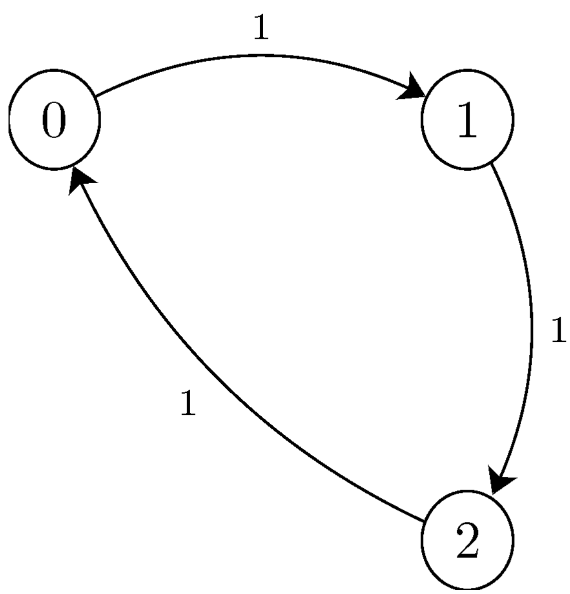
Week 3

Markov chains: Periodicity, recurrence and transience, positive and null recurrences

7 Markov chain: periodicity

Consider the Markov chain shown below:



There is a periodic pattern in this chain. Indeed, we have

$$
\begin{aligned}
p_{00}^{(n)} &= P(X_n = 0 \mid X_0 = 0) \neq 0\ (=1), \quad \text{if } n = 3, 6, 9, \cdots, \\
p_{00}^{(n)} &= P(X_n = 0 \mid X_0 = 0) = 0, \quad \text{if } n \neq 3, 6, 9, \cdots.
\end{aligned}
$$

Such a state is called a **periodic** state with period $d_0 = 3$.

The **period** of a state i, denoted by d_i, is defined as the greatest common divisor (gcd) of all $n \geq 1$ for which $p_{ii}^{(n)} > 0$, i.e.,

$$d_i = \gcd \{n : \; p_{ii}^{(n)} > 0\}.$$

If $p_{ii}^{(n)} = 0$, for all $n \geq 1$, then we let $d_i = 0$ (or $d_i = \infty$).

- If $d_i > 1$, we say that state i is **periodic**.
  In this case, $p_{ii}^{(kd_i)} \neq 0$ for all $k \geq k_0$, where $k_0 > 1$ is an integer, and $p_{ii}^{(n)} = 0$ when $n \neq kd_i$. Note that it happens $p_{ii}^{(d_i)} = 0$.
- If $d_i = 1$, we say that state i is **aperiodic**.

In mathematics, the greatest common divisor (gcd) of two or more integers, which are not all zero, is the largest positive integer that divides each of the integers. For example, the gcd of 8 and 12 is 4.

Th3.0. If $i \longleftrightarrow j$, then $d_i = d_j$.

Proof. Suppose $p_{ij}^{(m)} > 0$ and $p_{ji}^{(n)} > 0$. Then, by the C-K equation,

$$p_{ii}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{ki}^{(n)} > 0.$$

It follows that $m + n = kd_i$ for some $k \geq 1$. Similarly, supposing $p_{jj}^{(l)} > 0$, we have

$$\begin{aligned} p_{ii}^{(m+n+l)} &= \sum_{k \in S} p_{ik}^{(m)} p_{ki}^{(n+l)} \\ &\geq p_{ij}^{(m)} p_{jj}^{(l)} p_{ji}^{(n)} > 0. \end{aligned}$$

Hence $m + n + l = k_1 d_i$ for some $k_1 > k$. As a consequence, we have

$$l = (k_1 - k) d_i = c_1 d_i$$

where $c_1 \geq 1$ is an integer. Recall that $d_j = gcd\{l : p_{jj}^{(l)} > 0\}$, we have $d_i \leq d_j$. The same argument shows that $d_j \leq d_i$, i.e., $d_i = d_j$.

A class is said to be **periodic** if its states are periodic. Similarly, a **class** is said to be aperiodic if its states are aperiodic. Finally, a **Markov chain** is said to be **aperiodic** if all of its states are aperiodic.

Why is periodicity important?

As we will see later, it plays a role when we discuss limiting distributions. It turns out that in a typical problem, we are given an irreducible Markov chain, and we need to check if it is aperiodic.

How do we check that a Markov chain is aperiodic?

Consider a finite irreducible Markov chain X_n, i.e., the MC only has a one class with finite states:

- If there is a self-transition in the chain ($p_{ii} > 0$ for some i), then the chain is aperiodic.
- Suppose that you can go from state i to state i in l steps, i.e., $p_{ii}^{(l)} > 0$. Also suppose that $p_{ii}^{(m)} > 0$. If $\gcd(l, m) = 1$, then chain is aperiodic.

  Note: two numbers m and l are said to be **co-prime** if their greatest common divisor (gcd) is 1, i.e., $\gcd(l, m) = 1$.

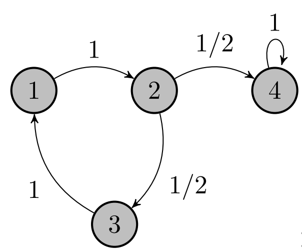
- The chain is aperiodic if and only if there exists a positive integer n such that all elements of the matrix P^n are strictly positive, i.e.,

$$p_{ij}^{(n)} > 0, \quad \text{for all } i, j \in S.$$

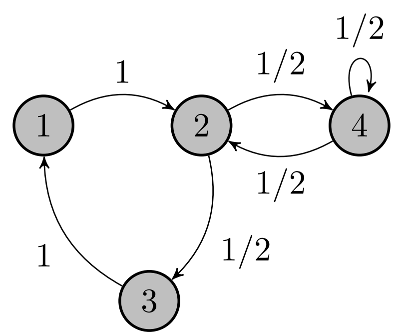
Example 3.1: Consider a MC with transition matrix:

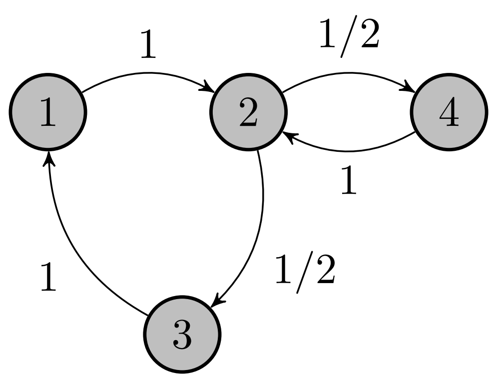
$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{bmatrix}; \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

The state diagrams:



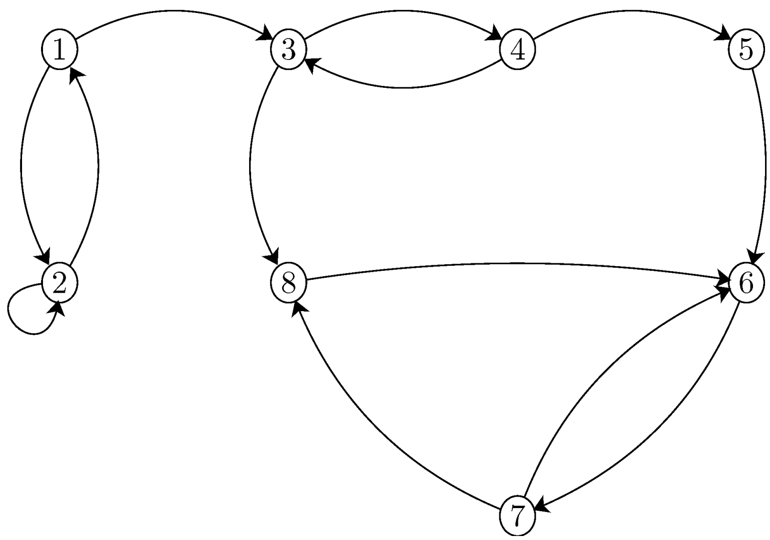
;





8 Recurrence and transience

Consider MC with the state transition diagram:



Here are the classes:

$$\begin{aligned}\text{Class } 1 &= \{1, 2\},\\ \text{Class } 2 &= \{3, 4\},\\ \text{Class } 3 &= \{5\},\\ \text{Class } 4 &= \{6, 7, 8\}.\end{aligned}$$

There are two kinds of classes:

- if MC enters Class 4, it will always stay in that class;
- not true for other classes.

The states in Class 4 are called **recurrent** states, while the other states in this chain are called **transient**.

In general, a state $i \in S$ is said to be **recurrent** if

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

i.e. a state is said to be **recurrent** if, any time that we leave that state, we will return to that state in the future with probability one.

A non-recurrent state is called **transient**.

Let f_{ij} denote the probability that the chain eventually arrives at state j from i, i.e.,

$$f_{ij} = P(X_n = j \text{ for some } n \geq 1 \mid X_0 = i),$$

A state i is recurrent if $f_{ii} = 1$; is transient if $f_{ii} < 1$.

Write

$$\begin{aligned} f_{ij}^{(1)} &= P(X_1 = j \mid X_0 = i), \\ f_{ij}^{(n)} &= P(X_n = j, X_{n-1} \neq j, ..., X_1 \neq j \mid X_0 = i), \quad \text{for } n \geq 2. \end{aligned}$$

Then $f_{ij}^{(n)}$ is the probability that, starting at i, the first visit to j takes place at the n-th transition. Note that

$$\big(X_n = j \text{ for some } n \geq 1\big) = \cup_{n=1}^{\infty}(X_n = j) = \cup_{n=1}^{\infty} A_n(j),$$

where $A_n(j) = (X_n = j, X_{n-1} \neq j, ..., X_1 \neq j)$ and $A_n(j)$ are disjoint. We have

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

We have the following identity:

$$p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)}. \tag{1}$$

Compare with the C-K equation:

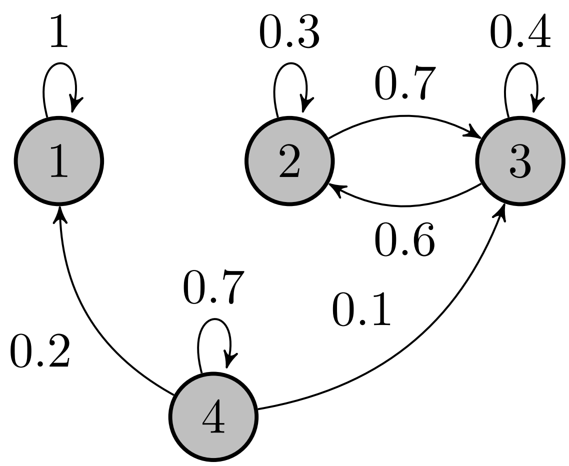
$$p_{ij}^{(n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n-m)}, \quad \text{for any } 1 \le m \le n.$$

Example 3.3: Consider a MC with transition matrix: for $0 < p < 1$

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.3 & 0.7 & 0 \\ 0 & 0.6 & 0.4 & 0 \\ 0.2 & 0 & 0.1 & 0.7 \end{bmatrix}$$

We may show that

$$f_{11} = 1, \quad f_{22} = 1, \quad f_{33} = 1, \quad f_{44} < 1.$$



Th3.1. A state i is recurrent (transient) if and only if $f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$ (< 1) or $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$ $(< \infty)$, where we recall

$$p_{ii}^{(n)} = P(X_n = i \mid X_0 = i).$$

Note that the f_{ii} is the probability that, starting at i, the state i is eventually re-entered; the $\sum_{n=1}^{\infty} p_{ii}^{(n)}$ is the expected (total) number of visits to the state i.

Indeed, by letting $N_i = \sum_{n=1}^{\infty} I_{(X_n = i)}$, then

$$E(N_i \mid X_0 = i) = \sum_{n=1}^{\infty} E\big(I_{(X_n = i)} \mid X_0 = i\big) = \sum_{n=1}^{\infty} p_{ii}^{(n)}.$$

Proof. difficult, omitted.

Th3.2. If state i is recurrent and $i \longleftrightarrow j$, then state j is recurrent.

This theorem shows that recurrence is a class property. Thus, we can extend the above definitions to classes.

A **class** is said to be **recurrent** if the states in that class are recurrent. If, on the other hand, the states are transient, the **class** is called transient. In general, a Markov chain might consist of several transient classes as well as several recurrent classes.

Proof. Suppose $p_{ij}^{(m)} > 0$ and $p_{ji}^{(n)} > 0$. Then, by the C-K equation,

$$p_{ii}^{(m+n+r)} \geq p_{ij}^{(m)} p_{jj}^{(r)} p_{ji}^{(n)}$$

So, if $\sum_{r=1}^{\infty} p_{jj}^{(r)} = \infty$, then $\sum_{r=1}^{\infty} p_{ii}^{(r)} = \infty$.

Th3.3. In a finite Markov chain, there is at least one recurrent class.

Proof. Consider a finite Markov chain with r states, $S = \{1, 2, \cdots, r\}$. Suppose that all states are transient. Then, starting from time 0, the chain might visit state 1 several times, but at some point the chain will leave state 1 and will never return to it. That is, there exists an integer $M_1 > 0$ such that $X_n \neq 1$, for all $n \geq M_1$. Similarly, there exists an integer $M_2 > 0$ such that $X_n \neq 2$, for all $n \geq M_2$, and so on. Now, if you choose

$$n \geq \max\{M_1, M_2, \cdots, M_r\},$$

then X_n cannot be equal to any of the states $1, 2, \cdots, r$. This is a contradiction, so we conclude that there must be at least one recurrent state, which means that there must be at least one recurrent class.

Let N_i be the number of times that the MC visits state i, i.e., $N_i = \sum_{n=1}^{\infty} I_{(X_n=i)}$. We have

Th3.4. If i is a transient state, then

$$E(N_i \mid X_0 = i) = f_{ii}/(1 - f_{ii}).$$

Proof. difficult, omitted.

9 Null recurrent

Recall the definition of $f_{ij}^{(n)}$:

$$\begin{aligned} f_{ij}^{(1)} &= P(X_1 = j \mid X_0 = i), \\ f_{ij}^{(n)} &= P(X_n = j, X_{n-1} \neq j, ..., X_1 \neq j \mid X_0 = i), \quad \text{for } n \geq 2. \end{aligned}$$

A recurrent state i is said to be

1. **positive recurrent** if $\sum_{n=1}^{\infty} n\, f_{ii}^{(n)} < \infty$;
2. **null recurrent** if $\sum_{n=1}^{\infty} n\, f_{ii}^{(n)} = \infty$;
3. **ergodic** if i is positive recurrent and aperiodic.

Write $T_j = \min\{n \geq 1 : X_n = j\}$; $T_j := \infty$ if for all $n \geq 1$, $X_n \neq j$.

T_j is the first visiting time to state j for the MC $\{X_n\}_{n \geq 0}$.

$\mu_i = E(T_i \mid X_0 = i)$ is the mean recurrence time of state i (the expected number of transitions needed to return to state i).

For a transient state i, $\mu_i = \infty$ as $P(T_i = \infty \mid X_0 = i) > 0$.

When i is a recurrent state, we have $f_{ii}^{(n)} = P(T_i = n \mid X_0 = i), n \geq 1$ so that $\mu_i = \sum_{n=1}^{\infty} n\, f_{ii}^{(n)}$. Hence, $\mu_i = \sum_{n=1}^{\infty} n\, f_{ii}^{(n)} = \infty$ for null recurrent, and $\mu_i < \infty$ for positive recurrent.

Compare with

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \sum_{n=1}^{\infty} P(X_n = i \mid X_0 = i),$$

the expected (total) number of visits to the state i.

We also have the following results:

Th3.5. (i) If $i \longleftrightarrow j$, then both i and j are transient or null recurrent or positive recurrent;

(ii) If $C_j = \{i : i \longleftrightarrow j\}$ and j is a recurrent, then C_j is a closed class.

Proof. (i) We make use of the following result: A state k is null recurrent iff $\sum_{n=1}^{\infty} p_{kk}^{(n)} = \infty$ and $p_{kk}^{(n)} \to 0$, as $n \to \infty$ (see next week notes).

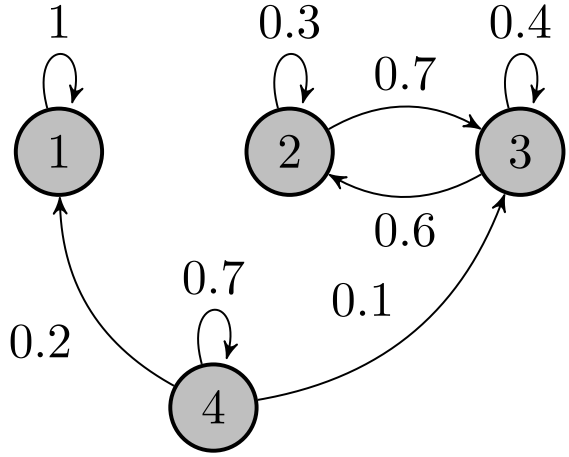
Suppose $p_{ij}^{(m)} > 0$ and $p_{ji}^{(n)} > 0$. By the C-K equation, we have

$$p_{ii}^{(m+n+r)} \geq p_{ij}^{(m)} p_{jj}^{(r)} p_{ji}^{(n)},$$

i.e, if $p_{ii}^{(n)} \to 0$, then $p_{jj}^{(n)} \to 0$. So, if both i and j are recurrent and i is null recurrent, then $p_{jj}^{(n)} \to 0$ and $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$, i.e., j is also null recurrent. Now (i) comes from Th3.2.

For (ii), we have to prove if $i \in C_j$ and $k \notin C_j$, then $p_{ik} = 0$. In fact, if $p_{ik} > 0$, then the MC has a positive probability (at least p_{ik}) so that it will never return to state i, due to the fact that $p_{kk}^{(n)} = 0$ for all $n \geq 0$ as $k \notin C_j$. This means the state i is not a recurrent, which is a contradiction.

An example:



Th3.6. If S is a finite state space of a MC, then at least one state is recurrent, and all recurrent states are positive recurrent.

Proof. By Th3.3, there is a state (j, say) to be recurrent. If j is null recurrent, then $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ and $p_{jj}^{(n)} \to 0$, as $n \to \infty$ (see next week notes), indicating that

$$p_{ij}^{(n)} = \sum_{k=1}^{n} f_{ij}^{(k)} p_{jj}^{(n-k)} \to 0, \quad \text{for any } i \in S,$$

since $\sum_{k=1}^{\infty} f_{ij}^{(k)} \leq 1$. But, for each $n \geq 0$,

$$\sum_{j \in S} p_{ij}^{(n)} = P(X_n \in S \mid X_0 = i) = 1.$$

It is impossible if S is finite, implying that j can only be positive recurrent.

Th3.7. Let C_0 be a closed class with finite states. Then all states in C_0 are positive recurrent.

Proof. Since C_0 is closed, for all $n \geq 1$, we have

$$p_{jk}^{(n)} = 0, \quad j \in C_0 \quad \text{and} \quad k \notin C_0.$$

This implies that, for all $n \geq 1$ and $j \in C_0$,

$$\sum_{k \in C_0} p_{jk}^{(n)} = 1.$$

Therefore, $\sum_{n=1}^{\infty} \sum_{k \in C_0} p_{jk}^{(n)} = \infty$ for $j \in C_0$. As a consequence, there exists $k_0 \in C_0$ so that $\sum_{n=1}^{\infty} p_{k_0 k_0}^{(n)} = \infty$, i.e., $k_0 \in C_0$ is a recurrent. Now the claim comes from Th3.5.

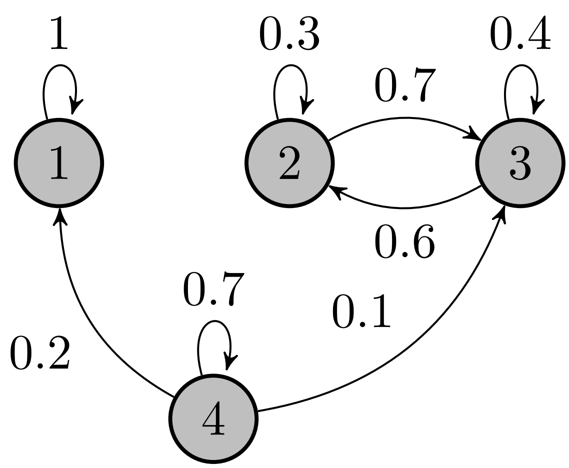
In summary:

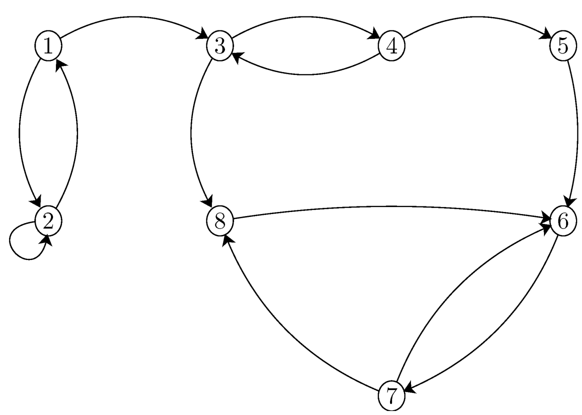
Let S be a state space of a MC $\{X_n\}_{n\geq 0}$.

S can be decomposed into a number of classes (easy by using the state diagram):

$$S = \big(T_1 \cup \cdots \cup T_r\big) \cup \big(C_1 \cup \cdots \cup C_m\big)$$

where T_j are disjoint not closed classes and C_k are disjoint closed classes. We can say C_k are positive recurrent classes and T_j are transient classes.





For an irreducible MC, it follows that

- all states are transient; or
- all states are positive recurrent; or
- all states are null recurrent.

Note: if S has infinite states, having no recurrent states in the chain is possible.

Example 3.4. MC with transition matrix:

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & ... \\ 0 & 0 & 1 & 0 & 0 & ... \\ 0 & 0 & 0 & 1 & 0 & ... \\ ... & & & & & \\ ... & & & & & \end{bmatrix}$$

Note: there is no relationship between aperiodic and recurrent (transient).