Reinforcement Learning Notes

Varad Vaidya

February 23, 2024

Lecture notes from the various YouTube playlist related to Reinforcement Learning, combined with the course E1277 Reinforcement Learning at IISc Bangalore by Prof. Gugan Thoppe. The plan is to merge the notes from david silver's course and the course at IISc Bangalore, at appropriate places, to make a single set of notes.

Since the notes are being merged from two different sources, proper crediting of the two (and many more) is hard. In general, the notes will follow, from the standard books of Reinforcement Learning, and are my interpretation of the same.

Disclaimer: This document will inevitably contain some mistakes— both simple typos and legitimate errors. Keep in mind that these are the notes of a graduate student in the process of learning the material, so take what you read with a grain of salt. If you find mistakes and feel like telling me, I will be grateful and happy to hear from you, even for the most trivial of errors. You can reach me by email at vaidyavarad2001@gmail.com.

For more notes like this, visit varadVaidya.

Varad Vaidya,
Fall Term: 2023,
Last Update: February 23, 2024,

Contents

Part I

E1277 — Reinforcement Learning

1 Markov Chains

1.1 Markov Processes or Markov Chains

A Markov chain is simply a Markov Decision Process without decision. It is one the most simplest stocastic process, and has no "memory" of the past. So, just the present stat determines its future dynamics. In this context, we will be considering Discrete Time Markov Chains (DTMCs).

DTMC involves two concepts:

- Discrete Time Stochastic Process (DTSP)
- Row Stochastic Matrix.

The two are defined as follows:

Definition 1.1 (Discrete Time Stochastic Process (DTSP)). A DTSP is a sequence $(X_n)_{n\geq 0}$ of random variables defined on the same proboability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in the same set $\mathcal{S}$, i.e.

$$X_n : \Omega \to \mathcal{S} \quad \forall n \geq 0$$

where $\mathcal{S}$ is called a state space, and in finite or countably infinite. We call the a variable as a "state" when the state is an element of the state space $\mathcal{S}$. The cardinality of the state space is denoted by $|\mathcal{S}|$.

Definition 1.2 (Row Stochastic Matrix). A matrix $\mathcal{P} \in \mathbb{R}^{|\mathcal{S}|\times|\mathcal{S}|}$ is called a row stochastic matrix if it satisfies the following conditions:

$$\mathcal{P}_{ij} \in [0, 1] \quad \forall i, j \in \mathcal{S}$$

$$\sum_{j=1}^{|\mathcal{S}|} \mathcal{P}_{ij} = 1 \quad \forall i \in \mathcal{S}$$

Thus, with these two definitions we can define a Markov Chain (or DTMC) as follows:

Definition 1.3 (Markov Chain). Let $\mathcal{S}$ be a finite state space, and ν be a distribution over $\mathcal{S}$.

$$\nu = (\nu_1, \nu_2, \ldots, \nu_{|\mathcal{S}|}) \quad \text{s.t} \quad \nu_i \in [0, 1]\ \forall i \in \mathcal{S}, \quad \sum_{i\in\mathcal{S}} \nu_i = 1$$

Further, let $\mathcal{P}$ be a row stocastic matrix over $\mathcal{S}$.

Then a DTSP $(X_n)_{n\geq 0}$ is called a Markov Chain with initial distribution ν and transition matrix $\mathcal{P}$ if it satisfies the following conditions:

- Initial state is distributed according to ν, i.e.

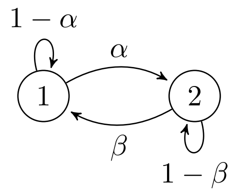
$$\mathbb{P}(X_0 = i) = \nu_i \quad \forall i \in \mathcal{S}$$

- Present is independent of the past given the present.

$$\begin{aligned}\mathbb{P}\left(X_{n+1}=i_{n+1}|X_n=i_n,X_{n-1}=i_{n-1},\ldots,X_0=i_0\right)&=\mathbb{P}\left(X_{n+1}=i_{n+1}|X_n=i_n\right)\\&=\mathcal{P}_{i_n i_{n+1}}\quad\forall n\geq 0\end{aligned}$$

Example (Markov Chains)**.** Let a Markov chain be defined as follows:

$$\mathcal{S}=\{1,2\}\quad \nu=(p,1-p)\quad \mathcal{P}=\begin{bmatrix}1-\alpha & \alpha\\ \beta & 1-\beta\end{bmatrix}$$

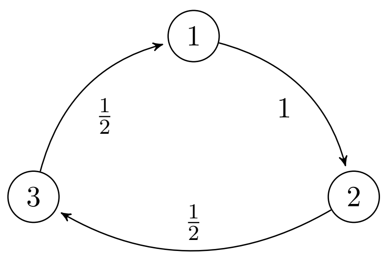


Thus, we can have say:

$$\mathbb{P}(X_3=1|X_0=1,X_1=1,X_2-2)=\beta$$

Another example that we can have is:

$$\mathcal{S}=\{1,2,3\}\quad \nu=(\nu_1,\nu_2,\nu_3)\quad \mathcal{P}=\begin{bmatrix}0 & 1 & 0\\ 0 & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & 0 & \frac{1}{2}\end{bmatrix}$$



given, $\nu=(1,0,0)$, the outcomes of the Markov chain if we sample it are:

$$\begin{aligned}X_n&=1,2,3,3,1,2,\ldots\\&=1,2,3,1,2,\ldots\end{aligned}$$

Note that, to define the DTMC, we need the initial distribution, but the results wont change on the choice of the distribution ν, assuming that the Markov chain is ergodic. Hence, while defining the markov chain, the initial distribution is not mentioned, and can be assumed arbitrarily.

Theorem 1.1 (Necessary and Sufficient Conditions for DTSP to be Markov Chains)**.** A DTSP $(X_n)_{n\geq 0}$ on $\mathcal{S}$ is a Markov chain $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ if and only if:

$$\mathbb{P}\left(X_n = i_n, X_n = i_n, \dots, X_0 = i_0\right) = \nu_{i_0}\mathcal{P}_{i_0 i_1}\mathcal{P}_{i_1 i_2}\dots\mathcal{P}_{i_{n-1}i_n} \quad \forall n \geq 0$$

Proof. To simplify the proof, assume that $\mathcal{P}_{ij} > 0 \quad \forall i, j \in \mathcal{S}$. The theorem remains true in the general case, but requires more book-keeping.

Suppose, $(X_n)_{n\geq 0}$ is a Markov chain $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$.

Using the fact,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A)$$
$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B|A)\mathbb{P}(C|A \cap B)$$

We get,

$$\begin{aligned}
\mathbb{P}(X_0 = i_0, \dots, X_n = i_n) &= \mathbb{P}\left(\{X_0 = i_0\} \cap \{X_1 = i_1\} \cap \dots \cap \{X_n = i_n\}\right) \\
&= \mathbb{P}(X_0 = i_0)\mathbb{P}(X_1 = i_1|X_0 = i_0)\dots\mathbb{P}(X_n = i_n|X_{n-1} = i_{n-1}) \\
&= \mathbb{P}(X_0 = i_0)\mathbb{P}(X_1 = i_1|X_0 = i_0)\dots\mathbb{P}(X_n = i_n|X_{n-1} = i_{n-1}) \\
&\dots \text{Using the memoeryless property of Markov chains} \\
&= \nu_{i_0}\mathcal{P}_{i_0 i_1}\dots\mathcal{P}_{i_{n-1}i_n}
\end{aligned}$$

This proves the forward claim. To show the reverse claim:

Put $n = 0$, in the claim, to trivially get the initial distribution back, showing one part of the definition. For the other part:

$$\begin{aligned}
\mathbb{P}(X_n = i_n|X_{n-1} = i_{n-1}, \dots, X_0 = i_0) &= \frac{\mathbb{P}(X_n = i_n, \dots, X_0 = i_0)}{\mathbb{P}(X_{n-1} = i_{n-1}, \dots, X_0 = i_0)} \\
&= \frac{\nu_{i_0}\mathcal{P}_{i_0 i_1}\dots\mathcal{P}_{i_{n-1}i_n}}{\nu_{i_0}\mathcal{P}_{i_0 i_1}\dots\mathcal{P}_{i_{n-1}i_n}} \\
&= \mathcal{P}_{i_{n-2}i_{n-1}}
\end{aligned}$$

Now we need to show:

$$\begin{aligned}
&\mathbb{P}(X_n = i_n|X_{n-1} = i_{n-1}) \\
&= \frac{\mathbb{P}(X_{n-1} = i_{n-1}, X_n = i_n)}{\mathbb{P}(X_{n-1} = i_{n-1})} \\
&= \frac{\sum_{i_0 \in \mathcal{S}} \mathbb{P}(X_0 = i_0 \dots X_{n-1} = i_{n-1}, X_n = i_n)}{\sum_{i_0 \in \mathcal{S}} \mathbb{P}(X_{n-1} = i_{n-1}, X_0 = i_0)} \\
&= \mathcal{P}_{i_{n-1}i_n}
\end{aligned}$$

😄

In the above proof we have used the following series of fact:

$$\begin{aligned}\mathbb{P}(X_2 = j) &= \mathbb{P}(\Omega \cap \{X_2 = j\}) \\ &= \mathbb{P}\left[\bigcup_{i=1}^{|S|}\{X_i = i\} \cap \{X_2 = j\}\right]\end{aligned}$$

Using the fact:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

we get:

$$\mathbb{P}(X_2 = j) = \sum_{i=1}^{|S|} \mathbb{P}\left(X_i = i, X_2 = j\right)$$

Theorem 1.2 (Markov Property)**.** Let $(X_n)_{n\geq 0}$ be the markov chain detnoted by $\langle \mathcal{S}, \mathcal{P}, \mu \rangle$, then condintional on $\{X_m = i\}$, $\{X_{m+n}\}$ is the markov chain $\langle \mathcal{S}, \mathcal{P}, \delta_i \rangle$ and is independent of $X_1, X_2, \dots, X_m$, where δ_i is the distribution that is 1 at i and 0 everywhere else, i.e

$$\delta_i(j) \equiv \delta_{ij} = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

Proof. It suffices to show the following, for any $n \geq 0$ and on an event A, related to $X_0, \dots, X_m$:

$$\begin{aligned}&\mathbb{P}\left[(X_m = i_m, X_{m+1} = i_{m+1}, \dots, X_{m+n} = i_{m+n}) \cap A \mid X_m = i_m\right] \\ &\qquad = \mathbb{P}(A \mid X_m = i_m)\delta_{ij}\mathcal{P}_{i_m i_{m+1}} \dots \mathcal{P}_{i_{m+n-1} i_{m+n}}\end{aligned}$$

The above equation is just conditional independence:

$$\mathbb{P}(E_1 \cap E_2 \mid X_m = i) = \mathbb{P}(E_1 \mid X_m = i)\mathbb{P}(E_2 \mid X_m = i)$$

where $E_2 = (X_m = i_m, X_{m+1} = i_{m+1}, \dots, X_{m+n} = i_{m+n})$ and $E_1 = A$. The proof of the above statement goes as follows:

Let A be an elementary event, i.e. $A = \{X_0 = j_0, X_1 = j_1, \dots, X_m = j_m\}$. Consider the LHS of the above equation: Then, we have two cases:

- **Case 1**: $j_m \neq i_m$. Then, the LHS is trivially 0, since A and E_2 are disjoint.
- **Case 2**: $j_m = i_m$. Then, we have further cases. When, $i \neq i_m = j_m$, then the LHS and RHS are again 0. But when $i = i_m = j_m$, then the LHS is:

$$\begin{aligned}&\mathbb{P}(X_0 = j_0, X_1 = j_1, \dots X_m = j_m = i, X_m = i_m = i, X_{m+1} = i_{m+1}, \\ &\qquad \dots, X_{m+n} = i_{m+n} \mid X_m = i_m = i)\end{aligned}$$

$$
\begin{aligned}
&= \frac{\nu(j_o)\mathcal{P}_{j_0 j_1}\dots\mathcal{P}_{j_{m-1}j_m}\mathcal{P}_{j_m i_{m+1}}\dots\mathcal{P}_{i_{m+n-1}i_{m+n}}}{\mathbb{P}(X_m = i_m = i)} \\
&= \frac{\mathbb{P}(A)\mathcal{P}_{i_m i_{m+1}}\dots\mathcal{P}_{i_{m+n-1}i_{m+n}}}{\mathbb{P}(X_m = i_m = i)} \\
&= \mathbb{P}(A \mid X_m = i_m = i)\mathcal{P}_{i_m i_{m+1}}\dots\mathcal{P}_{i_{m+n-1}i_{m+n}}
\end{aligned}
$$

Thus, completing the proof. This can be extended to non elementary event A by using the fact that any event A can be written as a union of elementary events, i.e

$$
A = \bigcup_{i=1}^{|S|}\{X_0 = i_0, X_1 = i_1, \dots, X_m = i_m\}
$$

😊

Theorem 1.3 (Linear Algebra and Markov Chains)**.** Let $(X_n)_{n\geq 0}$ be the marrkov chain $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$. Then:

1. $\mathbb{P}(X_n = j) = (\nu\mathcal{P}^n)_j$
2. $\mathbb{P}_i(X_n = j) = \mathbb{P}\left(X_{m+n} = j \mid X_m = i\right) = \mathcal{P}_{ij}^{(n)} \equiv$ ij entry of $\mathcal{P}^n$

Proof. The second statement is obivious from the first statement. The proof of the first statement goes as follows:

Proof by induction: For $n = 0$, we have:

$$
\mathbb{P}(X_0 = j) = \nu_j \equiv \nu(j)
$$

Now, assume that the statement is true for n, then we have:

$$
\begin{aligned}
\mathbb{P}(X_{n+1} = j) &= \mathbb{P}\left(\bigcup_{i=1}^{|S|}\{X_n = i, X_{n+1} = j\}\right) \\
&= \sum_{i=1}^{|S|}\mathbb{P}(X_n = i, X_{n+1} = j) \\
&= \sum_{i=1}^{|S|}\mathbb{P}(X_n = i)\mathbb{P}(X_{n+1} = j \mid X_n = i) \\
&= \sum_{i=1}^{|S|}\mathbb{P}(X_n = i)\mathcal{P}_{ij} \\
&= \sum_{i=1}^{|S|}(\nu\mathcal{P}^n)_j\mathcal{P}_{ij} \\
&= (\nu\mathcal{P}^{n+1})_j
\end{aligned}
$$

Thus, the proof is complete. ☺

1.2 Communicating Classes

Definition 1.4. Let $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ be a markov chain. Then, we say i leads to j denoted as $i \to j$ if:

$$\mathbb{P}_i(X_n = j \text{ for some } n \geq 0) > 0$$

which is equivalent to:

$$\mathbb{P}(X_{m+n} = j \text{ for some } n \geq 0 \mid X_m = i) > 0$$

Definition 1.5. Let $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ be a markov chain. Then, we say i and j communicate denoted as $i \leftrightarrow j$ if:

$$i \to j \text{ and } j \to i$$

Theorem 1.4. The following statements are equivalent:

- $i \leftrightarrow j$
- $\mathcal{P}_{i_0 i_1} \mathcal{P}_{i_1 i_2} \ldots \mathcal{P}_{i_{n-1} i_n} > 0$ for some $n \geq 0$ and some sequence of states $i_0, i_1, \ldots, i_n$ such that $i_0 = i$ and $i_n = j$.
- $\mathcal{P}_{ij}^{(n)} > 0$ for some $n \geq 0$

Proof. i did not understood this :(Revisit this later. ☺

Definition 1.6 (Communicating Class). Each equivalence class of the relation $\leftrightarrow$ is called a communicating class.

A communicating class is called closed if $\forall i \in C, i \to j \Rightarrow j \in C$.

Example. The properties that hold in this example are:

$$\begin{aligned} &1 \leftrightarrow 2 \\ &2 \leftrightarrow 3 \Rightarrow 1 \quad \leftrightarrow 3 \end{aligned}$$

We also have the following:

$$\begin{aligned} &3 \to 4, 5, 6 \\ &4 \to 5, 6 \\ &5 \leftrightarrow 6 \end{aligned}$$

Thus, the following can be concluded:

$$\begin{aligned} \text{Communicating Classes} &= \{1, 2, 3\}, \{4\}, \{5, 6\} \\ \text{Closed Groups} &= \{5, 6\} \end{aligned}$$

1.3 Strong Markov Property

Definition 1.7 (Stopping Time). A random variable $T : \Omega \to \{0, 1, \dots\}$ is called a stopping time if for any $n \geq 0$, the occurance or non occurance of the event $\{T = n\}$ can be determined based only on the values of $X_0, X_1, \dots, X_n$.

Example. The following are examples of stopping times:

First passage time:

$$T_j = \inf\{n \geq 1 : X_n = j\}$$

Last passage time:

$$T_j = \sup\{n \geq 0 : X_n = j\}$$

Definition 1.8 (Strong Markov Property). Let $(X_n)_{n\geq 0}$ be a markov chain $(\mathcal{S}, \mathcal{P}, \mu)$ and let T, be a stopping time. Then, conditional on $T < \infty$ and $X_T = i$, the process $(X_{T+n})_{n\geq 0}$ is a markov chain with initial distribution δ_i and is independent of $X_0, X_1, \dots, X_T$.

1.4 Recurrance and Transience

Let $(X_n)_{n\geq 0}$ be the markov chain $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$.

Definition 1.9 (Recurrence). A state i is called recurrent if:

$$\mathbb{P}_i(X_i = 1 \text{ for infinitely many } n) = 1$$

Definition 1.10 (Transience). A state i is called transient if:

$$\mathbb{P}_i(X_i = 1 \text{ for infinitely many } n) < 1$$

Loosly, a recurrent state is one that the markov chain keeps visiting, while a transient state is one that the markov chain never visits from some time on.

Definition 1.11 (Passage Time). First Passage Time to state j:

$$T_j \coloneqq \inf\{n \geq 1 : X_n = j\}$$

or,

$$T_j : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$$

rth passage time to state j: We can define this inductivly as follows:

$$T_j^{(0)} = 0 \Rightarrow T_j^{(1)} = T_j$$

$$\Rightarrow T_j^{(r+1)} = \inf\{n > T_j^{(r)} : X_n = j\}$$

rth excursion or sojorn time:

$$S_j^{(r)} = \begin{cases} T_j^{(r)} - T_j^{(r-1)} & \text{if } T_j^{(r-1)} < \infty \\ 0 & \text{otherwise} \end{cases}$$

Lemma 1.1. For $r = 2, 3, \cdots$ condintional on $T_i^{(r-1)} < \infty$, S_i^r is independent of X_m for $0 < m < T_i^{(r-1)}$. Further,

$$\mathbb{P}_i(S_i^r = n \mid T_i^{(r-1)} < \infty) = \mathbb{P}_i(T_i = n)$$

Proof.

$$\{T_i^{(r-1)}\} \subseteq \{X_{T_i^{(r-1)}}\}$$

$$\therefore \{T_i^{(r-1)}\} = \{T_i^{(r-1)}\} \cap \{X_{T_i^{(r-1)}} = i\}$$

This implies conditioning on the event $\{T_i^{(r-1)} < \infty\}$ is equivalent to conditioning on $\{T_i^{(r-1)}\} < \infty$ and $\{X_{T_i^{(r-1)}} = i\}$. Further T_i^{r-1} is a stopping time. Thus, conditional on $\{T_i^{(r-1)} < \infty\}$ and $\{X_{T_i^{(r-1)} = i}\}$

$$\left(X_{T_j^{(r-1)}}\right)_{(n \geq 0)} \quad \text{is a markov chain}$$

Now,

$$\begin{aligned} &\mathbb{P}(S_i^{(r)} = n \mid T_i^{(r-1)} < \infty) \\ &= \mathbb{P}(S_i^{(r)} = n \mid T_i^{(r-1)} < \infty, X_{T_i^{(r-1)}} = i) \\ &= \mathbb{P}(X_{T_i^{(r-1)}+1} \neq i, \dots, X_{T_i^{(r-1)}+n} = i \mid T_i^{(r-1)} < \infty, X_{T_i^{(r-1)}} = i) \\ &= \mathbb{P}(X_1 \neq i, \dots, X_n = i \mid X_0 = i) \\ &= \mathbb{P}_i(T_i = n) \end{aligned}$$

Completing the proof. 😃

Definition 1.12 (Number of Visits).

$$V_i = \sum_{n=1}^{\infty} \mathbb{I}\{X_n = i\}$$

Thus, we have the following:

$$\mathbb{E}_i(V_i) = \mathbb{E}_i \sum_{n=0}^{\infty} \mathbb{I}\{X_n = i\} = \sum_{n=0}^{\infty} \mathbb{P}_i(X_n = i) = \sum_{n=0}^{\infty} \mathcal{P}_{ii}^{(n)}$$

we can replace the sum with the expectation, using the monotone convergence theorem.

Lemma 1.2. Let $f_i := \mathbb{P}_i(T_i < \infty)$. Now for $r = 0, 1, \dots$ we have:

$$\mathbb{P}_i(V_i > r) = f_i^r$$

Proof. We can prove this using induction.

Base case: $r = 0$:

$$r = 0 \Rightarrow \mathbb{P}_i(V_i > 0) = 1 \text{ and } f_i^0 = 1$$

Thus, the base case holds true trivially.

Suppose, the given claim holds true for some r. Then,

$$\begin{aligned}
\mathbb{P}_i(V_i > r+1) &= \mathbb{P}_i(T_i^{(r+1)} < \infty) \\
&= \mathbb{P}_i(T_i^{(r)} < \infty, S_i^{(r+1)} < \infty) \\
&= \mathbb{P}_i(T_i^{(r)} < \infty)\mathbb{P}_i(S_i^{(r+1)} < \infty, T_i^{(r)} < \infty) \\
&= \mathbb{P}_i(V_i > r)\mathbb{P}_i(T_i < \infty) \\
&= f_i^r f_i = f_i^{r+1}
\end{aligned}$$

☺

Thus, using the above lemmas we can state the following theorem:

Theorem 1.5. Let $(X_n)_{n\geq 0}$ be the markov chain $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$. Then,

- for $i \in \mathcal{S}$, $\mathbb{P}_i(T_i < \infty) = 1 \Rightarrow i$ is recurrent, and

$$\sum_{n=0}^{\infty} \mathcal{P}_{ii}^{(n)} = \infty$$

- for $i \in \mathcal{S}$, $\mathbb{P}_i(T_i < \infty) < 1 \Rightarrow i$ is transient, and

$$\sum_{n=0}^{\infty} \mathcal{P}_{ii}^{(n)} < \infty$$

Proof. Consider the case of recurrence. We have:

$$\mathbb{P}_i(V_i = \infty) = \mathbb{P}_i\left(\bigcap_{r=1}^{\infty}\{V_i > r\}\right)$$

Since, $\{V_i > r\} \supseteq \{V_i > r+1\}$, using the continuity of probability measure, we get:

$$\mathbb{P}_i(V_i = \infty) = \lim_{r\to\infty} \mathbb{P}_i(V_i > r) = \lim_{r\to\infty} f_i^r = 1$$

$$\Rightarrow \mathbb{E}_i(V_i) = \infty \Rightarrow \sum_{n=0}^{\infty} \mathcal{P}_{ii}^{(n)} = \infty$$

Now, consider the case of transience. Suppose, $f_i = \mathbb{P}_i(T_i < \infty) < 1$. Then, we have:

$$\sum_{n=0}^{\infty} \mathcal{P}_{ii}^{(n)} = \sum_{n=0}^{\infty} \mathbb{P}_i(T_i = n) = \sum_{n=0}^{\infty} f_i^n = \frac{1}{1-f_i} < \infty$$

Hence, $\mathbb{P}_i(V_i = \infty) = 0$, which implies that i is transient. ☺

Theorem 1.6. Let C be a communicating class in a Markov Chain. Then either all states in C are recurrent or all states in C are transient.

Proof. Let C be a communicating class. Suppose $i \in C$ is transient.

$$\Rightarrow \sum_{n=0}^{\infty} \mathcal{P}_{ii}^{(n)} < \infty$$

Let $j \in C$ and $\exists n, m \geq 0$ such that $\mathcal{P}_{ij}^{(n)} > 0$ and $\mathcal{P}_{ji}^{(m)} > 0$. Then, for $r \geq 0$:

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

$$\Rightarrow \mathcal{P}_{jj}^{(r)} \leq \frac{\mathcal{P}_{ii}^{(n+r+m)}}{\mathcal{P}_{ii}^{(n)} \mathcal{P}_{ii}^{(m)}}$$

$$\Rightarrow \sum_{r=0}^{\infty} \mathcal{P}_{jj}^{(r)} \leq \frac{\sum_{r=0}^{\infty} \mathcal{P}_{ii}^{(n+r+m)}}{\mathcal{P}_{ij}^{(n)} \mathcal{P}_{ji}^{(m)}} < \infty$$

☺

Note that recurrence or transience is a class property. And thus the following theorem can be stated:

Theorem 1.7. Every recuurent class must be closed. And, Every finite closed class is recurrent.

Note that, infinite closed classes need not be necessarily recurrent. The classic random walk on a line of intergers is a classic example.

1.5 Invariant Distributions

Definition 1.13 (Invariant Distribution). A distribution Π on $\mathcal{S}$ is called invariant for the markov chain $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ if:

$$\Pi = \Pi \mathcal{P}$$

Thus, a stationery distribution is a left eigenvector of the transition matrix $\mathcal{P}$ with eigenvalue 1. Note that in contrast to linear algebra, vectors are represented as row vectors.

Theorem 1.8. Let $(X_n)_{n\geq 0}$ be the markov chain $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$. Then $X_{m+n_{n\geq 0}}$ is also a markov chain with the distribution $\langle \mathcal{S}, \mathcal{P}, \Pi \rangle$, where Π is the invariant distribution of $\mathcal{P}$.

Proof. We know that,

$$\mathbb{P}(X_n = j) = (\nu \mathcal{P}^n)_j \Rightarrow \mathbb{P}(X_n = j) = (\Pi \mathcal{P}^n)_j = \Pi_j$$

Now,

$$\begin{aligned} &\mathbb{P}(X_m = i_0, X_{m+1} = i_1, \dots, X_{m+n} = i_n) \\ &= \mathbb{P}(X_m = i_0)\mathbb{P}(X_{m+1} = i_1 \mid X_m = i_0) \dots \mathbb{P}(X_{m+n} = i_n \mid X_{m+n-1} = i_{n-1}) \\ &= \Pi_{i_0} \mathcal{P}_{i_0 i_1} \dots \mathcal{P}_{i_{n-1} i_n} \end{aligned}$$

$\Rightarrow (X_{m+n})_{n\geq 0}$ is a markov chain with the distribution Π 😃

Theorem 1.9. Every markov chain on a finite set space has at least one invariant distribution.

Proof. Since $\mathcal{P}$ is a row stocastic matrix:

$$\mathcal{P}\mathbf{1} = \mathbf{1}$$

Since $\mathcal{P}$ and $\mathcal{P}^\top$ have the same eigenvalues, we have:

$$\Rightarrow \exists \nu \neq 0 \text{ such that } \nu \mathcal{P} = \nu$$

Since, $\mathcal{P}$ is a real valued matrix, taking the complex conjugate of the above equation, we get:

$$\overline{\nu} \mathcal{P} = \overline{\nu}$$

Thus, adding and subtracting the above two equations, we get:

$$\Re(\nu)\mathcal{P} = \Re(\nu) \quad \text{and} \quad \Im(\nu)\mathcal{P} = \Im(\nu)$$

Since, $v \neq 0$, atleast one of $\Re(\nu)$ or $\Im(\nu)$ is non zero. Thus, if $\mathcal{P}$ has a complex left eigenvector, then it has a real left eigenvector.

Now, without loss of generality, let u be a real valued vector such that:

$$u\mathcal{P} = u$$

Defining,

$$u_+(i) := \max\{u(i), 0\} \quad \text{and} \quad u_-(i) := \max\{-u(i), 0\}$$

we have:

$$\Rightarrow u = u_+ - u_-$$

letting,

$$u_+\mathcal{P} =: u_+ \quad \text{and} \quad u_-\mathcal{P} =: u_-$$

we get,

$$u\mathcal{P} = u_+\mathcal{P} - u_-\mathcal{P} = y_+ - y_-$$

Suppose: $u_+(i) > 0 \Rightarrow u_-(i) = 0$.

$$\Rightarrow y_+(i) - y_-(i) = u_+(i) \Rightarrow y_+(i) = u_+(i) \text{ and } y_-(i) = 0$$

Similiarly

$$u_+(i) = 0 \Rightarrow u_-(i) > 0 \Rightarrow y_+(i) = 0 \text{ and } y_-(i) = u_-(i)$$

Thus, we can conclude:

$$u_+\mathcal{P} = u_+ \quad \text{and} \quad u_-\mathcal{P} = u_-$$

Since $u \neq 0$, either of u_+ or u_- is non zero. Thus, we have found a non zero left real valued eigenvector of $\mathcal{P}$.

Let z be the non zero vector among u_+ or u_-. Then, we have:

$$z\mathcal{P} = z \quad \text{and} \quad z \neq 0,\ z_i > 0 \Rightarrow \sum_i z_i > 0$$

Rescaling the vector we get,

$$\frac{z}{\sum_i z_i}\mathcal{P} = \frac{z}{\sum_i z_i}$$

Defining $\Pi(i) = \frac{z}{\sum_i z_i}$, we get:

$$\Pi\mathcal{P} = \Pi$$

☺

Definition 1.14. A markov chain is said tp be irreducible if the whole state space is one communicating class. i.e

$$i \leftrightarrow j \quad \forall\, i, j \in \mathcal{S}$$

Theorem 1.10. Let $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ be an irreducible and recurrent markov chian. Further let

$$\gamma_i^k = \mathbb{E}_k \sum_{n=0}^{T_k - 1} \mathbb{I}\{X_n = i\}$$

and $\gamma^k = (\gamma_1^k, \gamma_2^k, \ldots, \gamma_{|\mathcal{S}|}^k)$. Then, the following holds:

1. $\gamma_k^k = 1$
2. $\gamma^k P = \gamma^k$
3. $0 < \gamma^k < \infty$

Proof. **(1)**: This is true by the definition of γ_i^k.

(2): We have:

$$
\begin{aligned}
\gamma_j^k &= \mathbb{E}_k \sum_{n=0}^{T_k-1} \mathbb{I}\{X_n = j\} \\
&= \mathbb{E}_k \sum_{n=1}^{T_k} \mathbb{I}\{X_n = j\} = \mathbb{E}_k \sum_{n=0}^{\infty} \mathbb{I}\{X_n = j, n \leq T_k\} \\
&= \sum_{n=1}^{\infty} \mathbb{P}_k(X_n = j, n \leq T_k) \qquad \text{//by monotone convergence theorem} \\
&= \sum_{n=1}^{\infty} \sum_{i \in \mathcal{S}} \mathbb{P}_k(X_n = j, X_{n-1} = i, n \leq T_k) \\
&= \sum_{n=1}^{\infty} \sum_{i \in \mathcal{S}} \mathbb{P}_k(X_n = j \mid X_{n-1} = i, n \leq T_k)\mathbb{P}_k(X_{n-1} = i, n \leq T_k) \\
&= \sum_{n=1}^{\infty} \sum_{i \in \mathcal{S}} \mathcal{P}_{ij}\mathbb{P}_k(X_{n-1} = i, n \leq T_k)
\end{aligned}
$$

Since we are summing over non negative sum, we can interchange the summation. Thus, we get:

$$
\begin{aligned}
\gamma_j^k &= \sum_{i \in \mathcal{S}} \mathcal{P}_{ij} \sum_{n=1}^{\infty} \mathbb{P}_k(X_{n-1} = i, n \leq T_k) \\
&= \sum_{i \in \mathcal{S}} \mathcal{P}_{ij} \underbrace{\sum_{m=0}^{\infty} \mathbb{P}_k(X_m = i, m \leq T_k - 1)}_{\mathbb{E}_k \sum_{m=0}^{T_k-1} \mathbb{I}\{X_m = i\}} \\
&= \sum_{i \in \mathcal{S}} \mathcal{P}_{ij}\gamma_i^k \\
&\Rightarrow \gamma^k \mathcal{P} = \gamma^k
\end{aligned}
$$

(3): We will use the fact that the given markov chain is irreducible.

$$\Rightarrow \mathcal{P}_{ij}^{(n)} > 0 \quad \forall\, i, j \in \mathcal{S}$$

Since we know that γ^k is invariant $\Rightarrow \gamma^k = \gamma^k \mathcal{P} = \gamma^k \mathcal{P}^n$. Thus, we have:

$$\gamma_k^k = \sum_j \gamma_j^k \mathcal{P}_{jk}^{(n)} \geq \gamma_i^k \mathcal{P}_{ik}^{(n)}$$

The above statement holds, since sum of non negative quantities is greater than or equal to any of the individual quantities. Now,

$$\gamma_k^k = 1 \quad \because \mathcal{P}_{ik}^{(n)} > 0 \Rightarrow \gamma_i^k < \infty$$

Now, to show $\gamma_i^k > 0$. Let $\mathcal{P}_{ki}^{(m)} > 0$

$$\Rightarrow \gamma_i^k = \sum_j \gamma_j^k \mathcal{P}_{ji}^{(m)} \geq \gamma_k^k \mathcal{P}_{ki}^{(m)} = \mathcal{P}_{ki}^{(m)} > 0$$

$$\Rightarrow 0 < \gamma_i^k < \infty$$

Completing the proof. ☺

Theorem 1.11. Let $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ be an irreducible markov chain. Further, let λ be an invariant measure such that $\lambda_k = 1$, then $\lambda \geq \gamma^k$. If $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ is additionally recurrent, then $\gamma^k = \lambda$.

Proof. NOTE: This theorem also applies with inifinte markov chain.
Since λ, is a measure, thus, $\lambda(i)$ is non negtive. Now, since λ is invariant, we have:

$$\lambda = \lambda \mathcal{P} \Rightarrow \lambda(j) = \sum_{i \in \mathcal{S}} \lambda(i) \mathcal{P}_{ij} = \sum_{i_1 \neq k} \lambda(i) \mathcal{P}_{ij} + \mathcal{P}_{kj}$$

Massaging $\mathcal{P}_{kj}$, we get:

$$\mathcal{P}_{kj} = \mathbb{P}(X_1 = j \mid X_0 = k) = \mathbb{P}_k(X_1 = j) = \mathbb{P}_k(X_1 = j, T_k > 1)$$

Thus,

$$\begin{aligned}
\lambda_j &= \sum_{i_1 \neq k} \sum_{i_2 \in \mathcal{S}} \lambda_{i_2} \mathcal{P}_{i_2 i_1} \mathcal{P}_{i_1 j} + \mathbb{P}_k(X_1 = j, T_k > 1) \\
&= \sum_{i_1 \neq k} \sum_{i_2 \neq k} \lambda_{i_2} \mathcal{P}_{i_2 i_1} \mathcal{P}_{i_1 j} + \mathbb{P}_k(X_1 = j, T_k > 1) + \mathbb{P}_k(X_2 = j, T_k > 2)
\end{aligned}$$

From induction, we get:

$$\lambda_j = \underbrace{\text{some expression}}_{\text{non negative}} + \sum_{n=1}^{m} \mathbb{P}_k(X_n = j, T_k > n)$$

$$\Rightarrow \lambda_j \geq \sum_{n=1}^{\infty} \mathbb{P}_k(X_n = j, T_k > n) = \gamma_j^k \qquad //\text{by Theorem 1.10}$$

Thus, we have:

$$\lambda \geq \gamma^k$$

Now, if the markov chain is recurrent, then we have to show that $\lambda = \gamma^k$. Let $\mu = \lambda - \gamma^k$. Thus, $\mu \mathcal{P} = (\lambda - \gamma^k)\mathcal{P} = \lambda - \gamma^k = \mu$. Now,

$$\mu_k = \lambda_k - \gamma_k^k = 1 - 1 = 0 \quad \mu = \mu \mathcal{P}$$

Choose $i, k \in \mathcal{S}$ such that $\mathcal{P}_{ik}^{(n)} > 0$. Then,

$$\begin{aligned} 0 = \mu_k &= \sum_{j \in \mathcal{S}} \mu_j \mathcal{P}_{jk}^{(n)} \\ &\geq \mu_i \mathcal{P}_{ik}^{(n)} \Rightarrow \mu_i = 0 \forall\, i \in \mathcal{S} \\ &\Rightarrow \mu = 0 \Rightarrow \lambda = \gamma^k \end{aligned}$$

Completing the proof. 😄

Theorem 1.12. Let $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ be a finite irreducible markov chain. Then, Π is unique and $\Pi_i = \frac{1}{m_i}$, where $m_i = \mathbb{E}_i(T_i)$.

Proof. The intuitive meaning of m_i is the expected time to return to state i, starting from state i.
Since the markov chain is finite, there exits a invariant distribution Π such that $\Pi = \Pi\mathcal{P}$.
Now,

$$\pi_k = \sum_{j \in \mathcal{S}} \pi_j \mathcal{P}_{jk}^{(n)} \geq \pi_i \mathcal{P}_{ik}^{(n)} > 0$$

Thus, $\pi_i > 0$. Now, let $\lambda := \frac{\Pi}{\pi_k}$ Thus, $\lambda_k = 1$

$$\Rightarrow \lambda\mathcal{P} = \frac{1}{\pi_k}\Pi\mathcal{P} = \frac{1}{\pi_k}\Pi = \lambda$$

$$\Rightarrow \lambda \text{ is an invariant measure}$$

$\because$ chain is irreducible and from Theorem 1.10, $\lambda = \gamma^k$

$$\Rightarrow \frac{\Pi}{\pi_k} = \gamma^k \Rightarrow \sum_i \frac{\Pi_i}{\pi_k} = \sum_i \gamma_i^k \Rightarrow \frac{1}{\pi_k} = \sum_i \gamma^k$$

$$\because \gamma^k > 0 \Rightarrow \pi_k = \frac{1}{\sum_i \gamma^k}$$

Now,

$$\Rightarrow \sum_i \gamma_i^k = \sum_{i \in \mathcal{S}} \sum_{n=1}^{\infty} \mathbb{P}_k(X_n = i, n \leq T_k)$$

Now, sum is over all positive quantities, summation can be interchanged. Thus, we get:

$$\begin{aligned} \sum_i \gamma_i^k &= \sum_{n=1}^{\infty} \sum_{i \in \mathcal{S}} \mathbb{P}_k(X_n = i, n \leq T_k) \\ &= \sum_{n=1}^{\infty} \mathbb{P}_k(T_k \geq n) = \mathbb{E}_k(T_k) = m_k \Rightarrow \pi_k &= \frac{1}{\mathbb{E}_k(T_k)} \end{aligned}$$

Thus, completing the proof. 😄

Definition 1.15. A state i, is sais to be a aperiodic if $\mathcal{P}_{ii} > 0$ for sufficient large n.

Lemma 1.3. $\langle \mathcal{S}, \mathcal{P}, \sigma \rangle$ is an irreducible markov chain further, suppose $k \in \mathcal{S}$ is aperiodic. Thus, every $i \in \mathcal{S}$, is aperiodic. Infact, $\mathcal{P}_{ii}^{(n)} > 0$, $\forall i, j \in \mathcal{S}$ for sufficient large n.

Theorem 1.13. Let $\langle \mathcal{S}, \mathcal{P}, \nu \rangle$ be a finite irreducible and aperiodic markov chain. Then, irrespective of the initial distribution, the markov chain converges to the unique invariant distribution Π. i.e.

$$\lim_{n \to \infty} \mathbb{P}(X_n = i) = \Pi_i \quad \forall\, i \in \mathcal{S}$$

Further,

$$\lim_{n \to \infty} \mathcal{P}_{ij}^{(n)} = \Pi_j$$