Markov Chains: lecture 3.

Expected First Return Times for ergodic Markov Chains.

Definition: Let $T_y = \inf\{n \geq 1 \ : \ X_n = y\}$. Then $E(T_y \mid X_0 = y) = m_y$ is called the expected first return time for the state y.

Remark: We will write $E(T_y \mid X_0 = y)$ as $E_y(T_y)$, where $E_y(\dots)$ denotes the expectation with respect to the probabilities obtained when we condition on $X_0 = y$.

Basic Fact: For an ergodic Markov chain with finite state space, we have the mean first return times all finite. The next theorem shows us how to compute the mean first return times in terms of the stationary probability distribution for P.

Theorem: For an ergodic Markov chain, the expected first return time m_x for state x satisfies

$$m_x = 1/w_x,$$

where $w = [w_1, \dots, w_r]$ is the stationary probability vector for P.

"Proof": The weak law of large numbers for ergodic Markov chains says that if

$$H_j^n = \frac{Y_0 + \cdots + Y_n}{n+1} = \text{ average \# of times in state } j \text{ in the first } n \text{ steps,}$$

then

$$\lim_{n \to \infty} P(|H_j^n - w_j| > \epsilon) = 0,$$

for all $j = 1, 2, \dots, r$ and $\epsilon > 0$ provided $w = [w_1, \dots, w_r]$. Since

$$m_j = \text{ average number of steps to return to } j,$$

we have that n/m_j is approximately the average number of visits to j in the first n steps when we start at j. Therefore,

$$\frac{n+1}{m_j} \approx Y_0 + \cdots + Y_n,$$

or

$$\frac{1}{m_j} \approx \frac{Y_0 + \cdots + Y_n}{n+1} \approx w_j, \text{ as } n \to \infty$$

provided we start at state j.

Example: A die is rolled repeatedly. Let's show by using Markov chain theory that the mean time between occurrences of a given number is six. Is this natural? Why not 5?

Solution: What Markov chain do we want? Let X_n for $n \geq 1$ be the value of the nth roll. (This is Markovian, but the states actually move independently.) What is the transition Matrix?

$$P = \begin{bmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{bmatrix}$$

So $w = [1/6, 1/6, \dots, 1/6]$ is the stationary distribution and the previous theorem gives what we want.

Example: A certain experiment is believed to be described by a two state Markov chain with the transition matrix P, where

$$P = \begin{bmatrix} 1/2 & 1/2 \\ p & 1-p \end{bmatrix}$$

and the parameter p is not known. When the experiment is performed many times, the chain ends in state one approximately 20 percent of the time and in state two 80 percent of the time. Lets compute a sensible estimate for the unknown parameter p.

Solution:

By the form given for P we instantly see that the Markov chain is regular because each of the entries of P is strictly positive. Thus, a stationary distribution for the chain exists. Because we are told that "When the experiment is performed many times, the chain ends in state one approximately 20 percent of the time and in state two 80 percent of the time" we know that the stationary distribution is approximately $w = [.2 \quad .8]$. We also know that w is the unique solution that satisfies the equations:

$$w = wP, \qquad w_1 + w_2 = 1.$$

Therefore,

$$[.2 \quad .8] \begin{bmatrix} 1/2 & 1/2 \\ p & 1-p \end{bmatrix} = [.2 \quad .8].$$

Using only the first of the two equations gives us

$$\begin{aligned} & (1/2)(.2) + p(.8) = .2 \\ \Longrightarrow \quad & p = 1/8. \end{aligned}$$