Hand out. March 10, 2008 Math 504

Convergent Finite Markov Chains

1 Introduction

Consider a finite state Markov chain with one-step probability transition matrix P and state probability distribution vector $\pi^{(n)}$ at time $n \geq 0$. Then

$$\pi^{(n+1)} = \pi^{(n)} P \,, \quad n = 0, 1, \cdots \,.$$

A fundamental question is whether or not the process approaches a limit in the long-run. In other words, given an arbitrary initial state probability distribution $\pi^{(0)}$, do the state probability distributions $\pi^{(n)}$ converge as $n \to \infty$. Since

$$\pi^{(n)} = \pi^{(0)} P^n \,,$$

for each $n \geq 0$, and since $\pi^{(0)}$ is arbitrary, this question is equivalent to asking whether the powers of the transition matrix converge as $n \to \infty$. A Markov chain with transition matrix P will be called **convergent** if P^n converges as $n \to \infty$. In this case, we shall also refer to the transition matrix as being convergent.

2 Structure and Properties of Finite Markov Chains

The study of Markov chains hinges on the notion of recurrent and transient states. For a state i, let T_{ii} denote the time until the process first returns to state i, given that it starts in state i. Then state i is said to be **recurrent** if $P(T_{ii} < \infty) = 1$, and otherwise, if $P(T_{ii} < \infty) < 1$, state i is said to be **transient**. Thus, a state is transient if the process, having started in that state and perhaps having returned to that state a number of times, will eventually leave that state forever. It can be shown that if states i and j are transient then

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty \,.$$

This expression is the expected number of visits to state j, given that the process starts in state i. Thus, for transient states i and j, we have

$$p_{ij}^{(n)} \to 0 \,, \quad \text{as} \quad n \to \infty \,. \tag{1}$$

One important consequence of this result is that not all states of a finite chain can be transient.

The process moves among the states randomly, according to the transition probabilities. Two states i and j are said to **communicate** if

$$p_{ij}^{(n)} > 0 \,, \quad \text{and} \quad p_{ji}^{(m)} > 0 \,,$$

for some n and m. In other words, two states communicate if it is possible to travel from one state to the other and back again. If state n is recurrent, and it communicates with state m, then state m must be recurrent also.

A set of states S is **closed** if $p_{ij} = 0$ whenever $i \in S$ and $j \notin S$. Note that if the process enters a closed set, then it will never leave that set. A chain is said to be **irreducible** if there is no proper closed subset. Otherwise, the chain is called **reducible**. For a reducible chain, after possibly re-ordering the states, the transition matrix can be written in the form

$$P = \begin{bmatrix} P_1 & 0 \\ P_2 & P_3 \end{bmatrix} .$$

The matrix P_1 is the transition matrix for the sub-chain consisting of a closed set of states.

It follows from these concepts that for an arbitrary finite Markov chain, the probability transition matrix, again after possibly re-ordering the states, can be written in the so-called **canonical form**:

$$P = \begin{bmatrix} D & 0 \\ R & Q \end{bmatrix} , \tag{2}$$

where the matrix D is block diagonal

$$D = \begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & 0 & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & D_k \end{bmatrix} . \tag{3}$$

To see how this form comes about, start with a recurrent state, and find all states that communicate with it. These states form an irreducible, closed set. Let the matrix consisting of the one-step transition probabilities among these states form the first block D_1. Next, find a recurrent state (if any) that is not in the set obtained previously, and determine all states that communicate with it. This set of states forms another irreducible, closed set. Let the matrix consisting of the one-step transition probabilities among these states form the second block D_2. Continue in this fashion until no recurrent states are left. The remaining states are transient, and the matrix consisting of the one-step transition probabilities among these states is the matrix Q. The matrix R in (2) consists of the one-step transition probabilities from transient states to recurrent states. Finally, the matrix of zeros in (2) appears in the upper righthand corner, since it is not possible to go from a recurrent state to a transient state.

The canonical form reveals much about the behavior of the chain. If the process starts in an irreducible class, it will stay there and ultimately approach the limiting behavior of that class, if any. On the other hand, if the process starts in a transient state, it will ultimately move into one of the irreducible classes. In fact, as indicated by (1), $Q^n \to 0$ as $n \to \infty$, since the (i, j)-th entry of this matrix is the probability that the process is in transient state j, after n steps, given that it started in a transient state i. Further, entry (k, s) of the matrix $Q^n R$ is the probability that when the process starts in transient state k, it will enter for the first time, one of the irreducible sub-chains in $n+1$ steps, and will do so at state s. Summing over n, it follows that entry (k, s) of the matrix

$$R + QR + Q^2R + \cdots = (I + Q + Q^2 + \cdots)R \, ,$$

is the probability that when the process starts in transient state k, it will ultimately enter the recurrent classes for the first time at state s. Using the Neumann expansion $N = (I - Q)^{-1} = I + Q + Q^2 + \cdots$, we can write this matrix simply as NR.

Example Suppose the one-step transition matrix of a Markov chain is given by

$$P = \begin{bmatrix} 0.800 & 0.200 & 0 & 0 & 0 & 0 & 0 \\ 0.300 & 0.700 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.00 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.400 & 0.600 & 0 & 0 \\ 0 & 0 & 0 & 0.200 & 0.800 & 0 & 0 \\ 0.100 & 0.200 & 0.050 & 0.150 & 0.200 & 0.100 & 0.200 \\ 0.300 & 0.100 & 0.200 & 0.100 & 0.050 & 0.100 & 0.150 \end{bmatrix}.$$

The states have been numbered so that the matrix is in canonical form. The first recurrence class consists of states 1 and 2, the second recurrence class consists of state 3 only, and is absorbing, while the third and last recurrence class consists of states 4 and 5. States 6 and 7 are transient. The matrices R and Q are

$$R = \begin{bmatrix} 0.100 & 0.200 & 0.050 & 0.150 & 0.200 \\ 0.300 & 0.100 & 0.200 & 0.100 & 0.050 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.150 & 0.200 \\ 0.100 & 0.050 \end{bmatrix}.$$

Thus

$$NR = (I - Q)^{-1} R = \begin{bmatrix} 0.195 & 0.255 & 0.111 & 0.198 & 0.242 \\ 0.376 & 0.148 & 0.248 & 0.141 & 0.087 \end{bmatrix}.$$

Consider the first row of this matrix. The first entry shows that 19.5% of the entities that start in state 6, will enter the recurrent classes for the first time at state 1. Similarly, the second entry shows that 25.5% of the entities that start in state 6, will enter the recurrent classes for the first time at state 2. Thus, 19.5%+25.5% = 45% of the entities that start in state 6 will ultimately enter the first recurrence class. These entities however, once having entered this recurrent class, must then circulate among the states and ultimately become distributed according to limiting behavior of the class. The third entry in the first row shows that 11.1% of the entities that start in state 6 will enter the recurrent classes for the first time at state 3. Since this state is absorbing, the limiting behavior is clear. The entities that enter this state just stay there. A similar analysis can be done for entries 4 and 5 in this row.

In general, we will be able to determine the limiting distributions of the recurrent classes by finding the limit P^n as $n \to \infty$. In this case, the limit matrix is

$$W = \begin{bmatrix} 0.600 & 0.400 & 0 & 0 & 0 & 0 & 0 \\ 0.600 & 0.400 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.00 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.250 & 0.750 & 0 & 0 \\ 0 & 0 & 0 & 0.250 & 0.750 & 0 & 0 \\ 0.270 & 0.180 & 0.110 & 0.110 & 0.330 & 0 & 0 \\ 0.314 & 0.210 & 0.248 & 0.057 & 0.171 & 0 & 0 \end{bmatrix}.$$

Consider the 6-th row of this matrix. The first two entries show that 27% of the entities that start in state 6 ultimately go to state 1, and that 18% ultimately go to state 2. Thus, $27\% + 18\% = 45\%$ of the entities that start in state 6 ultimately move to the first recurrence class. We saw this result earlier using the matrix NR. From the 2×2 matrix in the upper left corner, we see that in the long-run, 60% of the entities in the first recurrence class will be in state 1, and 40% will be in state 2. In particular, the entities that reach this recurrence class from state 6 are proportioned this way: the fraction 0.270 is 60% of the total fraction 0.45, while the fraction 0.180 is 40% of the total fraction 0.45.

A similar analysis applies to the entities that go from state 6 to the third recurrence class. Indeed, the fourth and fifth entries show that 11% of the entities that start in state 6 ultimately go to state 4, and that 33% ultimately go to state 5. Thus, $11\% + 33\% = 44\%$ of the entities that start in state 6, ultimately move to the third recurrence class. From the 2×2 matrix in the middle of W, we see that in the long-run, 25% of the entities in the third recurrence class will be in state 4, and 75% will be in state 5. In particular, the entities that reach this recurrence class from state 6 are proportioned this way: the fraction 0.11 is 25% of the total fraction 0.44, while the fraction 0.33 is 75% of the total fraction 0.44.

So far, we have accounted for $45\% + 44\% = 89\%$ of the entities that start in transient state 6. But entry $(6, 3)$ of W shows that the remaining 11% go to state 3 and are absorbed. Thus, our accounting of the entities that start in transient state 6 is complete. The same analysis can be used to account for the entities that start in transient state 7.