CE 273
Markov Decision Processes

Lecture 2

Transient Behavior and Classification of States

Previously on Markov Decision Processes

Definition (Markov Property)

A stochastic process $\{X_n, n \geq 0\}$ with a countable state space S is called a DTMC if $\forall\, n \geq 0, i, j \in S$,

$$\mathbb{P}[X_{n+1} = j | X_n = i, X_{n-1}, X_{n-1}, \ldots, X_0] = \mathbb{P}[X_{n+1} = j | X_n = i]$$

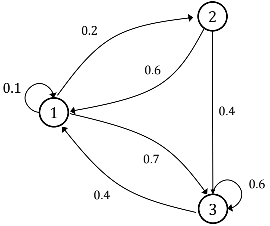
The probability with which the system moves from i to j, p_{ij}, is called the **transition probability** and the matrix of p_{ij} values is called the **one-step transition probability matrix**.

$$P = \left[p_{ij}\right]_{|S|\times|S|}$$

Previously on Markov Decision Processes

The transition probability matrix can also be visualized as a directed graph in which the states are nodes and an arc (i, j) exists only if $p_{ij} > 0$.

$$P = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.6 & 0 & 0.4 \\ 0.4 & 0 & 0.6 \end{bmatrix} \end{array}$$



The P matrix alone doesn't fully describe a DTMC. We'd also need to know the initial distribution.

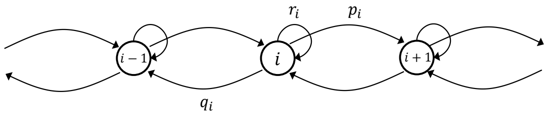
$$a_i = \mathbb{P}\left[X_0 = i\right] \forall\, i \in S$$

Let a be row vector of a_i's. A Markov chain can thus be fully specified using (S, P, a).

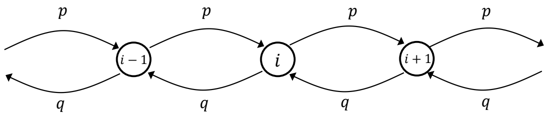
Previously on Markov Decision Processes

Most common variant of the random walk allows steps of size $0, 1$, and -1 on a 1D lattice.

State-Dependent Random Walk:



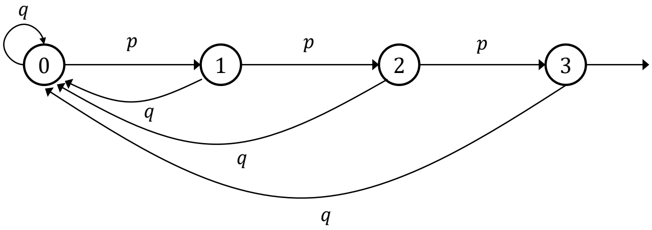
Simple Random Walk:



Previously on Markov Decision Processes

Suppose a coin is tossed repeatedly and the probability of seeing heads and tails is p and q. Assume a player wins ₹1 every time it is heads and looses his entire winnings if it is tails.

Let X_n represent the player's cash after n tosses. The DTMC can be represented using the following transition diagram.



Lecture Outline

Transient Behavior

Transient Behavior

Introduction

Given a Markov chain, we'll address the following two questions that reflect the transient behavior:

- What is the **marginal distribution** of X_n? We can subsequently address questions on the expected values etc.
- What is the expected time spent in various states up to time n. (Also called **occupancy times**.)

Transient Behavior

Marginal Distribution

Let $\{X_n, n \geq 0\}$ be a DTMC on $S = \mathbb{Z}^+$ with transition matrix P and initial distribution a. For a given n, the marginal distribution of X_n is

$$
\begin{aligned}
a_j^{(n)} &= \mathbb{P}\big[X_n = j\big] \, \forall j \in S \\
&= \sum_{i \in S} \mathbb{P}\big[X_n = j | X_0 = i\big] \mathbb{P}\big[X_0 = i\big] \text{ (Law of Total Probability)} \\
&= \sum_{i \in S} a_i p_{ij}^{(n)}
\end{aligned}
$$

where $p_{ij}^{(n)}$ is the probability of going from i to j in exactly n steps. Define the n-step transition matrix $P^{(n)}$ as

$$P^{(n)} = \big[p_{ij}^{(n)}\big]_{|S| \times |S|}$$

Hence, to compute the marginal distributions, we need to know the n-step transition matrices.

Transient Behavior

Marginal Distribution

Intuitively, to go from i to j in n steps, we need to transition from i to some state r in k steps and from r to j in remaining $(n-k)$ steps.

Theorem (Chapman-Kolmogorov Equations)

The n-step transition probabilities satisfy

$$p_{ij}^{(n)} = \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}, \forall\, i, j \in S, 0 \leq k \leq n$$

Proof.

$$\begin{aligned}
p_{ij}^{(n)} &= \mathbb{P}\left[X_n = j | X_0 = i\right] \\
&= \sum_{r \in S} \mathbb{P}\left[X_n = j, X_k = r | X_0 = i\right] \\
&= \sum_{r \in S} \mathbb{P}\left[X_n = j | X_k = r, X_0 = i\right] \mathbb{P}\left[X_k = r | X_0 = i\right] (\text{LoToP})
\end{aligned}$$

Proof.

$$= \sum_{r \in S} \mathbb{P}\big[X_n = j | X_k = r\big] \mathbb{P}\big[X_k = r | X_0 = i\big] \text{(Markov Property)}$$

$$= \sum_{r \in S} \mathbb{P}\big[X_{n-k} = j | X_0 = r\big] \mathbb{P}\big[X_k = r | X_0 = i\big] \text{(Time Homogeneity)}$$

$$= \sum_{r \in S} p_{ir}^{(k)} p_{rj}^{(n-k)}$$

∎

The CK equations can be compactly written as

$$P^{(n)} = P^{(k)} P^{(n-k)}$$

Applying the CK equations recursively,

Corollary

$$P^{(n)} = P^n$$

Thus, the probability mass function of X_n can be written as

$$a^{(n)} = aP^n$$

For example, consider the two-state DTMC with transition matrix

$$P = \begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix} \end{array} \qquad P^4 = \begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} 0.675 & 0.325 \\ 0.650 & 0.350 \end{bmatrix} \end{array}$$

Say $a = [0.5\ 0.5]$, then the pmf of X_4 is $[0.662\ 0.338]$

Transient Behavior

Computing Matrix Powers

The problem of finding the marginal distributions, thus simplifies to computing P^n. Regular multiplication often leads to numerical instability or is intractable.

Let's look at two cases:

- Finite state spaces
- Countably infinite state spaces

For finite state problems, we could use some tricks if we are lucky. Consider a square matrix A. Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of A (note that they satisfy $\det(\lambda I - A) = 0$). Suppose the right eigenvector of λ_j is x_j.

$$Ax_j = \lambda_j x_j$$

Define a matrix X whose columns are the right eigenvectors associated with $\lambda_1, \ldots, \lambda_m$.

$$X = \begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ x_1 & x_2 & \ldots & x_m \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix}$$

Then,

$$\begin{aligned} AX &= [Ax_1 \ Ax_2 \ \ldots \ Ax_m] \\ &= [\lambda_1 x_1 \ \lambda_1 x_2 \ \ldots \ \lambda_m x_m] = XD \end{aligned}$$

where $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$.

Suppose X is invertible. (A sufficient condition for this is that all eigenvalues are distinct.) In this case,

$$AX = XD$$
$$A = XDX^{-1}$$

If the above condition holds, we say that A is **diagonalizable**. In fact, we can also show that

$$X^{-1} = \begin{bmatrix} \longleftarrow & y_1 & \longrightarrow \\ \vdots & \vdots & \vdots \\ \longleftarrow & y_m & \longrightarrow \end{bmatrix}$$

where y_j is the left eigenvector of λ_j (i.e., $y_j A = \lambda_j y_j$). Computing powers of diagonalizable matrices is easy!

$$A^n = (XDX^{-1})(XDX^{-1}) \cdots (XDX^{-1}) = XD^n X^{-1}$$

Caution: Not all transition matrices are diagonalizable. For example,

$$\begin{bmatrix} 5/12 & 5/12 & 1/6 \\ 1/4 & 1/4 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

The method discussed above only works if they are diagonalizable.

Diagonalizable matrices offer more insight into the limiting behavior.

Theorem (Perron–Frobenius)

Let P be a transition matrix with m eigenvalues $\lambda_1, \ldots, \lambda_m$. Then

1. *At least one of them is 1. (Why?)*
2. $|\lambda_i| \leq 1 \, \forall \, i \in \{1, 2, \ldots, m\}$

Recall that $a^{(n)} = aP^n$. When the eigenvalues are unique, the set of eigenvectors form a basis, and thus we can write

$$a = c_1 y_1 + c_2 y_2 + \ldots + c_m y_m$$

Transient Behavior

Computing Matrix Powers

Post-multiplying by P,

$$aP = c_1 y_1 P + c_2 y_2 P + \ldots + c_m y_m P$$
$$aP = c_1 \lambda_1 y_1 + c_2 \lambda_2 y_2 + \ldots + c_m \lambda_m y_m$$

Post-multiplying by P repeatedly,

$$aP^2 = c_1 \lambda_1 y_1 P + c_2 \lambda_2 y_2 P + \ldots + c_m \lambda_m y_m P$$
$$aP^2 = c_1 \lambda_1^2 y_1 + c_2 \lambda_2^2 y_2 + \ldots + c_m \lambda_m^2 y_m$$
$$\vdots$$
$$aP^n = c_1 \lambda_1^n y_1 + c_2 \lambda_2^n y_2 + \ldots + c_m \lambda_m^n y_m$$

Using the Perron-Frobenius theorem, the terms containing λ's which have magnitudes strictly less than 1 will vanish as $n \to \infty$.

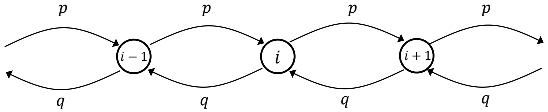
The rate of convergence will thus depend on the magnitude of the second largest eigenvalue. Smaller the better!

Transient Behavior

Computing Matrix Powers

Diagonalization-based methods cannot be applied to DTMCs with countably infinite states. One can use the problem structure to derive analytical expressions in most cases.

For instance, consider the simple random walk



What is the probability of returning to 0 starting from 0 in n steps $p_{00}^{(n)}$ if

- n is odd
- n is even

Generalize this result and write an expression for $p_{ij}^{(n)}$.

We can show that

$$p_{ij}^{(n)} = \begin{cases} \binom{n}{b} p^a q^b & \text{if } n+j-i \text{ is even} \\ 0 & \text{otherwise} \end{cases}$$

where $a = (n+j-i)/2$ and $b = (n+i-j)/2$.

Transient Behavior

Occupancy Times

We might sometimes be interested in the expected amount of time spent by the system in different states up to time n (e.g., parking).

Such metrics are called occupancy times. Let $V_j^{(n)}$ be the number of visits to j over $\{0, 1, \ldots, n\}$. Mathematically, occupancy time of j up to time n starting from i is

$$m_{ij}^{(n)} = \mathbb{E}\left[V_j^{(n)} | X_0 = i\right], \forall\, i, j \in S, n \geq 0$$

The matrix of $m_{ij}^{(n)}$ values, also called the occupancy time matrix, is represented by

$$M^{(n)} = \left[m_{ij}^{(n)}\right]_{|S| \times |S|}$$

The occupancy times matrix can be computed from the transition matrix!

Transient Behavior

Occupancy Times

Theorem

Let $P^0 = I$. *For a fixed* n, $M^{(n)} = \sum_{r=0}^{n} P^r$

Proof.

Fix a $j \in S$. Define a random variable Z_r which is 1 if $X_r = j$ and is 0 otherwise. Then, $V_j^{(n)} = \sum_{r=0}^{n} Z_r$.

$$
\begin{aligned}
m_{ij}^{(n)} &= \mathbb{E}\big[V_j^{(n)} | X_0 = i\big] \\
&= \mathbb{E}\Big[\sum_{r=0}^{n} Z_r | X_0 = i\Big] \\
&= \sum_{r=0}^{n} \mathbb{E}\big[Z_r | X_0 = i\big] \\
&= \sum_{r=0}^{n} \mathbb{P}\big[X_r = j | X_0 = i\big] = \sum_{r=0}^{n} p_{ij}^{(r)}
\end{aligned}
$$

■

Transient Behavior

Occupancy Times

We'll now extend this analysis to instances in which n is very large. Specifically, we will look at the following limits

$$\lim_{n\to\infty} p_{ij}^{(n)}$$

$$\lim_{n\to\infty} m_{ij}^{(n)}$$

and derive conditions under which they exist and find methods to compute them.

But first, we will classify the Markov chain into different classes, which will help us in studying these limits.

Lecture Outline

Classification of States

Definition (Accessibility)

A state j is said to be accessible from i if $\exists\, n \geq 0$ such that $p_{ij}^{(n)} > 0$ and we write $i \to j$

Definition (Communicating)

States i and j are said to be communicating (denoted by $i \leftrightarrow j$) if $i \to j$ and $j \to i$.

It is easy to verify these properties by checking for directed paths in the transition diagram.

Proposition

- *Reflexivity:* $i \leftrightarrow i$
- *Symmetry:* $i \leftrightarrow j \Leftrightarrow j \leftrightarrow$
- *Transitivity:* $i \leftrightarrow j, j \leftrightarrow k \Rightarrow i \leftrightarrow k$

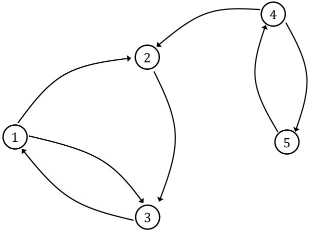
Classification of States

Communicating Classes

Definition (Communicating Class)

A set $C \subseteq S$ is a communicating class if

1. $i \in C, j \in C \Rightarrow i \leftrightarrow j$ (any two states in C must communicate)
2. $i \in C, i \leftrightarrow j \Rightarrow j \in C$ (C is maximal)



- $\{1, 2, 3\}$ and $\{4, 5\}$ are communicating classes.
- $\{1, 3\}$ is not.

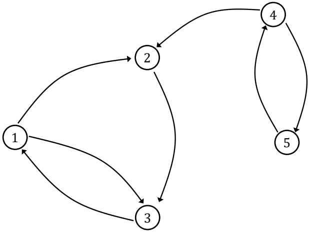
Classification of States

Communicating Classes

Definition (Closed Communicating Class)

A communicating class is said to be closed if $i \in C$ and $j \notin C$, then $i \nrightarrow j$.

Think of closed communicating classes as a black hole. Once the DTMC enters it, it can never leave.



$\{1, 2, 3\}$ and $\{4, 5\}$ are communicating classes but only $\{1, 2, 3\}$ is closed.

Classification of States

Irreducibility

The state space S of a DTMC can be partitioned as

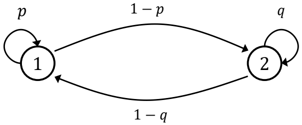
$$S = C_1 \cup C_2 \cup \ldots \cup C_k \cup \mathcal{C}$$

where $C_1, \ldots, C_k$ is the set of closed communicating classes. (k in the above equation can potentially be ∞.) $\mathcal{C}$ is assumed to be the set of states belonging to communicating classes that are not closed.

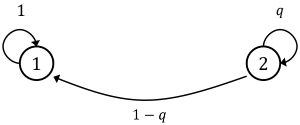
Definition (Irreducibility)

A DTMC is said to be irreducible if its state space S is a closed communicating class. Else it is called reducible.

Classification of States



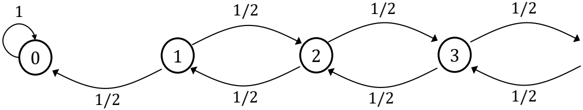
Irreducible DTMC



Reducible DTMC

For the figure on the left $S = C_1 = \{1, 2\}$. For that on the right, $S = C_1 \cup \mathcal{C} = \{1\} \cup \{2\}$.

Is this DTMC irreducible?



Note that all properties discussed so far depend only on the signs of the transition probabilities and not on the magnitudes.

Classification of States

Recap

- Accessibility
- Communicating
- Communicating Class
- Closed Communicating Class
- Irreducibility

We'll now look at other ways to classify states that **may** depend on the magnitudes of the transition probabilities.

Classification of States

Recurrence, Transience, and Periodicity

To study this new classification, let's first define a random variable called the **passage time**

$$\tilde{T}_i = \min\{n > 0 : X_n = i\}$$

It represents the time step when the process visits i for the first time (ignoring the initial state).What is the support of $\tilde{T}_i$?

Given a random variable, we are typically interested in its pmf and expected value. In the context of DTMCs, the following functions are of interest.

1. **Probability that the return time is finite**

$$\tilde{u}_i = \mathbb{P}\big[\tilde{T}_i < \infty | X_0 = i\big]$$

2. **Expected return time**

$$\tilde{m}_i = \mathbb{E}\big[\tilde{T}_i | X_0 = i\big]$$

Interpretation

$\tilde{u}_i$ can also be viewed as the probability with which i is revisited and $\tilde{m}_i$ is the expected time between consecutive visits.

Classification of States

Recurrence, Transience, and Periodicity

$$\tilde{T}_i = \min\{n > 0 : X_n = i\}$$
$$\tilde{u}_i = \mathbb{P}\left[\tilde{T}_i < \infty | X_0 = i\right]$$
$$\tilde{m}_i = \mathbb{E}\left[\tilde{T}_i | X_0 = i\right]$$

- What can you say about $\tilde{m}_i$ if $\tilde{u}_i < 1$? ($\tilde{u}_i < 1 \Rightarrow \tilde{m}_i = \infty$)
- What can you say about $\tilde{m}_i$ if $\tilde{u}_i = 1$? ($\tilde{u}_i = 1 \not\Rightarrow m_i \neq \infty$)

Definition

A state i is said to be **recurrent** if $\tilde{u}_i = 1$. It is **transient** if $\tilde{u}_i < 1$.

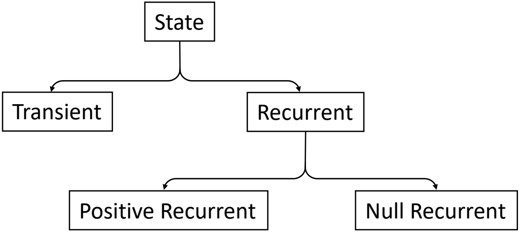
Definition

A recurrent state i is said **positive recurrent** if $\tilde{m}_i < \infty$. It is **null recurrent** if $\tilde{m}_i = \infty$.

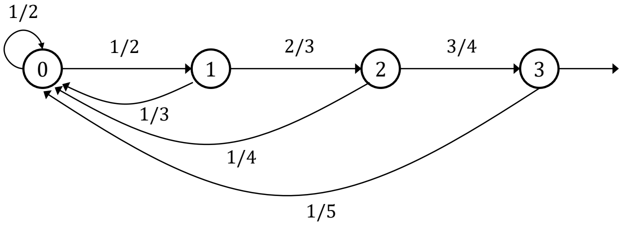
Classification of States

Recap

- Accessibility
- Communicating
- Communicating Class
- Closed Communicating Class
- Irreducibility



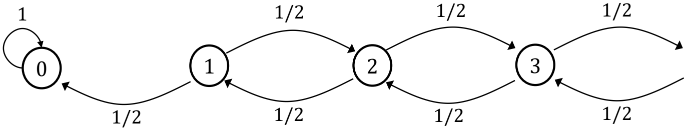
Consider the modified-success runs with the following transition probabilities.



Compute

- $\mathbb{P}\big[\tilde{T}_0 = n | X_0 = 0\big]$, where $n \geq 1$
- $\tilde{u}_0$ and $\tilde{m}_0$

Consider the following Markov chain



Which of the following is true for state 0

1. A closed communicating class
2. Transient state
3. Positive recurrent
4. Null recurrent

What about state 1?

Recurrence and transience of states can also be established using the following set of results.

These results will also come in handy when we study the limiting behavior of DTMCs.

First, recall from the discussion on occupancy times that $V_i^{(n)}$ is the number of visits to i over $\{0, 1, \dots, n\}$. And the occupancy time of i starting from i up to time n is

$$m_{ii}^{(n)} = \mathbb{E}\big[V_i^{(n)} | X_0 = i\big], \forall\, i \in S, n \geq 0$$

Also, recall that

$$m_{ii}^{(n)} = \sum_{r=0}^{n} p_{ii}^{(r)}$$

Classification of States

Recurrence, Transience, and Periodicity

Theorem

A state i is recurrent $\Leftrightarrow \sum_{r=0}^{\infty} p_{ii}^{(r)} = \infty$ *and is transient* $\Leftrightarrow \sum_{r=0}^{\infty} p_{ii}^{(r)} < \infty$

Proof (sketch).

Define a random variable V_i as the number of visits to i over the infinite horizon. $V_i^{(n)} \to V_i$ almost surely. Hence, we can write

$$\begin{aligned}\mathbb{E}\big[V_i | X_0 = i\big] &= \lim_{n\to\infty} \mathbb{E}\big[V_i^{(n)} | X_0 = i\big] \\ &= \lim_{n\to\infty} m_{ii}^{(n)} \\ &= \sum_{r=0}^{\infty} p_{ii}^{(r)}\end{aligned}$$

So we need to show that if state i is recurrent $\Leftrightarrow \mathbb{E}\big[V_i | X_0 = i\big] = \infty$ and is transient $\mathbb{E}\big[V_i | X_0 = i\big] < \infty$

Proof (sketch).

By definition, $\mathbb{E}\big[V_i|X_0 = i\big] = \sum_{k=1}^{\infty} k\mathbb{P}\big[V_i = k|X_0 = i\big]$

If state i is **recurrent**, $\mathbb{P}\big[V_i = \infty|X_0 = i\big] = 1$. Hence,

$$\mathbb{E}\big[V_i|X_0 = i\big] = \infty$$

Recall that $\tilde{u}_i$ can also be interpreted as the probability of coming back to i from i. Hence, if i is **transient**, $\mathbb{P}\big[V_i = k|X_0 = i\big] = \tilde{u}_i^{k-1}(1 - \tilde{u}_i)$. Therefore,

$$\begin{aligned}
\mathbb{E}\big[V_i|X_0 = i\big] &= \sum_{k=1}^{\infty} k\tilde{u}_i^{k-1}(1 - \tilde{u}_i) \\
&= (1 - \tilde{u}_i)\sum_{k=1}^{\infty} k\tilde{u}_i^{k-1} \text{ (arithmetico–geometric series)} \\
&= (1 - \tilde{u}_i)\left[\frac{1}{1 - \tilde{u}_i} + \frac{\tilde{u}_i}{(1 - \tilde{u}_i)^2}\right] = (1 - \tilde{u}_i)^{-1} < \infty
\end{aligned}$$

■

Similar conditions can be derived to distinguish between null and positive recurrent states.

To do so, recall that $\tilde{m}_i$ is the expected time between consecutive visits. Hence, $1/\tilde{m}_i$ is the number of visits to state i in unit time, which can be mathematically written as

$$\lim_{n\to\infty} \frac{m_{ii}^{(n)}}{n+1} = \lim_{n\to\infty} \frac{1}{n+1} \sum_{r=0}^{n} p_{ii}^{(r)}$$

Theorem

A recurrent state is

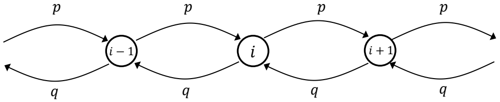
- *Positive recurrent* $\Leftrightarrow \lim_{n\to\infty} \frac{1}{n+1} \sum_{r=0}^{n} p_{ii}^{(r)} > 0$
- *Null recurrent* $\Leftrightarrow \lim_{n\to\infty} \frac{1}{n+1} \sum_{r=0}^{n} p_{ii}^{(r)} = 0$

Classification of States

Periodicity

If the period of a state i is d, then its return times are integer multiples of d. For any d' that is not an integer multiple of d, $p_{ii}^{(d')} = 0$.

What is the period of the states in the simple random walk?



Definition

Let i be a recurrent state and d be the largest positive integer such that

$$\sum_{k=1}^{\infty} \mathbb{P}\left[\tilde{T}_i = kd \middle| X_0 = i\right] = 1$$

- i is aperiodic if $d = 1$
- i is periodic with period d if $d > 1$

Note that periodicity does not depend on the magnitude of transition probabilities. Is success runs aperiodic?

Classification of States

Class Properties

Theorem

If two states $i \leftrightarrow j$, and i is transient (positive recurrent, null recurrent, periodic), then j is transient (positive recurrent, null recurrent, periodic)

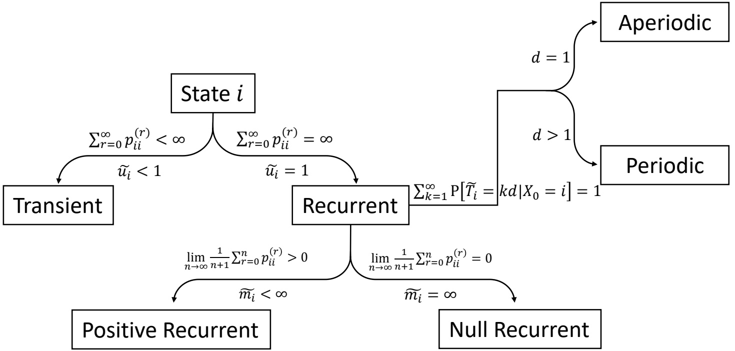
If one state in a communicating/closed communicating class/irreducible DTMC has one of the above properties, then all the remaining states are guaranteed to have the same property.

For this reason, these classification are also referred to as class properties.

Classification of States

Summary

- Accessibility ($i \rightarrow j$)
- Communicating ($i \leftrightarrow j$)
- Communicating Class (All states communicate and the set is maximal)
- Closed Communicating Class ('Blackhole')
- Irreducibility (The entire DTMC is a 'blackhole')



Your Moment of Zen