Markov Chains (continued)

Although we have worked quite a bit with the (first) return time T_i we do not yet know its expected value. Indeed, we do not even know whether and when it is finite. It turns out that this expectation relates to the important concept of stationary (or invariant) distribution that we now turn to.

We say that a distribution $\pi = (\pi_i)_{i \in I}$ is stationary if it satisfies π=πT. The first question that we address is whether such an object exist for any transition matrix T (at least for a finite I). On p. 35 of Norris, one finds the following intriguing statement:

"(..) for a finite state-space I, the existence of an invariant row vector is a simple piece of linear algebra: the row sums of T are all 1, so the column vector of ones is an eigenvector with eigenvalue 1, so T must have a row eigenvector with eigenvalue 1." What does that mean and how does that help us?

First, if we let $v = (1, .., 1)^T$ it is clear that $v = Tv$ (as stated above). So, $\lambda = 1$ is an eigenvalue of T (a statement we have made repeatedly without proof in the past). Therefore, $(T - I)$ is a singular matrix with determinant 0 and so must be the transpose $(T - I)^T = (T^T - I)$. So, the transpose T^T also has eigenvalue 1 and must have an eigenvector (column) $w = T^T w$. But if we let $u = w^T$ (a row vector) and transpose the last equation, we have $u = uT$. This is the meaning of T has a "row eigenvector." What is the catch? Well, a stationary distribution is clearly a row eigenvector but the converse is not necessarily true! So, we need a little more than just a simple piece of linear algebra to answer the existence question. There are several ways to go. Here I adopt one that introduces an important theorem of topology.

Fixed Point Theorem (Brouwer): Let K be a convex, closed, and bounded subset of $\mathcal{R}^n$, and let f:$K \to K$ be a continuous map. Then there must exist a (fixed) point $X \in K$ such that $f(X) = X$.

Let us apply this result to our problem: let us denote by K the set of distributions $\{p = (p_i)_{i \in I} : p_i \geq 0, \sum_{i \in I} p_i = 1\}$. This set K is clearly convex, closed and bounded (it is often called a simplex). Now consider the map f defined by $f(p) = pT$ (with T our transition matrix). This f is clearly continuous and maps K into itself. So there must be a fixed point $\pi = \pi T$ which is our stationary distribution.

Since Brouwer's theorem says nothing about uniqueness, we cannot say anything more at this point. Indeed, it is a difficult question. Moreover, this says nothing about the infinite (countable) case.

To make the desired connection with the expected return time $E(T_i)$, we need some further concepts. Consider some state $i \in I$ and define $\gamma_j^{(i)}$ as the expected number of turns (time) spent in j between consecutive visits to i. Formally

$$\gamma_j^{(i)} = E\left(\sum_{n=0}^{T_i - 1} 1_{\{X_n = j\}}\right)$$

Clearly with this definition $\gamma_i^{(i)} = 1$. Moreover $\gamma_j^{(i)} = E\left(\sum_{n=1}^{T_i} 1_{\{X_n = j\}}\right)$. Consider this last expression and write[1]

$$\begin{aligned}
\gamma_j^{(i)} &= E\sum_{n=1}^{T_i}(1_{\{X_n=j\}}) = E\sum_{n=1}^{\infty}(1_{\{X_n=j,n\le T_i\}}) = \sum_{n=1}^{\infty}P(X_n = j, n \le T_i) \\
&= \sum_{n=1}^{\infty}\sum_{k\in I}P(X_{n-1} = k, X_n = j, n \le T_i) \\
&= \sum_{n=1}^{\infty}\sum_{k\in I}p_{kj}P(X_{n-1} = k, n \le T_i) = \sum_{k\in I}p_{kj}\sum_{n=1}^{\infty}P(X_{n-1} = k, n \le T_i) \\
&= \sum_{k\in I}p_{kj}\sum_{m=0}^{\infty}P(X_m = k, m \le T_i - 1) = \sum_{k\in I}p_{kj}E\sum_{m=0}^{T_i-1}(1_{\{X_m=j\}}) = \sum_{k\in I}p_{kj}\gamma_k^{(i)}
\end{aligned}$$

So, if we let $\gamma^{(i)} = (\gamma_j^{(i)})_{j\in I}$, this vector satisfies $\gamma^{(i)} = \gamma^{(i)}T$. In fact, this is really useful when all these expectations are finite. So, let us make the assumption that the Markov chain is recurrent and irreducible (is a single communicating class). In that case, one easily proves (see homework) that all $\gamma_j^{(i)}$ are positive and finite. Thus, we may consider

$$M_i = \sum_{j\in I}\gamma_j^{(i)}$$

It is not difficult to see that M_i is in fact our desired expected return time to i (see homework). In the case of a finite state-space I, M_i is clearly finite. With an infinite I there are two cases: (1) M_i is finite and the state i is called *positive recurrent*; (2) M_i is infinite and i is called *null recurrent*. We are only concerned from now on with the positive recurrent case (which happens when I is finite).

We may now divide $\gamma^{(i)}$ by M_i to obtain a distribution that is clearly invariant. Indeed, it is possible to prove that, in the positive recurrent case, there is only one such invariant distribution. So, it has to be the same as π. Interestingly, since $\gamma_i^{(i)} = 1$, we find that $\frac{1}{M_i} = \pi_i$. So, the expected return time to i in a positive recurrent case is nothing but one over the i component of the (unique) stationary distribution.

---

[1]Everywhere here the probabilities are conditioned into $X_0 = i$.

One may naturally ask how to obtain π and therefore the expected return times. In practice, one has to solve the system of linear equations $\pi = \pi T$. In a lot of (positive recurrent) cases, there is more than just an existence and uniquenes result on π. It is also often a limit that is approached from any initial distribution. In other words one finds that $\lim_{n \to \infty} pT^n = \pi$. However, there are exceptions (called periodic cases). Here are two examples of positive recurrent Markov chains.

Let $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. One finds easily that $T^{2n} = I$ and $T^{2n+1} = T$. So, there can never be any convergence.

Let instead $T = \begin{pmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{pmatrix}$. One finds that $T^n \to \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$ the matrix whose rows are the stationary distribution. A sufficient condition for this to hold is when T is *positive* (that is all its entries are strictly positive).

I finally state without proof an *Ergodic Theorem* which can be viewed as an application of the famous *Strong Law of Large Numbers*.

Ergodic Theorem: Assume that T is irreducible and positive recurrent. Then

1) if $V_i(n) = \sum_{k=0}^{n-1} 1_{\{X_k = i\}}$ is the number of visits to i before time n, then

$$P\left(\frac{V_i(n)}{n} \to \pi_i \text{ as } n \to \infty\right) = 1$$

2) for any bounded function $f: I \to \mathcal{R}$

$$P\left(\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \to \sum_{i \in I} \pi_i f(i)\right) = 1$$

Homework

1. Prove that if T is recurrent and irreducible, then all $\gamma_j^{(i)}$ are positive and finite. Hint: for any $n > 0$ $\gamma^{(i)} = \gamma^{(i)} T$. Therefore, there exists n such that $\gamma_j^{(i)} = \gamma_i^{(i)} p_{ij}^{(n)}$.

2. Show that M_i is the expected return time. Hint: write $\sum_{j \in I} E\left(\sum_{n=1}^{\infty} 1_{\{X_n = j, n \le T_i\}}\right)$ as $\sum_{n=1}^{\infty} n P(X_0 = i, X_1 \neq i, .., X_{n-1} \neq i, X_n = i)$.