Markov Chains (continued)

One important issue arises when in the method discussed in the last class when the eigenvalues of T include complex numbers. How do we interpret the effect of such eigenvalues on probabilities that are real by definition? Let us discuss the following example

$$T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

The eigenvalues are the solutions of $\mathrm{p}(\mu) = -\mu(\mu - \frac{1}{4}) + \frac{1}{4} = 0$, or $\mu = 1, \pm\frac{i}{2}$. Suppose we are interested in the probability of being at state 1, starting from state 1, after n turns. In the notations of Norris, this reads $\mathrm{p}_{11}^{(n)}$. Using the same argument as in the last class, this must read

$$\mathrm{p}_{11}^{(n)} = \mathrm{a} + \mathrm{b}\left(\frac{i}{2}\right)^n + \mathrm{c}\left(-\frac{i}{2}\right)^n \tag{1}$$

However, by Euler's formula, $\left(\pm\frac{i}{2}\right)^n = \left(\frac{1}{2}\right)^n\left(\cos\frac{n\pi}{2} \pm i\sin\frac{n\pi}{2}\right)$. Combining with (1) yields

$$\mathrm{p}_{11}^{(n)} = \mathrm{a} + \left(\frac{1}{2}\right)^n\left\{(\mathrm{b}+\mathrm{c})\cos\frac{n\pi}{2} + i(\mathrm{b}-\mathrm{c})\sin\frac{n\pi}{2}\right\} \tag{2}$$

So, if b and c are complex conjugates, we can write β=(b + c) and $\gamma = i$(b − c), both real, and the expression in (2) is real. The same technique as before applies. We write

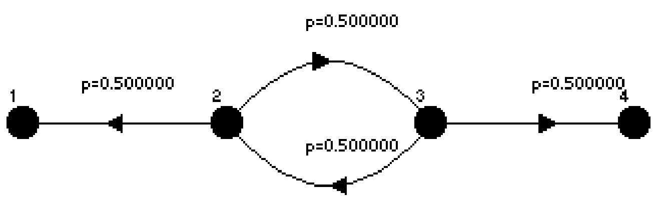
$\mathrm{p}_{11}^{(0)} = \mathrm{a} + \beta = 1$ (since we start at state 1)
$\mathrm{p}_{11}^{(1)} = \mathrm{a} + \frac{1}{2}\gamma = 0$ (since the $\mathrm{p}_{11} = 0$)
$\mathrm{p}_{11}^{(2)} = \mathrm{a} - \frac{1}{4}\beta = 0$ (by calculating λT^2)

It follows that $\mathrm{a} = \frac{1}{5}$, $\beta = \frac{4}{5}$, and $\gamma = -\frac{2}{5}$. In particular, the long run probability of being at state 1 is $\frac{1}{5}$.

In general, we can calculate the probability $\mathrm{p}_{ij}^{(n)}$ of reaching state j from state i after n turns in the very same way. We can then ask whether it is *possible* to reach state j from state i which, of course, requires that $\mathrm{p}_{ij}^{(n)} > 0$ for some n. We will say that i *leads to* j (denoted $i \to j$) whenever that probability is positive for some n. We similarly say that i communicates with j (denoted $i \leftrightarrow j$) when both $i \to j$ and $j \to i$. This is clearly an equivalence relation that defines so-called *communicating* classes. Such a class is *closed* if whenever $i \to j$ and i is in the class then j is in the class (in other words, you cannot leave a closed class). A state i that forms a closed class is called *absorbing*. If a Markov chain is such that I is a single class, then it is called *irreducible*.

We can also define the *hitting time* H_i^j as the minimum time n it takes for X_n to be in state j, starting from $X_0 = i$ (with the convention that this is ∞ if j can never be reached from i). This is clearly a random variable. We can similarly define H_i^A for a subset $A \subseteq I$. We may now define the probability $h_i^A = P(H_i^A < \infty)$ that, starting from state i, $(X_n)_{n \geq 0}$ ever hits A. When A is a closed class, this will be called its *absorbsion* probability.

With the example of last time, it is clear that the whole set $I = \{1,2,3\}$ forms an irreducible chain since all states communicate with each other. Therefore, the hitting probabilities h_i^j are all trivially 1. A less trivial example is given by the following diagram (from Norris p. 12)



Clearly, the probability, starting from state 1, that the chain ever hits anything but state 1 is nil. But what is, for instance, the probability that starting from 2, the chain hits (and is absorbed into) 4? h_i^4 denotes the probability that, starting from i, the chain hits 4. Clearly, $h_1^4 = 0$ and $h_4^4 = 1$. Moreover:

$$h_2^4 = \tfrac{1}{2}h_1^4 + \tfrac{1}{2}h_3^4 \qquad \text{and} \qquad h_3^4 = \tfrac{1}{2}h_2^4 + \tfrac{1}{2}h_4^4 \tag{3}$$

Solving (3) yields $h_2^4 = \frac{1}{3}$.

A related question is *how long* it takes to reach 4 starting from 2. In other words, what is the *mean time* in which one reaches j from i? This is simply the expectation of the random variable H_i^j that reads (and may be infinite)

$$k_i^j = E(H_i^j) = \sum_{n<\infty} nP(H_i^j = n) \; + \infty P(H_i^j = \infty) \tag{4}$$

With the same example, we find $k_1^4 = \infty$. Moreover $k_2^4 = 1 + \frac{1}{2}k_1^4 + \frac{1}{2}k_3^4 = \infty$. So, the expected time it takes to reach 4 from 1 is infinite because of the possibility to be absorbed in 1 from which it become impossible to reach 4. If, however, we ask the mean time of reaching $A = \{1{,}4\}$ from 2 we have $k_1^A = k_4^A = 1$ and

$$k_2^A = 1 + \tfrac{1}{2}k_1^A + \tfrac{1}{2}k_3^A$$
$$k_3^A = 1 + \tfrac{1}{2}k_2^A + \tfrac{1}{2}k_4^A$$

which yields $k_2^A = 2$.

More generally, we can obtain the following

Theorem 1: The vector of hitting probabilities $h^A = (h_i^A : i \in A)$ is the minimal non-negative solution of the system of linear equations

$$h_i^A = 1 \qquad \text{for } i \in A$$
$$h_i^A = \sum_{j \in I} \mathrm{p}_{ij} h_j^A \text{ for } i \notin A$$

Proof: Norris, pp. 13-14.

Theorem 2: The vector of mean hitting times $k^A = (k_i^A : i \in I)$ is the minimal non-negative solution of the system of linear equations

$$k_i^A = 0 \qquad \text{for } i \in A$$
$$k_i^A = 1 + \sum_{j \notin A} \mathrm{p}_{ij} k_j^A \qquad \text{for } i \notin A$$

Proof: Norris, pp. 17-18.

Homework

1. Problem 1.3.2 p. 18 of Norris.