Markov Chains

For finite state automata, we considered models with transitions between states based on input data. Next we will consider a collection of states with probabilistic transitions. We allow for a fixed start state or a random start state from some initial distribution. We draw this again as a directed graph where a directed edge e_{ij} is labeled with the probability of moving from i to j see Figure 1.

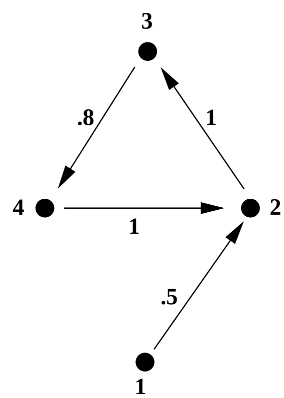

Figure 1: A Markov chain with 4 states.

For each state, the sum of the probabilities on the edges directed out of any given state must sum to one. In the figure above we do not see this explicitly. For example, state 3 has only one edge directed out of it and this edge has probability .8. We assume that there is a loop at state 3 with probability .2. Hence we can stay at a given state. In general, we will not draw the loops but always assume they exist with probability equal to 1 minus the sum of probabilities on outward edges.

Now we formally define a *Markov chain*. Suppose we have a collection of random variables $\{X_n\}$. We think of X_i as the state after i transitions and X_0 as the initial distribution. We require the following *Markov property*: $p_{ij} = p(X_{n+1} = j | X_n = i) = p(X_{n+1} = j | X_n = i, X_{n-1} = i-1, \ldots, X_0)$. This property states that the probability of moving to state j from state i

does not depend on which states we were in previous to state i. We think of this as a property of *no memory*. Hence we only need to know the one-step transition probabilities and initial distribution. We often record these transition probabilities in a matrix where entry ij is the probability of moving from state i to state j.

Example The transition matrix for the chain in Figure 1 is:

$$\mathbf{T} = \begin{pmatrix} .5 & .5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & .2 & .8 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Notice that these transition matrices will always have row sum equal to one. Any such matrix is called *stochastic*.

We will often be interested in knowing the probability of being in a certain state after many steps. Let $r_{ij}(n) = p(X_n = j|X_0 = i)$. Hence $r_{ij}(n)$ is the probability of being in state j in n steps having started in state i. In general we have the following recursion: $r_{ij}(n) = r_{i1}p_{1j} + r_{i2}(n-1)p_{2j} + \ldots = \Sigma_{k=1}^{m} r_{ik}(n-1)p_{kj}$. If we have a random initial state, we get: $p(X_n = j) = \Sigma_{i=1}^{m} p(X_0 = i)r_{ij}(n)$.

We need to consider various types of states. A state i is *recurrent* if from i wherever you go there is always some way back to state i. A state is *transient* if it is not recurrent. In the limit as n tends toward infinity, $p(X_n = i)$ tends toward 0 if i is a transient state. A state i is *periodic* if it is recurrent and there exists a number d s.t. $p_{ii}(n) = 0$ if n is not a multiple of d. A *recurrent class* is a set of recurrent states which can all reach each other. These states are said to *communicate*.

A Markov chain is called *ergodic* if it has only one recurrent class and no periodic states.