

STA447

Lecture Notes

Yuchen Wang

February 5, 2020

Contents

1	Markov Chain Probabilities	2
1.1	Markov Chain examples	2
1.2	Elementary Computations	3
1.3	Recurrence and Transience	4
1.4	Communicating States and Irreducibility	6
2	Markov Chain Convergence	10
2.1	Stationary Distributions	10
2.2	Searching for Stationary	10
2.3	Obstacles to Convergence	12
2.4	Convergence Theorem	12

1 Markov Chain Probabilities

Notation 1.1.

$$P(X_{n+1} = j | X_n = i) = p_{ij}$$

Definition 1.1 (Markov chain). A (discrete time, discrete space, time homogeneous) Markov chain is specified by three ingredients:

- A state space S , any non-empty finite or countable set.
- Initial probabilities $\{v_i\}_{i \in S}$, where v_i is the probability of starting at i (at time 0). (So $v_i \geq 0$ and $\sum_i v_i = 1$)
- Transition probabilities $\{p_{ij}\}_{i,j \in S}$, where p_{ij} is the probability of jumping to j if you start at i . (So $p_{ij} \geq 0$, and $\sum_j p_{ij} = 1$ for all i)

Remark 1.1 (Markov property).

$$P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, X_2 = i_2, \dots, X_n = i_n) = P(X_{n+1} = j | X_n = i_n) = p_{i_n j}$$

i.e. The probabilities at time $n + 1$ depend only on the state at time n .

Remark 1.2.

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = v_{i_0} p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n}$$

1.1 Markov Chain examples

Example 1.1 (the Frog Walk). Let $X_n :=$ pad index the frog is at after n steps.

$$\begin{aligned} S &= \{1, 2, 3, \dots, 20\} \\ v_{20} &= 1, v_i = 0 \forall i \neq 20 \\ p_{ij} &= \begin{cases} \frac{1}{3}, & |j - i| \leq 1 \text{ or } |j - i| = 19 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

Example 1.2 (Bernoulli process).

$$\begin{aligned} S &= \{1, 2, 3, \dots\} \\ v_0 &= 1, v_i = 0 \forall i \neq 0 \\ p_{ij} &= \begin{cases} p, & j = i + 1 \\ 1 - p, & j = i \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

where $0 < p < 1$.

Example 1.3 (Simple random walk (s.r.w.)). Let $X_n :=$ net gain (in dollars) after n bets

$$\begin{aligned} S &= \{0, 1, 2, 3, \dots\} \\ v_a &= 1, v_i = 0 \forall i \neq a \\ p_{ij} &= \begin{cases} p, & j = i + 1 \\ 1 - p, & j = i - 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

where $0 < p < 1, a \in \mathbb{Z}$.

Special case: When $p = 1/2$, call it simple symmetric random walk.

Example 1.4 (Ehrenfest's Urn). Let $X_n := \#$ balls in Urn 1 at time n .

We have d balls in total, divided into two urns. At each time, we choose one of the d balls uniformly at random, and move it to the other urn.

$$\begin{aligned} S &= \{1, 2, 3, \dots, d\} \\ v_a &= 1, v_i = 0 \forall i \neq a \\ p_{ij} &= \begin{cases} (d-i)/d, & j = i+1 \\ i/d, & j = i-1 \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

1.2 Elementary Computations

Notation 1.2.

$$\mu_i^{(n)} := P(X_n = i)$$

Notation 1.3.

$$\begin{aligned} m &:= |S| && \text{(the number of elements in S, could be infinity)} \\ \mu^{(n)} &= (\mu_1^{(n)}, \mu_2^{(n)}, \mu_3^{(n)}, \dots) && (m \times 1) \\ v &= (v_1, v_2, v_3, \dots) && (m \times 1) \\ P &= (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & \\ & \ddots & & \\ p_{m1} & \dots & & p_{mm} \end{pmatrix} && (m \times m \text{ matrix}) \end{aligned}$$

Fact 1.1.

$$\begin{aligned} \mu^{(1)} &= vP = \mu^{(0)}P \\ \mu^{(n)} &= vP^n = \mu^{(0)}P^n \end{aligned}$$

Notation 1.4.

$$p_{ij}^{(n)} := P(X_n = j, X_0 = i) = P(X_{m+n} = j | X_m = i) \quad (\text{for any } m \in \mathbb{N})$$

Fact 1.2.

$$\begin{aligned} \sum_{j \in S} p_{ij}^{(n)} &= 1 \\ p_{ij}^{(1)} &= p_{ij} \\ P^{(n)} &= P^n \end{aligned} \quad (\text{for all } n \in \mathbb{N})$$

Notation 1.5.

$$\begin{aligned} P^0 &:= I \\ P^{(0)} &:= I \\ p_{ij}^{(0)} &= \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Theorem 1.1 (Chapman-Kolmogorov equations).

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$$

$$P_{ij}^{(m+s+n)} = \sum_{k \in S} \sum_{l \in S} p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(n)}$$

Matrix form:

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

$$P^{(m+s+n)} = P^{(m)} P^{(s)} P^{(n)}$$

Theorem 1.2 (Chapman-Kolmogorov Inequality).

$$p_{ij}^{(m+n)} \geq p_{ik}^{(m)} p_{kj}^{(n)} \quad (\text{for all } k \in S)$$

$$P_{ij}^{(m+s+n)} \geq p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(n)} \quad (\text{for any } k, l \in S)$$

1.3 Recurrence and Transience

Notation 1.6.

$$P_i(\dots) \equiv P(\dots | X_0 = i)$$

$$E_i(\dots) \equiv E(\dots | X_0 = i)$$

$$N(i) = \#\{n \geq 1 : X_n = i\} \quad (\text{total number of times that the chain hits } i, \text{ not counting time } 0)$$

Definition 1.2 (**return probability**). Let f_{ij} be the return probability from i to j .

$$f_{ij} := P_i(X_n = j \text{ for some } n \geq 1) \equiv P_i(N(j) \geq 1)$$

Fact 1.3.

$$1 - f_{ij} = P_i(X_n \neq j \text{ for all } n \geq 1) \quad (1)$$

$$P_i(N(i) \geq k) = (f_{ii})^k \quad (2)$$

$$P_i(N(j) \geq k) = f_{ij} (f_{jj})^{k-1} \quad (3)$$

$$f_{ik} \geq f_{ij} f_{jk} \quad (4)$$

Fact 1.4. $f_{ij} > 0$ iff $\exists m \geq 1$ with $p_{ij}^{(m)} > 0$, i.e., there is some time m for which it is possible to get from i to j in m steps.

Definition 1.3 (**recurrent and transient states**). A state i of a Markov chain is recurrent if $f_{ii} = 1$. Otherwise, i is transient if $f_{ii} < 1$.

Proposition 1.1. If Z is a non-negative integer, then

$$E(Z) = \sum_{k=1}^{\infty} P(Z \geq k)$$

Theorem 1.3 (**Recurrent State Theorem**). As follows

- State i is recurrent $\iff P_i(N(i) = \infty) = 1 \iff \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$
- State i is transient $\iff P_i(N(i) = \infty) = 0 \iff \sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$

Proof.

$$\begin{aligned}
 P_i(N(i) = \infty) &= \lim_{k \rightarrow \infty} P_i(N(i) \geq k) && \text{(by continuity of probabilities)} \\
 &= \lim_{k \rightarrow \infty} (f_{ii})^k && (P_i(N(i) \geq k) = (f_{ii})^k) \\
 &= \begin{cases} 1, & f_{ii} = 1 \\ 0, & f_{ii} < 1 \end{cases}
 \end{aligned}$$

■

Therefore,

$$\begin{aligned}
 \sum_{n=1}^{\infty} p_{ii}^{(n)} &= \sum_{n=1}^{\infty} P_i(X_n = i) \\
 &= \sum_{n=1}^{\infty} E_i(\mathbf{1}\{X_n = i\}) \\
 &= E_i\left(\sum_{n=1}^{\infty} \mathbf{1}\{X_n = i\}\right) \\
 &= E_i(N(i)) \\
 &= \sum_{k=1}^{\infty} P_i(N(i) \geq k) && \text{(by proposition 1.1)} \\
 &= \sum_{k=1}^{\infty} (f_{ii})^k \\
 &= \begin{cases} \infty, & f_{ii} = 1 \\ \frac{f_{ii}}{1-f_{ii}} < \infty, & f_{ii} < 1 \end{cases}
 \end{aligned}$$

Example 1.5 (simple random walk). If $p = 1/2$ then $\forall i, f_{ii} = 1$. If $p \neq 1/2$, then $\forall i, f_{ii} < 1$

Proof. Consider state 0. We need to check if $\sum_{n=1}^{\infty} p_{00}^{(n)} = \infty$.

If n is odd, then $p_{00}^{(n)} = 0$.

If n is even, $p_{00}^{(n)} = P(\frac{n}{2} \text{ heads and } \frac{n}{2} \text{ tails on first } n \text{ tosses})$.

This is a Binomial(n, p) distribution, so

$$\begin{aligned}
 p_{00}^{(n)} &= \binom{n}{n/2} p^{n/2} (1-p)^{n/2} \\
 &= \frac{n!}{[(n/2)!]^2} p^{n/2} (1-p)^{n/2} \\
 &= \frac{(n/e)^n \sqrt{2\pi n}}{[(n/2e)^{n/2} \sqrt{2\pi n/2}]^2} p^{n/2} (1-p)^{n/2} && \text{(Stirling's approximation)} \\
 &= [4p(1-p)]^{n/2} \sqrt{2/\pi n}
 \end{aligned}$$

Case 1: If $p = 1/2$, then $4p(1-p) = 1$, so

$$\begin{aligned}
 \sum_{n=1}^{\infty} p_{00}^{(n)} &= \sum_{n=2,4,6,\dots}^{\infty} \sqrt{2/\pi n} \\
 &= \sqrt{2/\pi} \sum_{n=2,4,6,\dots}^{\infty} n^{-1/2} \\
 &= \sqrt{2/\pi} \sum_{n=1}^{\infty} 2k^{-1/2} \\
 &= \infty
 \end{aligned}$$

Therefore, state 0 is recurrent.

Case 2: If $p \neq 1/2$, then $4p(1-p) < 1$, so

$$\begin{aligned} \sum_{n=1}^{\infty} p_{00}^{(n)} &= \sum_{n=2,4,6,\dots} [4p(1-p)]^{n/2} \sqrt{2/\pi n} \\ &< \sum_{n=2,4,6,\dots} [4p(1-p)]^{n/2} && \text{(Geometric Series)} \\ &= \frac{4p(1-p)}{1-4p(1-p)} \\ &< \infty \end{aligned}$$

Therefore, the state 0 is transient.

The same exact calculation applies to any other state i . ■

Theorem 1.4 (f-Expansion).

$$f_{ij} = p_{ij} + \sum_{k \in S, k \neq j} p_{ik} f_{kj}$$

Proof.

$$\begin{aligned} f_{ij} &= P_i(\exists n \geq 1 : X_n = j) \\ &= \sum_{k \in S} P_i(X_1 = k, \exists n \geq 1 : X_n = j) \\ &= P_i(X_1 = j, \exists n \geq 1 : X_n = j) + \sum_{k \neq j} P_i(X_1 = k, \exists n \geq 1 : X_n = j) \\ &= P_i(X_1 = j)P_i(\exists n \geq 1 : X_n = j | X_1 = j) + \sum_{k \neq j} P_i(X_1 = k)P_i(\exists n \geq 1 : X_n = j | X_1 = k) \\ &= p_{ij}(1) + \sum_{k \neq j} p_{ik}(f_{kj}) \end{aligned}$$
■

Remark 1.3. The f-Expansion shows that $f_{ij} \geq p_{ij}$.

Remark 1.4. It essentially follows from logical reasoning: from i , to get to j eventually, we have to either jump to j immediately (with probability p_{ij}), or jump to some other state k (with probability p_{ik}) and then get to j eventually (with probability p_{kj})

1.4 Communicating States and Irreducibility

Definition 1.4 (communicating states). State i communicates with state j , written $i \rightarrow j$, if $f_{ij} > 0$.

Remark 1.5. i.e. if it is possible to get from i to j .

Notation 1.7. Write $i \leftrightarrow j$ if both $i \rightarrow j$ and $j \rightarrow i$.

Definition 1.5 (irreducibility). A Markov chain is irreducible if $i \rightarrow j$ for all $i, j \in S$, i.e., if $f_{ij} > 0$ for all $i, j \in S$. Otherwise, the chain is reducible.

Lemma 1.1 (Sum Lemma). If $i \rightarrow k$, and $l \rightarrow j$, and $\sum_{n=1}^{\infty} p_{kl}^{(n)} = \infty$, then $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$

Proof. Since $i \rightarrow k$, and $l \rightarrow j$, there exists $m, r \geq 1$ s.t. $p_{ik}^{(m)} > 0$ and $p_{lj}^{(r)} > 0$.

By the Chapman-Kolmogorov inequality,

$$p_{ij}^{(m+s+r)} \geq p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(r)}$$

Hence

$$\begin{aligned}
\sum_{n=1}^{\infty} p_{ij}^{(n)} &\geq \sum_{n=m+r+1}^{\infty} p_{ij}^{(n)} \\
&= \sum_{s=1}^{\infty} p_{ij}^{(m+s+r)} && (s = n - m - r) \\
&\geq \sum_{s=1}^{\infty} p_{ij}^{(m)} p_{kl}^{(s)} p_{lj}^{(r)} \\
&= \underbrace{p_{ij}^{(m)}}_{+} \underbrace{p_{lj}^{(r)}}_{+} \underbrace{\sum_{s=1}^{\infty} p_{kl}^{(s)}}_{=\infty} \\
&= \infty
\end{aligned}$$

■

Corollary 1.1 (Sum Corollary). If $i \leftrightarrow k$, then i is recurrent iff k is recurrent.

Proof. Setting $j = i$ and $l = k$ in the Sum Lemma: If $i \leftrightarrow k$, then $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty \iff \sum_{n=1}^{\infty} p_{kk}^{(n)} = \infty$. ■

Theorem 1.5 (Cases Theorem). For an **irreducible** Markov chain, either

- (a) $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$ for all $i, j \in S$, and all states are recurrent (recurrent Markov chain); or
- (b) $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$ for all $i, j \in S$, and all states are transient (transient Markov chain).

Theorem 1.6 (Finite Space Theorem). An irreducible Markov chain on a **finite** state space always falls into case (a), i.e., $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$ for all $i, j \in S$, and all states are recurrent.

Proof. Choose any state $i \in S$. We have

$$\begin{aligned}
\sum_{j \in S} \sum_{n=1}^{\infty} p_{ij}^{(n)} &= \sum_{n=1}^{\infty} \sum_{j \in S} p_{ij}^{(n)} && \text{(exchanging the sums)} \\
&= \sum_{n=1}^{\infty} 1 \\
&= \infty
\end{aligned}$$

Then if S is finite, it follows that there must exist at least one $j \in S$ with $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$. So we must be in case (a). ■

Notation 1.8. For $i \neq j$, let H_{ij} be the event that the chain hits the state i before returning to j , i.e.,

$$H_{ij} = \{\exists n \in \mathbb{N} : X_n = i, \text{ but } X_m \neq j \text{ for } 1 \leq m \leq n-1\}$$

Lemma 1.2 (Hit Lemma). If $j \rightarrow i$ with $j \neq i$, then $P_j(H_{ij}) > 0$.

Proof. Since $j \rightarrow i$, there is some possible path from j to i . i.e., there is $m \in \mathbb{N}$ and x_0, x_1, \dots, x_m with $x_0 = j$ and $x_m = i$ and $p_{x_r x_{r+1}} > 0$ for all $0 \leq r \leq m-1$.

Let $S = \max\{r : x_r = j\}$ be the last time this path hits j .

Then x_S, x_{S+1}, \dots, x_m is a possible path which goes from j to i without first returning to j .

Hence $P_j(H_{ij}) \geq P(x_0, x_1, \dots, x_m) = p_{x_S x_{S+1}} p_{x_{S+1} x_{S+2}} \cdots p_{x_{m-1} x_m} > 0$ ■

Remark 1.6. If it is possible to get from j to i at all, then it is possible to get from j to i without first returning to j .

Intuitively obvious: If there is some path from j to i , then the final part of the path (starting with the last time it visits i) is a possible path from j to i which does not return to j .

Lemma 1.3 (f-Lemma). If $j \rightarrow i$ and $f_{jj} = 1$, then $f_{ij} = 1$

Proof. If $i = j$ it is trivial, so assume $i \neq j$.

Since $j \rightarrow i$, we have $P_j(H_{ij}) > 0$ by the Hit Lemma.

But one way to never return to j is to first hit i and then from i never return to j :

$$P_j(\text{never return to } j) \geq P_j(H_{ij})P_i(\text{never return to } j)$$

Therefore

$$1 - f_{jj} \geq P_j(H_{ij})(1 - f_{ij})$$

Since $f_{jj} = 1$, then $\underbrace{P_j(H_{ij})}_{>0}(1 - f_{ij}) = 0$

Hence $f_{ij} = 1$. ■

Lemma 1.4 (Infinite Returns Lemma). For an irreducible Markov chain, if it is recurrent, then

$$P_i(N(j) = \infty) = 1$$

for all $i, j \in S$.

But if it is transient, then $P_i(N(j) = \infty) = 0$ for all $i, j \in S$.

Proof. Let $i, j \in S$. If the chain is recurrent, then $f_{ij} = f_{jj} = 1$ by the f-Lemma.

Then

$$\begin{aligned} P_i(N(j) = \infty) &= \lim_{k \rightarrow \infty} P_i(N(j) \geq k) \\ &= \lim_{k \rightarrow \infty} f_{ij}(f_{jj})^{k-1} \\ &= \lim_{k \rightarrow \infty} (1)(1)^{k-1} \\ &= 1 \end{aligned}$$

If the chain is transient, then $f_{jj} < 1$, then

$$\begin{aligned} P_i(N(j) = \infty) &= \lim_{k \rightarrow \infty} P_i(N(j) \geq k) \\ &= \lim_{k \rightarrow \infty} f_{ij}(f_{jj})^{k-1} \\ &= \lim_{k \rightarrow \infty} (1)(f_{jj})^{k-1} \\ &= 0 \end{aligned}$$
■

Theorem 1.7 (Recurrence Equivalence Theorem). If a chain is irreducible, then the following are equivalent (and all correspond to case (a)):

1. There are $k, l \in S$ with $\sum_{n=1}^{\infty} p_{kl}^{(n)} = \infty$.
2. For all $i, j \in S$, we have $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$.
3. There is $k \in S$ with $f_{kk} = 1$, i.e. k is recurrent.

4. For all $j \in S$, we have $f_{jj} = 1$, i.e. all states are recurrent.
5. For all $i, j \in S$, we have $f_{ij} = 1$.
6. There are $k, l \in S$ with $P_k(N(l) = \infty) = 1$.
7. For all $i, j \in S$, we have $P_i(N(j) = \infty) = 1$.

Proof. Follow from results that we have already proven

- 1 \implies 2: Sum Lemma.
- 2 \implies 4: Recurrent State Theorem (with $i = j$).
- 4 \implies 5: f-Lemma.
- 5 \implies 3: immediate.
- 3 \implies 1: Recurrent State Theorem (with $l = k$).
- 4 \implies 7: Infinite Returns Lemma.
- 7 \implies 6: Immediate.
- 6 \implies 3: Recurrent State Theorem (with $l = k$).

■

Theorem 1.8 (Transience Equivalence Theorem). If a chain is **irreducible**, then the following are equivalent (and all correspond to case (b)):

1. There are $k, l \in S$ with $\sum_{n=1}^{\infty} p_{kl}^{(n)} < \infty$.
2. For all $i, j \in S$, we have $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$.
3. For all $k \in S$, we have $f_{kk} < 1$, i.e. k is transient.
4. There is $j \in S$ with $f_{jj} < 1$, i.e. some state is recurrent.
5. There are $i, j \in S$ with $f_{ij} < 1$.
6. For all $k, l \in S$, $P_k(N(l) = \infty) = 0$.
7. There are $i, j \in S$ with $P_i(N(j) = \infty) = 0$.

Remark 1.7 (closed subset note). Suppose a chain is reducible, but it has a closed subset $C \subseteq S$ (i.e. $p_{ij} = 0$ for $i \in C$ and $j \notin C$) on which it is irreducible (i.e. $i \rightarrow j$ for all $i, j \in C$). Then, the Recurrence Equivalence Theorem and other results about irreducible chains still apply to the chain when **restricted** to C .

Proposition 1.2. For simple random walk with $p > 1/2$, $f_{ij} = 1$ whenever $j > i$. (Similarly, if $p < 1/2$ and $j < i$, then $f_{ij} = 1$.)

Proof. Let $X_0 = 0$, and $Z_n = X_n - X_{n-1}$ for $n = 1, 2, \dots$, so that $X_n = \sum_{i=1}^n Z_i$. Since Z_n s iid with $P(Z_n = 1) = p$ and $P(Z_n = -1) = 1 - p$, then by Law of Large Numbers,

$$\lim_{n \rightarrow \infty} \frac{1}{n} (Z_1 + Z_2 + \dots + Z_n) \stackrel{p}{=} E(Z_1) = p(1) + (1-p)(-1) = 2p - 1 > 0$$

$$\begin{aligned} \implies \infty &= \lim_{n \rightarrow \infty} (Z_1 + Z_2 + \dots + Z_n) \\ &= \lim_{n \rightarrow \infty} X_n - X_0 \\ &= \lim_{n \rightarrow \infty} X_n \end{aligned}$$

But if $i < j$, then to go from i to ∞ , the chain must pass through j , so $f_{ij} = 1$.

■

2 Markov Chain Convergence

2.1 Stationary Distributions

Definition 2.1 (stationary distributions). If π is a probability distribution on S (i.e. $\pi_i \geq 0$ for all $i \in S$, and $\sum_{i \in S} \pi_i = 1$), then π is stationary for a Markov chain with transition probabilities (p_{ij}) if $\sum_{i \in S} \pi_i p_{ij} = \pi_j$ for all $j \in S$ (or $\pi P = \pi$, in matrix notation).

Remark 2.1. Intuitively, π being stationary means if the chain starts with probabilities $\{\pi_i\}$, then it will keep the same probabilities one time unit later.

Definition 2.2 (doubly stochastic). A Markov Chain is doubly stochastic if in addition to the usual condition that $\sum_{j \in S} p_{ij} = 1$ for all $i \in S$, $\sum_{i \in S} p_{ij} = 1$ for all $j \in S$.

Remark 2.2. This holds for the Frog Example.

Proposition 2.1. If a Markov chain with states S satisfies $|S| < \infty$ and is doubly stochastic, then the uniform distribution on S is a stationary distribution.

Proof. Let $\{\pi_i\}$ be a distribution such that $\pi_i = \frac{1}{|S|}$.
Then

$$\begin{aligned} \sum_{i \in S} \pi_i p_{ij} &= \sum_{i \in S} \frac{1}{|S|} p_{ij} \\ &= \frac{1}{|S|} \sum_{i \in S} p_{ij} \\ &= \frac{1}{|S|} (1) && \text{(doubly stochastic)} \\ &= \frac{1}{|S|} \\ &= \pi_j \end{aligned}$$

Then $\{\pi_i\}$ is stationary. ■

2.2 Searching for Stationary

Definition 2.3 (reversibility). A Markov chain is reversible (or time reversible, or satisfies detailed balance) with respect to a probability distribution $\{\pi_i\}$ if $\pi_i p_{ij} = \pi_j p_{ji}$ for all $i, j \in S$.

Proposition 2.2. If a chain is reversible with respect to π , then π is a stationary distribution.

Proof. Reversibility means $\pi_i p_{ij} = \pi_j p_{ji}$, so then for $j \in S$,

$$\sum_{i \in S} \pi_i p_{ij} = \sum_{i \in S} \pi_j p_{ji} = \pi_j \sum_{i \in S} p_{ji} = \pi_j (1) = \pi_j$$
■

Lemma 2.1 (M-test). Let $\{x_{nk}\}_{n,k \in \mathbb{N}}$ be a collection of real numbers. Suppose that $\lim_{n \rightarrow \infty} x_{nk}$ exists for each fixed $k \in \mathbb{N}$. Suppose further that $\sum_{k=1}^{\infty} \sup_n |x_{nk}| < \infty$. Then $\lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} x_{nk} = \sum_{k=1}^{\infty} \lim_{n \rightarrow \infty} x_{nk}$.

Proposition 2.3 (Vanishing Probabilities Proposition). If a Markov chain's transition probabilities satisfy that $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ for all $i, j \in S$, then the chain does **not** have a stationary distribution.

Proof. Suppose for contradiction that there is a stationary distribution π . Then we would have $\pi_j = \sum_{i \in S} \pi_i p_{ij}^{(n)}$ for any n , so

$$\begin{aligned} \pi_j &= \lim_{n \rightarrow \infty} \pi_j = \lim_{n \rightarrow \infty} \sum_{i \in S} \pi_i p_{ij}^{(n)} \\ &= \lim_{n \rightarrow \infty} \sum_{i \in S} \pi_i p_{ij}^{(n)} \\ &= \sum_{i \in S} \lim_{n \rightarrow \infty} \pi_i p_{ij}^{(n)} \quad (\text{exchange the sum and the limit, which is valid by M-test}) \\ &= \sum_{i \in S} \pi_i \lim_{n \rightarrow \infty} p_{ij}^{(n)} \\ &= \sum_{i \in S} 0 \\ &= 0 \end{aligned}$$

So we would have $\pi_j = 0$ for all j . But this means that $\sum_j \pi_j = 0$, which is a contradiction. ■

Lemma 2.2 (Vanishing Lemma). If a Markov chain has some $k, l \in S$ with $\lim_{n \rightarrow \infty} p_{kl}^{(n)} = 0$, then for any $i, j \in S$ with $k \rightarrow i$ and $j \rightarrow l$, $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

Proof. Since $k \rightarrow i$ and $j \rightarrow l$, we can find $r, s \in \mathbb{N}$ with $p_{ki}^{(r)} > 0$ and $p_{jl}^{(s)} > 0$. Then by the Chapman-Kolmogorov Inequality,

$$p_{kl}^{(r+n+s)} \geq p_{ki}^{(r)} p_{ij}^{(n)} p_{jl}^{(s)}$$

Hence

$$p_{ij}^{(n)} \leq p_{kl}^{(r+n+s)} / p_{ki}^{(r)} p_{jl}^{(s)}$$

But the assumptions imply that

$$\lim_{n \rightarrow \infty} \left[p_{kl}^{(r+n+s)} / p_{ki}^{(r)} p_{jl}^{(s)} \right] = 0$$

Hence

$$\begin{aligned} 0 &\leq \lim_{n \rightarrow \infty} p_{ij}^{(n)} \leq 0 \\ \implies \lim_{n \rightarrow \infty} p_{ij}^{(n)} &= 0 \end{aligned}$$
■

Corollary 2.1 (Vanishing Together Corollary). For an **irreducible** Markov chain, either

1. $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ for all $i, j \in S$, or
2. $\lim_{n \rightarrow \infty} p_{ij}^{(n)} \neq 0$ for all $i, j \in S$

Corollary 2.2 (Vanishing Probabilities Corollary). If an **irreducible** Markov chain's transition probabilities satisfy that $\lim_{n \rightarrow \infty} p_{kl}^{(n)} = 0$ for some $k, l \in S$, then the chain does not have a stationary distribution.

Lemma 2.3. If the x_n s are non-negative, and $\sum_{n=1}^{\infty} x_n < \infty$, then $\lim_{n \rightarrow \infty} x_n = 0$.

Corollary 2.3 (Transient Not Stationary Corollary). A Markov chain which is **irreducible and transient** cannot have a stationary distribution.

Proof. If a chain is irreducible and transient, then by the Transience Equivalence Theorem, $\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$ for all $i, j \in S$. Hence $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ for all $i, j \in S$.

Thus by the Vanishing Probabilities Corollary, there is no stationary distribution. ■

2.3 Obstacles to Convergence

Definition 2.4 (period). The period of a state i is the greatest common divisor (gcd) of the set $\{n \geq 1 : p_{ii}^{(n)} > 0\}$, i.e. the largest number m such that all the values of n with $p_{ii}^{(n)} > 0$ are all integer multiples of m . If the period of each state is 1, we say the chain is aperiodic; otherwise we say the chain is periodic.

Remark 2.3. Intuitively, the period of a state i is the pattern of returning to i from i . e.g. If the period of i is 2, then it is only possible to get from i to i in an even numbers of steps.

Fact 2.1. If state i has period t , and $p_{ii}^{(m)} > 0$, then m is an integer multiple of t , i.e., t divides m .

Fact 2.2. If $p_{ii} > 0$, then the period of state i is 1.

Fact 2.3. If $p_{ii}^{(n)} > 0$ and $p_{ii}^{(n+1)} > 0$, then the period of state i is 1.

Lemma 2.4 (Equal Periods Lemma). If $i \leftrightarrow j$, then the periods of i and of j are equal.

Proof. Let the periods of i and j be t_i and t_j . Since $i \leftrightarrow j$, we can find $r, s \in \mathbb{N}$ with $p_{ij}^{(r)} > 0$ and $p_{ji}^{(s)} > 0$. Then

$$p_{ii}^{(r+s)} \geq p_{ij}^{(r)} p_{ji}^{(s)} > 0$$

Therefore by Fact 2.1, t_i divides $r + s$.

Suppose now that $p_{jj}^{(n)} > 0$. Then

$$p_{ii}^{(r+n+s)} \geq p_{ij}^{(r)} p_{jj}^{(n)} p_{ji}^{(s)} > 0$$

So t_i divides $r + n + s$.

Since t_i divides both $r + n + s$ and $r + s$, then it must divide n as well.

Since this is true for any n with $p_{jj}^{(n)} > 0$, it follows that t_i is a common divisor of $\{n \in \mathbb{N} : p_{jj}^{(n)} > 0\}$.

But t_j is the **greatest** such common divisor, so $t_j \geq t_i$.

Similarly we can show that $t_i \geq t_j$, so we have $t_i = t_j$. ■

Corollary 2.4 (Equal Periods Corollary). If a chain is **irreducible**, then all states have the same period.

Corollary 2.5. If a chain is **irreducible** and $p_{ii} > 0$ for some state i , then the chain is **aperiodic**.

2.4 Convergence Theorem

Theorem 2.1 (Markov Chain Convergence Theorem). If a Markov chain is **irreducible**, **aperiodic**, and has a **stationary distribution** $\{\pi_i\}$, then $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j$ for all $i, j \in S$, and $\lim_{n \rightarrow \infty} P(X_n = j) = \pi_j$ for any initial probabilities $\{v_i\}$.

Theorem 2.2 (Stationary Recurrence Theorem). If chain **irreducible** and has a **stationary distribution**, then it is **recurrent**.

Proof. The Transient Not Stationary Corollary says that a chain cannot be irreducible, transient and have a stationary distribution.

Therefore, if a chain is irreducible and has a stationary distribution, then it cannot be transient, i.e. it must be recurrent. ■

Lemma 2.5 (Number Theory Lemma). If a set A of positive integers is non-empty, and satisfies additivity, and $\gcd(A) = 1$, then there is some $n_0 \in \mathbb{N}$ s.t. for all $n \geq n_0$ we have $n \in A$ i.e. the set A includes all of the integers $n_0, n_0 + 1, n_0 + 2, \dots$

Proposition 2.4. If a state i has $f_{ii} > 0$ and is aperiodic, then there is $n_0(i) \in \mathbb{N}$ such that $p_{ii}^{(n)} > 0$ for all $n \geq n_0(i)$

Proof. Let $A = \{n \geq 1 : p_{ii}^{(n)} > 0\}$. Since $f_{ii} > 0$, then A is not empty.

If $m, n \in A$, then

$$p_{ii}^{(m+n)} \geq p_{ii}^{(m)} p_{ii}^{(n)} > 0$$

So $m + n \in A$, which shows that A satisfies additivity. Also $\gcd(A) = 1$ since the state i is aperiodic. Hence from the Number Theory Lemma, there is $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, we have $n \in A$ i.e. $p_{ii}^{(n)} > 0$. ■

Corollary 2.6. If a chain is irreducible and aperiodic, then for any states $i, j \in S$, there is $n_0(i, j) \in \mathbb{N}$ s.t. $p_{ij}^{(n)} > 0$ for all $n \geq n_0(i, j)$

Proof. Find $n_0(i)$ as in Proposition 2.3, and find $m \in \mathbb{N}$ with $p_{ij}^{(m)} > 0$.

Then let $n_0(i, j) = n_0(i) + m$

Then if $n \geq n_0(i, j)$, then $n - m \geq n_0(i)$, so $p_{ij}^{(n)} \geq p_{ii}^{(n-m)} p_{ij}^{(m)} > 0$. ■

Lemma 2.6 (Markov Forgetting Lemma). If a Markov chain is irreducible and aperiodic, and has stationary distribution $\{\pi_i\}$, then for all $i, j, k \in S$,

$$\lim_{n \rightarrow \infty} |p_{ik}^{(n)} - p_{jk}^{(n)}| = 0$$

Remark 2.4. Intuitively, after a long time n , the chain “forgets” whether it started from state i or from state j .

Proof. _____ ■

long

Proof of Markov Chain Convergence Theorem _____

long

Corollary 2.7. If a chain is irreducible, then it has at most one stationary distribution.

Proof. By Markov Chain Convergence Theorem, any stationary distribution that it has must be equal to $\lim_{n \rightarrow \infty} P(X_n = j)$, so it is unique. ■