B.Sc./Grad. Dip.: Probability Models and Time Series
MAS programmes: Stochastic Models and Time Series

2 Markov Chains: Part II

2.1 Classification of States

Consider the question of how likely it is that a Markov Chain will ever re-visit a particular state given that it is known when it was last in that state. We address this question by classifying the states according to those which are *recurrent* and those which are *transient*.

Definition 2.1.1 (recurrent and transient states)
State i is *recurrent* if

$$\mathbb{P}(X_n = i \text{ for some } n \geq 1 | X_0 = i) = 1,$$

otherwise i is said to be *transient*.

Thus, if i is recurrent, then the probability of eventual return to i, having started from i is 1. A goal of this section will be to characterize recurrence and transience in terms of the $\{p_{ij}(n)\}$. To this end we introduce the *first passage times and probabilities.*

Definition 2.1.2 (first passage probability)
For $n \geq 1$, let

$$f_{ij}(n) = \mathbb{P}(X_1 \neq j, X_2 \neq j, \ldots, \ldots, X_{n-1} \neq j, X_n = j | X_0 = i)$$

and

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n).$$

An interpretation of these quantities would be that $f_{ij}(n)$ represents the probability that the first visit to state j, starting from i, takes place at the n-th step, and that f_{ij} represents the probability that the chain ever hits state j, starting from i.

The obvious conclusion is that

$$\text{state } j \text{ is recurrent} \iff f_{jj} = 1.$$

We can express this result in terms of the n-step transition probabilities, namely the $\{p_{ij}(n)\}$.

Proposition 2.1.3 (necessary and sufficient condition for recurrence)
state j is recurrent $\iff \sum_{n=1}^{\infty} p_{jj}(n) = \infty$.

Proposition 2.1.4 (other properties of n-step transition probabilities)
(a) If state j is recurrent & state i is such that $f_{ij} > 0$, then $\sum_n p_{ij}(n) = \infty$.

(b) If state j is transient then:
(i) $\sum_n p_{ij}(n) < \infty$ for all i;
(ii) $p_{ij}(n) \to 0$ as $n \to \infty$.

Remarks 2.1.5

• Let $N(i)$ be the number of times the chain visits starting point i. Then if the state i is recurrent then $\mathbb{P}(N(i) = \infty) = 1$; however, if it is transient, then $\mathbb{P}(N(i) = \infty) = 0$.

• Just because a state is recurrent, this does not imply that the *expected time* between subsequent visits to the state is finite! And so we have to make a further distinction between states which are *positive recurrent* and those which are *null recurrent* (to be defined shortly).

Definition 2.1.6 (First visit time)
Let
$$T_j = \min\{n \geq 1 : X_n = j\}$$
be the time of the first visit to state j, and set $T_j = \infty$ if the set $\{n \geq 1 : X_n = j\}$ is empty.

So note that state i is transient

$$\iff \mathbb{P}(T_i = \infty | X_0 = i) > 0 \Rightarrow E[T_i | X_0 = i] = \infty.$$

Definition 2.1.7 (mean recurrence time)
The *mean recurrence time* of state i is defined
$$\mu_i = E[T_i | X_0 = i] = \begin{cases} \sum_n n f_{ii}(n) & \text{if state } i \text{ is recurrent} \\ \infty & \text{if state } i \text{ is transient} \end{cases}.$$

One can see that even if state i is recurrent (thus returning to state i an infinite number of times, we could still have either $\mu_i < \infty$ or $\mu_i = \infty$.

Definition 2.1.8 (null and positive recurrence)
A recurrent state i is *null recurrent* if $\mu_i = \infty$ and *positive recurrent* if $\mu_i < \infty$.

The following theorem provides a test for determining whether a recurrent state is null or not.

Theorem 2.1.9 (test for nullity of a recurrent state)
A recurrent state, state i say, is null recurrent

$$\iff p_{ii}(n) \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Also, if state i is null recurrent then

$$p_{ji}(n) \longrightarrow 0 \text{ as } n \longrightarrow \infty \text{ for all } j.$$

It is sometimes also important to monitor the feasible opportunities for which a chain could visit a particular state.

Definition 2.1.10 (periodicity)
The *period*, $d(i)$, of state i is given by

$$d(i) = \gcd\{n : p_{ii}(n) > 0\}$$

which is the greatest common divisor of the epochs at which return to state i is possible.
If $d(i) > 1$, then state i is *periodic*.
On the other hand, if $d(i) = 1$, then state i is *aperiodic*.

As an example, for the symmetric random walk that we considered previously, each state is periodic, with $d(i) = 2$ for all $i \in \mathbb{S}$. It is only possible to return to a state after an even number of steps.

Positive recurrent states that are also aperiodic can be analyzed in interesting ways. This property is called *ergodicity*.

Definition 2.1.11 (ergodicity)
A state is *ergodic* if it is positive recurrent and aperiodic.

Example 2.1.12 (symmetric random walk)
To recall, state space $\mathbb{S} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ and

$$p_{ij} = \begin{cases} p & \text{if} \quad j = i+1 \\ 1-p & \text{if} \quad j = i-1 \\ 0 & \quad\text{otherwise} \end{cases}$$

where $0 < p < 1$.

State conditions on the parameter p for which state 0 is:
(i) recurrent; (ii) transient.
Is state 0 ergodic?

Solution:
State 0 (as do all the states) have period 2 (as noted earlier). Therefore, state 0 is **not** ergodic.
As a matter of fact

$$p_{00}(2n+1) = 0, \qquad n = 0, 1, 2, \ldots$$

and

$$p_{00}(2n) = \binom{2n}{n} p^n (1-p)^n = \frac{(2n)!}{n!n!} p^n (1-p)^n, \qquad n = 1, 2, \ldots$$

To answer parts (i) and (ii), let us consider *Sterling's formula* (which is presented here without proof) which says that for large n

$$n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}.$$

Plugging this formula into the expression for $p_{00}(2n)$, we find that for large n,

$$p_{00}(2n) \sim \frac{(4p(1-p))^n}{\sqrt{n\pi}}.$$

• Looking at this expression, note that $0 < p(1-p) < 1$ and that $p(1-p)$ has a maximum value of $\frac{1}{4}$ attained at $p = \frac{1}{2}$.

• Hence, at $p = \frac{1}{2}$,

$$p_{00}(2n) \sim \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\pi}}$$

and since $\sum_{m=1}^{\infty} \frac{1}{\sqrt{m}} = \infty$, it follows that $\sum_{n=0}^{\infty} p_{00}(n) = \infty$. Thus the chain is recurrent.

• If $p \neq \frac{1}{2}$, then

$$p_{00}(2n) \sim \frac{x^n}{\sqrt{\pi n}} \quad \text{where} \quad 0 < x < 1.$$

But

$$\sum_{n=1}^{\infty} \frac{x^n}{\sqrt{n}} < \infty \quad \text{i.e. convergent}^1.$$

In this case, state 0 (as for each other state within the chain) is transient.

Remarks 2.1.13 (symmetric random walks in higher dimensions)
• *2-dimensional random walk*: for any point on the lattice, the probabilities of moving $(1,0)$, $(-1,0)$, $(0,1)$ and $(0,-1)$ relative to the current position at the next step are each equal to $\frac{1}{4}$. In this case, all states are recurrent.

• *3-dimensional random walk*: again for any point on the lattice, the probabilities of moving $(1,0,0)$, $(-1,0,0)$, $(0,1,0)$, $(0,-1,0)$, $(0,0,-1)$ and $(0,0,1)$ are each equal to $\frac{1}{6}$. In this case, all states are transient!!! The states are similarly transient for all other n-dimensional analogues for $n \geq 4$.

2.2 Classification of chains

Let us first find a way of characterizing the "probabilistic interconnectedness" of the states within a chain.

Definition 2.2.1 (communicating states)
• i communicates with j, written as $i \to j$, if $p_{ij}(m) > 0$ for some $m \geq 0$.
• i and j intercommunicate if $i \to j$ and $j \to i$. Write as $i \leftrightarrow j$.

It can be shown, by way of an alternative characterization, that for $i \neq j$, $i \to j \iff f_{ij} > 0$.

The next theorem gives us a clue as to how we might partition and classify all the states within a chain without having to check every single state in exactly the same detail.

[1] By the "ratio test" - look it up on Wikipedia - $a_n = \frac{x^n}{\sqrt{n}}$, so $\frac{a_{n+1}}{a_n} = \frac{x^{n+1}}{\sqrt{n+1}} \times \frac{\sqrt{n}}{x^n} = x\sqrt{\frac{n}{n+1}} \longrightarrow x$ as $n \longrightarrow \infty$ s.t. $0 < x < 1$.

Theorem 2.2.2 (properties of intercommunicating states)
If $i \leftrightarrow j$, then
(a) i and j have the same period;
(b) i is recurrent if and only if j is recurrent;
(c) i is null recurrent if and only if j is null recurrent.

Proof [Part (b) only]
Since $i \leftrightarrow j$, there exist $m, n \geq 0$ such that $p_{ji}(m) > 0$ and $p_{ij}(n) > 0$. By the Chapman-Kolmogorov equations,

$$p_{jj}(m+n+s) = \sum_k p_{jk}(m+s)p_{kj}(n) = \sum_k \left\{ \sum_h p_{jh}(m)p_{hk}(s) \right\} p_{kj}(n)$$
$$\geq \sum_k p_{jk}(m)p_{kk}(s)p_{kj}(n) \geq p_{ji}(m)p_{ii}(s)p_{ij}(n).$$

Summing over s:

$$\sum_{s=1}^{\infty} p_{jj}(s) \geq \sum_{s=1}^{\infty} p_{jj}(m+n+s) \geq p_{ji}(m)p_{ij}(n) \sum_{s=1}^{\infty} p_{ii}(s).$$

But since i is recurrent then, from Proposition 2.1.3, $\sum_{s=1}^{\infty} p_{ii}(s) = \infty$, which, by the above inequality, implies that $\sum_{s=1}^{\infty} p_{jj}(s) = \infty$ also. Hence, again, by Proposition 2.1.3, state j is recurrent. We can also adapt the proof in the other direction to show that if j is recurrent then i is also recurrent.

Definition 2.2.3 (closedness and irreducibility)
A set $\mathbb{C}$ of states is called
(a) *closed* if $p_{ij} = 0$ for $i \in \mathbb{C},\ j \notin \mathbb{C}$.
(b) *irreducible* if $i \leftrightarrow j$ for all $i, j \in \mathbb{C}$.

Remarks 2.2.4
• If the chain enters a closed set of states, then it will never leave that set. If a closed set contains only one state, then that state is called *absorbing*.
• An entire irreducible set can be classified according to periodicity, recurrence, transience etc. according to Theorem 2.2.2.
• If the entire state space $\mathbb{S}$ is irreducible, then the chain itself is said to be irreducible: thus, in such cases, one speaks of an *irreducible Markov chain*.

We can partition the state space according to the following theorem:

Theorem 2.2.5 (decomposition of the state space)
The state space $\mathbb{S}$ can be partitioned uniquely as

$$\mathbb{S} = \mathbb{T} \cup \mathbb{C}_1 \cup \mathbb{C}_2 \cup \ldots$$

where $\mathbb{T}$ is the set of transient states and the $\mathbb{C}_i$, $i = 1, 2, \ldots$ are irreducible closed sets of recurrent states.

Let us reflect a little on the implications of this theorem. If at stage 0, the chain resides in the set of transient states $\mathbb{T}$, then either it stays in $\mathbb{T}$ for ever (which could happen if this set is countably infinite) or it eventually gets sucked into precisely one of the closed sets, $\mathbb{C}_i$, where it resides forever never to leave. On the other hand, if the chain resides in one of the closed irreducible sets at time 0, then the chain will never venture out of that set, and so that set will become, in effect, the state space of the chain.

We can refine these results even further for chains with a finite state space.

Lemma 2.2.6 (decomposition for finite chains) $\hookrightarrow \mathbb{S} = \mathbb{T} + \mathbb{C}_1 + \mathbb{C}_2 + \ldots$

Suppose that $\mathbb{S}$ is finite. Then:

(a) at least one state is recurrent;

(b) all recurrent states are positive recurrent.

Example 2.2.7 (Markov chain with finite state space)

Let $\mathbb{S} = \{a, b, c, d, e, f\}$ and

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Classify the states of this chain.

<u>Solution</u>:

$\{a, b\}$ and $\{e, f\}$ constitute irreducible closed sets, both of which contain only positive recurrent states. States c and d are transient because $c \to d \to f$, however there is no route from e or f back to either c or d. There is positive probability of being in the current state at the next stage in the chain, i.e. $p_{ii} > 0$ for all i: therefore the period of all states, $d(i)$, is equal to 1.

It is also the case that states $\{a, b, e, f\}$ are all ergodic (see Definition 2.1.11).

Definition 2.2.8 (stationary distribution)

The vector $\boldsymbol{\pi}$ is called a *stationary distribution* of the chain $\mathbf{X}$ if the entries of $\boldsymbol{\pi}$, $\{\pi_j : j \in \mathbb{S}\}$, are such that

(I) $\pi_j \geq 0$ for all $j \in \mathbb{S}$ and $\sum_{j\in\mathbb{S}} \pi_j = 1$;

(II) $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$ i.e. $\pi_j = \sum_{i\in\mathbb{S}} \pi_i p_{ij}$ for all $j \in \mathbb{S}$.

Thus if the chain behaves in accordance with its stationary distribution at a particular stage, then it will behave in accordance to that distribution for all later stages. To see this note that if $\boldsymbol{\pi}$ is the stationary distribution of the chain, then

$$\boldsymbol{\pi}\mathbf{P}^2 = (\boldsymbol{\pi}\mathbf{P})\mathbf{P} = \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}.$$

And so more generally

$$\boldsymbol{\pi}\mathbf{P}^n = (\boldsymbol{\pi}\mathbf{P})\mathbf{P}^{n-1} = \boldsymbol{\pi}\mathbf{P}^{n-1} = (\boldsymbol{\pi}\mathbf{P})\mathbf{P}^{n-2} = \boldsymbol{\pi}\mathbf{P}^{n-2} = \ldots = \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}.$$

Working with Lemma 1.2.9, it is clear that if the mass function of X_0 is $\boldsymbol{\pi}$, then, for any $n \geq 1$, the mass function of X_n is also $\boldsymbol{\pi}$.

But are there any circumstances under which we can be sure that a stationary distribution actually exists (and that it is perhaps unique) prior to carrying out any calculations? It turns out that an answer to this question is available for chains that are irreducible i.e. ones for which all the states intercommunicate.

Theorem 2.2.9 (existence and uniqueness of a stationary distribution)
Suppose that $\mathbf{X}$ is an irreducible Markov chain on $\mathbb{S}$. Also let μ_i be the *mean recurrence time* of state $i \in \mathbb{S}$, i.e. the expected number of steps between successive visits to the state.

Then $\mathbf{X}$ has a stationary distribution $\boldsymbol{\pi}$ if and only if all the states in $\mathbb{S}$ are positive recurrent.
Further, if $\mathbf{X}$ does have a stationary distribution, $\boldsymbol{\pi}$ say, then it is unique and its components are given by

$$\pi_i = \frac{1}{\mu_i} \qquad i \in \mathbb{S}.$$

All states in the sample space S must be positive recurrent for the MC $\underset{\sim}{X}$ to have a unique stationary distribution $\underset{\sim}{\pi}$.

Example 2.2.10 (finite chain revisited)
Recall that

$$\mathbb{C}_1 = \{a, b\} \text{ and } \mathbb{C}_2 = \{e, f\}$$

constitute the closed, irreducible, sets.

Now consider $X_0 \in \mathbb{C}_1$. Then $X_n \in \mathbb{C}_1$ for all n. Also, the states of $\mathbb{C}_1$ are positive recurrent. The state space is in effect $\mathbb{C}_1$ and the associated transition matrix is

$$\mathbf{P}_{\mathbb{C}_1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}.$$

Thus the unique stationary distribution is found by solving

$$(\pi_a, \pi_b) = (\pi_a, \pi_b) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} \quad \text{which implies } (\pi_a, \pi_b) \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{4} \end{pmatrix} = (0, 0).$$

Thus

$$\begin{aligned} \frac{1}{2}\pi_a - \frac{1}{4}\pi_b &= 0 \\ -\frac{1}{2}\pi_a + \frac{1}{4}\pi_b &= 0 \end{aligned} \quad .$$

Either of the two equations yields that $\pi_b = 2\pi_a$, and so we conclude that $(\pi_a, \pi_b) = A \times (1, 2)$ where A is a constant. But since $\pi_a + \pi_b = 1$, then $A = \frac{1}{3}$: thus $(\pi_a, \pi_b) = (\frac{1}{3}, \frac{2}{3})$.

So we know that if a chain starts off in its stationary distribution at a particular point in time, then it will maintain that distribution at all later stages. In other words, $\mathbf{X}$ will have already attained its *limiting distribution* from the very beginning. But what can be said about the limiting behaviour of a chain which does not start off in its stationary distribution? What conditions are needed to make any kind of meaningful statement?

First note that periodicity (i.e. $d(i) > 1$) can certainly be a reason for a limiting distribution not to exist as the next example shows.

Example 2.2.11 (no limiting distribution for periodic chain)
Suppose $\mathbb{S} = \{v, w\}$ and $p_{vw} = p_{wv} = 1$. Then clearly

$$p_{vv}(n) = p_{ww}(n) = \begin{cases} 0 & \text{if} \quad n \text{ odd} \\ 1 & \text{if} \quad n \text{ even} \end{cases}.$$

Thus, for $i \in \{v, w\}$, $p_{ii}(n)$ does **not** converge as $n \longrightarrow \infty$. So we ought to restrict attention to aperiodic chains (where $d(i) = 1$ for all $i \in \mathbb{S}$).

Theorem 2.2.12 (limiting distribution)
For an irreducible aperiodic chain,

$$p_{ij}(n) \longrightarrow \frac{1}{\mu_j} \quad \text{as} \quad n \longrightarrow \infty \;\text{ for all } i, j.$$

Let us interpret and consider some of the implications of this very important theorem.

Remarks 2.2.13
(a) If the chain is transient or null recurrent, then $\mu_j = \infty$ for all $j \in \mathbb{S}$, in which case $p_{ij}(n) \longrightarrow 0$ as $n \longrightarrow \infty$.

(b) If the chain is positive recurrent, then

$$p_{ij}(n) \longrightarrow \pi_j = \frac{1}{\mu_j}$$

where $\boldsymbol{\pi} = \{\pi_j : j \in \mathbb{S}\}$ is the unique stationary distribution.

(c) Not only does the chain forget its starting point $X_0 = i$ (since $\frac{1}{\mu_j}$ does not involve i), but it also forgets the distribution of X_0 (if indeed X_0 was permitted to be random at stage 0). To see this, note that

$$\mathbb{P}(X_n = j) = \sum_i \mathbb{P}(X_0 = i) p_{ij}(n) \longrightarrow \frac{1}{\mu_j} \sum_i \mathbb{P}(X_0 = i) = \frac{1}{\mu_j} \;\text{ as }\; n \longrightarrow \infty.$$

(d) If $\mathbf{X} = \{X_n\}$ is an irreducible chain but with period d, then one can extract a subsequence of $\mathbf{X}$ which is both irreducible and aperiodic by setting $\mathbf{Y} = \{Y_n = X_{nd} : n \geq 0\}$. It can be shown that

$$p_{jj}(nd) = \mathbb{P}(Y_n = j | Y_0 = j) \longrightarrow \frac{d}{\mu_j} \;\text{ as }\; n \longrightarrow \infty.$$