so on: M^kp will result in a vector whose i-th element represents the probability that the system is in state i after k timesteps.

- If p is a stable state for the distribution of the system, i.e., p doesn't change as timesteps go by, what kind of vector is p (in addition to being a probability vector)?

Now consider the Markov matrix

$$M = \begin{pmatrix} 0.8 & 0.5 \\ 0.2 & 0.5 \end{pmatrix}$$

Suppose that you have a system that at time zero is known to be in state #2. That is, it is in state #1 with probability 0 and in state #2 with probability 1.

- Write p down at time zero. What is p after one timestep? What is p after two timesteps? Which way is the probability flowing as time elapses: from state #1 to state #2, or from state #2 to state #1? As time goes to infinity, will *all* the probability flow to state #1, that is, will $p=(1, 0)^T$ become a stable state for the distribution of the system? Is $p=(1, 0)^T$ an eigenvector of M?

M has two eigenvectors; with their corresponding eigenvalues, these are

Let's express p as a weighted sum of these two eigenvectors; specifically, $p = v_1 + w\, v_2$.

$$v_1 = \begin{pmatrix} 0.7143 \\ 0.2857 \end{pmatrix}, \quad \lambda_1 = 1 \qquad v_2 = \begin{pmatrix} -0.5 \\ 0.5 \end{pmatrix}, \quad \lambda_2 = 0.3$$

- Compute w. (You don't need to compute it exactly, approximately is o.k.) Then write an expression for Mp in terms of the two eigenvectors. Then write an expression for $M^k p$ in terms of the two eigenvectors. What is $M^k p$ as k goes to infinity? As time goes to infinity, what's the probability of finding the system in state #1? Is your answer here consistent with your answer to the immediately previous bullet?

Assuming you've been paying attention, at this point you've realized that computing eigenvectors and eigenvalues of Markov matrices is pretty useful. There are some nice general properties that can be proved for them.

- Assume that p is both a probability vector *and* an eigenvector of M. Recall that the sum of the elements of p is 1. What is the sum of the elements of Mp? Based on your answer, can p have an eigenvalue bigger than 1? Can it have an eigenvalue smaller than 1? Why?

- Recall that the sum of each column of M equals 1. Assume that v is an eigenvector of M, although it's not necessarily a probability vector. (For example, some elements of v might be negative.) In other words,

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

Show that either

$$\lambda = 1 \quad \text{or} \quad \sum_i v_i = 0.$$

Is your answer consistent with your answer to the immediately previous bullet?

For bonus points, consider the fact that l*a+bl <= lal + lbl.*

- Prove that for any Markov matrix, all its eigenvalues λ satisfy lλl <= 1. What does this tell us about which eigenvectors we should pay attention to as the number of timesteps goes to infinity?

Only eigenvectors with eigenvalue 1 will matter for the long-term state of the system. Components along any other eigenvector will gradually fade away asymptotically to zero; none will grow with time.