Applied Math 9

Handout 2 for Markov Chains: Some Basic Properties

In class we discussed several examples of discrete state random processes. The *state space* is defined to be the set of values the processes can take, and among the examples discussed in class were

$$S = \{\ldots, -1, 0, 1, \ldots\} \text{ (the integers)},$$

$$S = \left\{\ldots, \frac{1}{c^2}, \frac{1}{c}, 1, c, c^2, \ldots\right\} \text{ (for some } c > 1),$$

$$S = \{a, b, c, \ldots, y, z, A, B, \ldots, Z, ., !, ?, , \}.$$

We will study *finite* state Markov chains, and to simplify notation we will always label the state space as

$$S = \{1, 2, \ldots, M\}.$$

A key item in the description of a Markov chain is the set of *transition probabilities.* These quantities are conditional probabilities, and for a Markov chain X_n (with n denoting the discrete time index) are defined for $i \in S, j \in S$ by

$$p_{ij} = P\{X_{n+1} = j | X_n = i\}.$$

We will also assume that these conditional probabilities do not depend on n, in which case the chain is said to be *stationary.*

The Markov property means that even if you tell me everything that lead up to where the process is at time n, all that matters for the future evolution is the position of the process at time n. Thus

$$\begin{aligned} & P\{X_{n+1} = j | X_n = i, X_0 = s_0, X_1 = s_1, \ldots, X_{n-1} = s_{n-1}\} \\ &= P\{X_{n+1} = j | X_n = i\} \\ &= p_{ij}, \end{aligned}$$

no matter what $s_0, s_1, \ldots, s_{n-1}$ are. It will be convenient to work with the matrix

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1M} \\ \vdots & & \vdots \\ p_{M1} & \cdots & p_{MM} \end{pmatrix}.$$

Thus the starting state corresponds to the row value and the ending state corresponds to the column value, and p_{ij} is the probability that in one step you get to j, given that you start at i. Since they are conditional probablilities, each row i of this matrix must satisfy

$$p_{ij} \geq 0, j \in S,$$

and

$$\sum_{j=1}^{M} p_{ij} = 1.$$

Example. Consider the random walk where the walker is not allowed to go outside at the end points 1 and M. Then

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & \cdots & & & 0 \\ 1/2 & 0 & 1/2 & 0 & \cdots & & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & & 0 & 1/2 & 0 & 1/2 \\ 0 & \cdots & & & 0 & 1/2 & 1/2 \end{pmatrix}.$$

How can we compute the n-step transition probabilities, i.e.,

$$p_{ij}^{(n)} = P\{X_{l+n} = j | X_l = i\}?$$

We will use two elementary properties of probabilities. First, if events E_1 and E_2 are disjoint (i.e., both can never happen simultaneously, and often written $E_1 \cap E_2 = \emptyset$), then

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2)$$

(E_1 or E_2 is often written $E_1 \cap E_2 = \emptyset$). The second fact is that for any events, the definition of conditional probability asserts that $P(E_1 \text{ and } E_2) = P(E_1|E_2)P(E_2)$. Consider $n = 2$. Since the process must be *somewhere* at time $l + 1$,

$$\begin{aligned} p_{ij}^{(2)} &= P\{X_{l+2} = j | X_l = i\} \\ &= \sum_{k=1}^{M} P\{X_{l+2} = j \text{ and } X_{l+1} = k | X_l = i\} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{M} P\{X_{l+2} = j | X_{l+1} = k \text{ and } X_l = i\} \; P\{\, X_{l+1} = k | X_l = i\} \\
&= \sum_{k=1}^{M} P\{X_{l+2} = j | X_{l+1} = k\} \; P\{\, X_{l+1} = k | X_l = i\} \\
&= \sum_{k=1}^{M} p_{kj} p_{ik} \; .
\end{aligned}
$$

¿From the definition of matrix multiplication, we should recognize this as the ijth entry of PP. Thus if $P^{(2)}$ is the matrix of two step transition probabilities, then $P^{(2)} = PP$. We write P^2 for PP. A similar argument shows $P^{(3)} = P^3$ and in general $P^{(n)} = P^n$.

To compute the distribution of the chain at time n, we need to specify its distribution at some earlier time. Typically, this earlier time is 0. Suppose we let

$$v_i = P\{X_0 = i\},$$

and then let $\mathbf{v} = (v_0, v_1, \ldots, v_M)$ (a row vector). We can compute the distribution of the chain at time n:

$$
\begin{aligned}
P\{X_n = j\} &= \sum_{k=1}^{M} P\{X_n = j \text{ and } X_0 = k\} \\
&= \sum_{k=1}^{M} P\{X_n = j | X_0 = k\} \; P\{\, X_0 = k\} \\
&= \sum_{k=1}^{M} \; p_{kj}^{(n)} \; v_k .
\end{aligned}
$$

If we arrange the values $P\{X_n = j\}$ as a row vector, then the last equation tells us that this vector equals

$$\mathbf{v} P^n .$$

Example. Consider the random walk with $M = 5$. Then

$$
P = \begin{pmatrix}
.5 & .5 & 0 & 0 & 0 \\
.5 & 0 & .5 & 0 & 0 \\
0 & .5 & 0 & .5 & 0 \\
0 & 0 & .5 & 0 & .5 \\
0 & 0 & 0 & .5 & .5
\end{pmatrix} .
$$

The 4-step transition matrix is

$$P^4 = \begin{pmatrix} .375 & .25 & .25 & .0625 & .0625 \\ .25 & .375 & .0625 & .25 & .0625 \\ .25 & .0625 & .375 & .0625 & .25 \\ .0625 & .25 & .0625 & .375 & .25 \\ .0625 & .0625 & .25 & .25 & .373 \end{pmatrix}.$$

If the distribution at time 0 is $\mathbf{v} = (0, .3, 0, 0, .7)$, then the distribution at time 4 is $\mathbf{v}P^4$, which equals

$$(0.1187, 0.1563, 0.1937, 0.2500, 0.2812).$$