11.3 Some Mathematical Conditions

There is a rich mathematical theory regarding the asymptotic behavior of Markov chains. We will not present such material here in this brief introduction, but we will give an example of the implications the theory can have.

(Note: Due to the large number of definitions and properties in this section, they are highlighted via bulleting.)

- A state in a Markov chain is called **recurrent** if it has the property that if we start at that state, we are guaranteed to return to the state.

- A nonrecurrent state is called **transient**.

Let $T_{ii}$ denote the time needed to return to state $i$ if we start there. Keep in mind that $T_{ii}$ is the time from one entry to state $i$ to the next entry to state $i$. So, it includes time spent in $i$, which is 1 unit of time for a discrete-time chain and a random exponential amount of time in the continuous-time case, and then time spent away from $i$, up to the time of next entry to $i$. Note that an equivalent definition of recurrence is that $P(T_{ii} < \infty) = 1$, i.e. we are sure to return to $i$. By the Markov property, if we are sure to return once, then we are sure to return again once after that, and so on, so this implies infinitely many visits.

- A recurrent state $i$ is called **positive recurrent** if $E(T_{ii}) < \infty$.

- A state which is recurrent but not positive recurrent is called **null recurrent**.

- In the discrete time case, a state $i$ is recurrent if and only if

\[
\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty \tag{11.50}
\]

where $p_{uv}^{(k)}$ means the row $u$, column $v$ element of the $k^{th}$ power of the transition matrix of the chain, i.e. the probability of being at state $v$ after $n$ steps if one starts at $u$.

This last can be easily seen in the “only if” case: Let $A_n$ denote the indicator random variable for the event $T_{ii} = n$ (Section 3.8). Then

\[
p_{ii}^{(n)} = E A_n, \tag{11.51}
\]

---

8Some random variables are finite with probability 1 but have infinite expected value. For instance, suppose $M = 1, 2, 3, \ldots$ with probabilities $1/2, 1/4, 1/8\ldots$ Those probabilities sum to 1, so $P(M < \infty) = 1$ yet $EM = \infty$. 

---
so the left-hand side of (11.50) is

\[ E \left( \sum_{n=1}^{\infty} A_n \right) \quad (11.52) \]

i.e. the expected value of the total number of visits to state i. If state i is recurrent, then we will visit i infinitely often, and thus that sum should be equal to infinity.

- A chain is **irreducible** if we can get from any state to any other state (though not necessarily in one step).
- One can show that in an irreducible chain, if one state is recurrent then they all are. The same statement holds if “recurrent” is replaced by “positive recurrent.”

Again, this last bullet should make intuitive sense to you for the recurrent case: We make infinitely many visits to state i, and each time we have a nonzero probability of going to state j from there. Thus eventually we will visit j, and indeed each time we visit i we will subsequently visit j at some time. Thus we will make infinitely many visits to j as well, i.e. j is recurrent.

### 11.3.1 Example: Random Walks

Consider the famous **random walk** on the full set of integers: At each time step, one goes left one integer or right one integer (e.g. to +3 or +5 from +4), with probability 1/2 each. In other words, we flip a coin and go left for heads, right for tails.

If we start at 0, then we return to 0 when we have accumulated an equal number of heads and tails. So for even-numbered n, i.e. \( n = 2m \), we have

\[ p_{ii}^{(n)} = P( \text{m heads and m tails} ) = \binom{2m}{m} \frac{1}{2^{2m}} \quad (11.53) \]

One can use Stirling’s approximation,

\[ m! \approx \sqrt{2\pi e^{-m}} m^{m+1/2} \quad (11.54) \]

to show that the series (11.50) diverges in this case. So, this chain (meaning all states in the chain) is recurrent. However, it turns out not to be not positive recurrent, as we’ll see below.

---

9Some readers may see the similarity to the notion of a connected graph.
The same is true for the corresponding random walk on the two-dimensional integer lattice (moving up, down, left or right with probability 1/4 each). However, in the three-dimensional and higher cases, the chain is not even null recurrent; it is transient.

This last point is rather startling to most people. What is so special about dimensions 3 and higher? UCD grad student Jeremy Bottleson has given a very insightful explanation as follows:

In the one-dimensional case, say we started at position 0. Will we return? Wherever we are at some time, we have a 50% chance of getting closer to 0 on our next step, and 50% of going further away. What about two dimensions, starting at (0,0)? Say we are now at (8,8). We have a 25% chance of next going to (7,8), thus closer to (0,0), and the same holds for (8,7). In the other two cases, we go further away from (0,0).

But things change in three dimensions. Say we start at (0.0.0), and are currently at (8,8,8). Now our probability of moving closer to (0,0,0) in our next step is only \(3 \cdot \frac{1}{8} = \frac{3}{8}\), while we have a \(\frac{5}{8}\) chance of moving further away. Intuitively, then, we have a built-in tendency to drift further away from (0,0,0) over time, hence transience.

### 11.3.2 Finding Hitting and Recurrence Times

We will use the symbol \(T_{ij}\) to denote the time it takes to get to state \(j\) if we are now in \(i\), known as the hitting time, analogous to the recurrence times \(T_{ii}\). Note that this is measured from the time we enter state \(i\) to the time we enter state \(j\).

### 11.3.3 Recurrence Times and Stationary Probabilities

Consider a continuous-time chain. For a positive recurrent state \(i\), it turns out that

\[
\pi_i = \frac{1/\lambda_i}{E(T_{ii})}
\]

(11.55)

To see this, we take an approach similar to that of Section [11.1.5.1]. Define alternating On and Off subcycles, where On means we are at state \(i\) and Off means we are elsewhere. Define a full cycle to consist of an On subcycle followed by an Off subcycle. Note again that \(T_{ii}\) is measured from the time we enter state \(i\) once until the time we enter it again.

Then intuitively the long-run proportion of time we are in state \(i\) is

\[
\pi_i = \frac{E(\text{On})}{E(T_{ii})}
\]

(11.56)
But an On subcycle has an exponential distribution, with mean duration \( 1/\lambda_i \). That gives us (11.55).

In the discrete-time case, an On subcycle always has duration 1. Thus we have

\[
\pi_i = \frac{1}{E(T_{ii})}
\]  

(11.57)

Thus positive recurrence means that \( \pi_i > 0 \). For a null recurrent chain, the limits in Equation (11.3) are 0, which means that there may be rather little one can say of interest regarding the long-run behavior of the chain.

Nevertheless, note that (11.57) makes sense even in the null recurrent case. The right-hand side will be 0, and if one views \( \pi_i \) in terms of (11.3), that means that in the long-run we don’t spend a positive proportion of time at this state.

Equation (11.57) also shows that random walk in one or two dimensions, though recurrent, is null recurrent. If they were positive recurrent, then by symmetry all the \( \pi_i \) would have to be equal, and yet they would have to sum to 1, a contradiction. So, \( ET_{ii} = \infty \) for all i.

11.3.3.1 Hitting Times

We are often interested in finding quantities of the form \( E(T_{ij}) \). We can do so by setting up systems of equations similar to the balance equations used for finding stationary distributions.

First consider the discrete case. Conditioning on the first step we take after being at state i, and using the Law of Total Expectation, we have

\[
E(T_{ij}) = \sum_k p_{ik} E(T_{ij} | W = k)
\]  

(11.58)

where W is the first state we go to after leaving state i. There are two cases for \( E(T_{ij} | W = k) \):

- **W = j:**
  
  This is the trivial case; \( E(T_{ij} | W = k) = 1 \).

- **W ≠ j:**
  
  If our first step is to some state k other than j, it takes us 1 unit of time to get to k, at which point “time starts over” (Markov property), and our expected remaining time is \( ET_{kj} \). So, in this case \( E(T_{ij} | W = k) = 1 + ET_{kj} \).
11.3. SOME MATHEMATICAL CONDITIONS

So, using the Law of Total Expectation, we have

\[
E(T_{ij}) = \sum_k p_{ik} E(T_{ij} | W = k) = \sum_{k \neq j} p_{ik} [1 + E(T_{kj})] + p_{ij} \cdot 1 = 1 + \sum_{k \neq j} p_{ik} E(T_{kj}) \tag{11.59}
\]

By varying \(i\) and \(j\) in (11.59), we get a system of linear equations which we can solve to find the \(ET_{ij}\). Note that (11.57) gives us equations we can use here too.

The continuous version uses the same reasoning:

\[
E(T_{ij}) = \sum_{k \neq j} p_{ik} \left[ \frac{1}{\lambda_i} + E(T_{kj}) \right] + p_{ij} \cdot \frac{1}{\lambda_i} = \frac{1}{\lambda_i} + \sum_{k \neq j} p_{ik} E(T_{kj}) \tag{11.60}
\]

One can use a similar analysis to determine the probability of ever reaching a state, in chains in which this probability is not 1. (Some chains have have transient or even absorbing states, i.e. states \(u\) such that \(p_{uv} = 0\) whenever \(v \neq u\).)

For fixed \(j\) define

\[
\alpha_{ij} = P(T_{ij} < \infty) \tag{11.61}
\]

Then denoting by \(S\) the state we next visit after \(i\), we have

\[
\alpha_{ij} = P(T_{ij} < \infty) = \sum_k P(S = k \text{ and } T_{ij} < \infty) \tag{11.62}
\]

\[
= \sum_{k \neq j} P(S = k \text{ and } T_{kj} < \infty) + P(S = j) \tag{11.63}
\]

\[
= \sum_{k \neq j} P(S = k) \, P(T_{kj} < \infty | S = k) + P(S = j) \tag{11.64}
\]

\[
= \sum_{k \neq j} p_{ik} \alpha_{kj} + p_{ij} \tag{11.65}
\]

So, again we have a system of linear equations that we can solve for the \(\alpha_{ij}\).