Markov Chains

{Section 1}

• Markov Property: Let $X = (X_0, X_1, \dots)$ be sequence of rv's taking values in state space, S. (assumed to be countable). Then X has the markov property if:

$$\left\{ \forall n \geq 0, i_0, \dots, i_{n+1} \in S: \quad \mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n) \right\}$$

If this holds, then X is a Markov Chain.

X is homogenous if $\mathbb{P}(X_{n+1} = j | X_n = i)$ is independent of n.

To fully specify a homogeneous Markov Chain, we require:

(i) The initial distribution, $\lambda = (\lambda_i : i \in S)$, $\lambda_i = \mathbb{P}(X_0 = i)$.

(ii) Transition probabilities, $p_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i)$.

{N.B We see that, $\sum_j p_{i,j} = \sum_j \mathbb{P}(X_1 = j | X_0 = i) = 1$.}

Thm: Let λ be a distribution (on S) and P a stochastic matrix. Then $X = (X_0, \dots)$ is a MC with initial distribution λ and transition matrix P iff:

$$\left\{ \mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \lambda_{i_0} p_{i_0, i_1} \cdots p_{i_{n-1}, i_n} \quad (*) \quad \forall n, i_0, \dots, i_n \right\}$$

Pf: Let A_k be event $X_k = i_k$, then we have:

$$\mathbb{P}(A_0 \cap A_1 \cap \dots \cap A_n) = \lambda_{i_0} p_{i_0, i_1} \cdots p_{i_{n-1}, i_n}$$

First, assume X is a MC. Now prove $(*)$ by induction on n.

$n = 0$: $\mathbb{P}(A_0) = \lambda_{i_0}$, which is true by definition of λ.

Assume true $\forall n < N$, then:

$$\mathbb{P}(A_0 \cap \dots \cap A_N) = \mathbb{P}(A_0, \dots, A_{N-1}) \mathbb{P}(A_0, \dots, A_N | A_0, \dots, A_{N-1})$$

$$= \lambda_{i_0} p_{i_0,i_1} \cdots p_{i_{N-2},i_{N-1}} \mathbb{P}(A_N | A_0, \dots, A_{N-1})$$

$$= \lambda_{i_0} p_{i_0,i_1} \cdots p_{i_{N-2},i_{N-1}} \mathbb{P}(A_N | A_{N-1})$$ by Markov property.

$$= \lambda_{i_0} p_{i_0,i_1} \cdots p_{i_{N-2},i_{N-1}} p_{i_{N-1},i_N}$$

So true by induction.

Conversely, suppose (*) holds, then for $n=0$, $\mathbb{P}(X_0=i_0)=\lambda_{i_0}$

Otherwise:

$$\mathbb{P}(X_n=i_n | X_0=i_0, \dots, X_{n-1}=i_{n-1}) = \mathbb{P}(A_n | A_0 \cap \dots \cap A_{n-1})$$

$$= \frac{\mathbb{P}(A_0 \cap \dots \cap A_n)}{\mathbb{P}(A_0 \cap \dots \cap A_{n-1})}$$

$$= p_{i_{n-1},i_n}$$

which is independent of $\{i_0, \dots, i_{n-2}\}$. So X is Markov.

Thm: Extended Markov property: for $n \geq 0$, any H given in terms of $\{X_i : i < n\}$ and any F given in terms of $\{X_i : i > n\}$, we have:

$$\{ \ \mathbb{P}(F | X_n = i, H) = \mathbb{P}(F, X_n = i) \ \}$$

· Transition Probability: $p_{i,j} = \mathbb{P}(X_1 = j | X_0 = i)$, however we also define the n-step transition probability: $p_{i,j}(n) = \mathbb{P}(X_n = j | X_0 = i)$

We have:

Chapman-Kolmagorov Equations.

$$\left\{ \begin{aligned} p_{i,j}(m+n) &= \mathbb{P}(X_{m+n} = j | X_0 = i) \\ &= \sum_{k \in S} \mathbb{P}(X_{m+n} = j | X_m = k, X_0 = i) \mathbb{P}(X_m = k | X_0 = i) \\ &= \sum_{k \in S} \mathbb{P}(X_{m+n} = j | X_m = k) \mathbb{P}(X_m = k | X_0 = i) \\ &= \sum_{k \in S} p_{i,k}(m) p_{k,j}(n) \end{aligned} \right\}$$

In matrix form, this gives us: $P(m+n) = P(m)P(n) \Rightarrow P(n) = P^n$.

When finding P^n, we may either find the eigenvalues and consider certain entries e.g $p_{1,2}(0) = 0$ etc. or we may apply the Chapman-Kolmogorov equations to generate a difference equation.

Distribution of X_n: Consider X_0 with distribution λ:

$$\mathbb{P}(X_1 = j) = \sum_{i \in S} \mathbb{P}(X_1 = j \mid X_0 = i)\,\mathbb{P}(X_0 = i) = \sum_{i \in S} \lambda_i p_{i,j}$$ So, X_1 has distribution λP (row vector).

Similarly, X_n has the distribution λP^n.

{Section 2}

• Communicating Classes: Focus attention on irreducible chains, where we may freely move between different states. {unique communicating class.}

We define: $i \to j$ for $i, j \in S$ if, $\exists n \geq 0 : p_{i,j}(n) > 0$.

$i \leftrightarrow j$ if $i \to j$, $i \leftarrow j$. $\Rightarrow$ i, j communicate.

We may see that $\leftrightarrow$ is an equivalence relation:

Pf: 1) $i \leftrightarrow i$ (reflexive)

2) (symmetric) by definition

3) Suppose we have, $i \leftrightarrow j$ and $j \leftrightarrow k$, then $\exists m, n \geq 0$: $p_{i,j}(m) > 0$, $p_{j,k}(n) > 0$. Then:

$$p_{i,k}(m+n) = \sum_{r \in S} p_{i,r}(m)\, p_{r,k}(n) \geq p_{i,j}(m)\, p_{j,k}(n) > 0 \text{ (transitive).}$$

The equivalence classes of this relation are precisely the communicating classes of the Markov Chain. We see these are not completely isolated, but once we have left a class, we may not return. Thus, for a state space with finitely many communicating classes, there will be an ultimate class which the chain remains within.

Definition: A subset $C \subseteq S$ is closed if $p_{i,j} = 0$ $\forall i \in C$, $\forall j \notin C$.

Proposition: C is closed $\Leftrightarrow$ "$i \in C$, $i \to j \Rightarrow j \in C$"

Pf: Assume C closed, let $i \in C$, $i \to j$, so $\exists m \geq 0$: $p_{i,j}(m) > 0$

So: $$p_{i,j}(m) = \sum_{i_1, \dots, i_{m-1} \in S} \{p_{i,i_1} \cdots p_{i_{m-1},j}\} > 0$$

So, there is some route, $i \to i_1 \to \cdots \to i_{m-1} \to j$ such that $p_{i,i_1} \cdots p_{i_{m-1},j} > 0$. Since $p_{i,i_1} > 0 \Rightarrow i_1 \in C$, then by induction, $j \in C$.

Conversely: for any $i \in C, j \notin C$, $i \not\to j$. So in particular $p_{i,j} = 0$

• Recurrence or transience: a state is recurrent if it is certain we will return, otherwise it is transient. We make some notational definitions:

$$\left\{ \begin{array}{l} 1)\ \mathbb{P}_i(A) = \mathbb{P}(A \mid X_0 = i) \\ 2)\ \mathbb{E}_i(Z) = \mathbb{E}(Z \mid X_0 = i) \end{array} \right\}$$

Definition: The first-passage time is the number of steps it takes to reach a state j, from i, for the first time. Defined by:

$$\left\{ T_j = \min\{n \geq 1 : X_n = j\} \right\}$$

The first-passage probability is:

$$\left\{ f_{ij}(n) = \mathbb{P}_i(T_j = n) \right\}$$

Definition: A state $i \in S$ is recurrent if:

$$\left\{ \mathbb{P}_i(T_i < \infty) = 1 \right\} \quad \text{else, it is transient.}$$

We may show that if i is recurrent, the probability that we return infinitely often to i is 1.

Thm: $\left\{ i \text{ is recurrent iff } \sum_{n=0}^{\infty} p_{i,i}(n) = \infty \right\}$

We must establish a number of pre-requisites first:

We define two generating functions:

$$1)\ P_{i,j}(s) = \sum_{n=0}^{\infty} p_{i,j}(n) s^n$$

$$2)\ F_{i,j}(s) = \sum_{n=0}^{\infty} f_{i,j}(n) s^n$$

From which we note that $p_{i,j}(0) = \delta_{i,j}$, $f_{i,j}(0) = 0$.

Thm: $\left\{ P_{i,j}(s) = \delta_{i,j} + F_{i,j}(s)\, P_{j,j}(s) \right\}$ for $-1 < s \leq 1$

Pf: By law of total probability:

$$p_{i,j}(n) = \sum_{m=1}^{n} \mathbb{P}_i(X_n = j \mid T_j = m)\, \mathbb{P}_i(T_j = m)$$

By the Markov property:

$$p_{i,j}(n) = \sum_{m=1}^{n} \mathbb{P}_i(X_n = j \mid X_m = j)\, \mathbb{P}_i(T_j = m)$$

$$= \sum_{m=1}^{n} p_{j,j}(n-m)\, f_{i,j}(m)$$

Multiplying through by s^n and summing:

$$\underbrace{\sum_{n=1}^{\infty} p_{i,j}(n) s^n}_{} = \underbrace{\sum_{n=1}^{\infty} \sum_{m=1}^{n} p_{j,j}(n-m) s^{n-m} f_{i,j}(m) s^m}_{}$$

convolution of power series $P_{j,j}(s)$, $F_{i,j}(s)$

$$P_{i,j}(s) - \delta_{i,j} = P_{j,j}(s) F_{i,j}(s)$$

$$\Rightarrow P_{i,j}(s) = \delta_{i,j} + P_{j,j}(s) F_{i,j}(s).$$

We may now prove the original theorem, using Abel's Lemma:

$$u(s) = \sum u_n s^n \Rightarrow \lim_{s \to 1^-} \{u(s)\} = \sum u_n.$$

Pf: Take $j = i$ in above formula:

$$P_{i,i}(s) = \frac{1}{1 - F_{i,i}(s)}$$

We need to check that $F_{i,i}(s) \neq 1$, else we have an issue.

$$F_{i,i}(s) = \sum_{n=1}^{\infty} f_{i,i}(n) s^n$$

By definition, $F_{i,i}(1) = \sum_n f_{i,i}(n) = \mathbb{P}(\text{ever returning to } i) \leq 1$

So, for $|s| < 1$, $F_{i,i}(s) < 1$. Now, we take limit as $s \to 1$

Then: $\lim_{s \to 1} P_{i,i}(s) = P_{i,i}(1) = \sum_n p_{i,i}(n)$

and $\lim_{s \to 1} \left\{ \frac{1}{1 - F_{i,i}(s)} \right\} = \frac{1}{1 - \sum f_{i,i}(n)}$

So: $\sum' p_{i,i}(n) = \dfrac{1}{1-\sum' f_{i,i}(n)}$

Thus, $\sum' p_{i,i}(n) = \infty \iff \mathbb{P}(\text{ever returning to } i) = 1$

i.e state i is recurrent $\square$

Thm: Let C be a communicating class:

i) Either every state is recurrent, or every state is transient

ii) If C contains a recurrent state, then C is closed.

Pf: Let $i \leftrightarrow j$, $i \neq j$. So, $\exists m: p_{i,j}(m) = \alpha > 0$, $\exists n: p_{j,i}(n) = \beta > 0$

So, for each k:

$p_{i,i}(m+k+n) \geq p_{i,j}(m) p_{j,j}(k) p_{j,i}(n) = \alpha\beta p_{j,j}(k)$

So, if $\sum'_k p_{j,j}(k) = \infty$, then $\sum'_r p_{i,i}(r) = \infty$. So j recurrent $\iff$ i recurrent.

For second statement, if C is not closed, there is non-zero probability we leave class and never return.

Thm: In a finite state space:

i) There exists at least one recurrent state.

ii) If the chain is irreducible, every state is recurrent.

This is really a statement about probability flow, in a finite state space, the probability cannot "drift off" to infinity.

Pf: ii) follows from i) since either all states are transient or recurrent. So we prove i).

Recall $P_{i,j}(s) = \delta_{i,j} + P_{j,j}(s) F_{i,j}(s)$

If j is transient, $\sum'_n p_{j,j}(n) = P_{j,j}(1) < \infty$. Also $F_{i,j}(1) < \infty$ since it is the probability of ever reaching j from i. So:

$P_{i,j}(1) < \infty$. By Abel's Lemma, $P_{i,j}(1)$ is given by:

$P_{i,j}(1) = \sum'_n p_{i,j}(n) < \infty \Rightarrow p_{i,j}(n) \to 0$.

Since we know that $\sum'_{j \in S} p_{i,j}(n) = 1$, if every state is transient, then $\sum' p_{i,j}(n) \to 0$ as $n \to \infty$ which is a contradiction.

Thm: (Pólya's Theorem). Consider $\mathbb{Z}^d = \{(x_1, \dots, x_d) : x_i \in \mathbb{Z}\}$. This generates a graph with x adjacent to y if $|x-y| = 1$.

Consider a random walk on $\mathbb{Z}^d$:

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} \frac{1}{2d} & \text{if } |j-i| = 1 \\ 0 & \text{otherwise} \end{cases}$$

This is an irreducible chain, so either all states are recurrent or all are transient. The theorem says the chain is recurrent iff $d = 1, 2$.

Pf: Start with $d=1$. We wish to show $\sum_i p_{0,0}(n) = \infty$ (i.e origin is recurrent). We may simplify this by noting it is impossible to return in an odd no. of steps. So, consider $\sum_i p_{0,0}(2n)$.

But: $p_{0,0}(2n) = \mathbb{P}(n \text{ steps to left}, n \text{ steps to right}) = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}$

We use Stirling's formula, $n! \simeq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ to get:

$p_{0,0}(2n) \sim \frac{1}{\sqrt{\pi n}}$. $\Rightarrow \sum_i p_{0,0}(2n) = \infty$.

Now, $d=2$ case: suppose after $2n$ steps I have made r steps right, l steps left, u steps up and d steps down. So $r+l+u+d = 2n$.

To return, we need $r = l$, $u = d$. So let $r = l = m$, $u = d = n-m$.

Then:
$$\begin{aligned}
p_{0,0}(2n) &= \left(\frac{1}{4}\right)^{2n} \sum_{m=0}^{n} \binom{2n}{m, m, n-m, n-m} \\
&= \left(\frac{1}{4}\right)^{2n} \sum_{m=0}^{n} \frac{(2n)!}{(m!)^2((n-m)!)^2} \\
&= \left(\frac{1}{4}\right)^{2n} \binom{2n}{n} \sum_{m=0}^{n} \left(\frac{n!}{m!(n-m)!}\right)^2 \\
&= \left(\frac{1}{4}\right)^{2n} \binom{2n}{n} \underbrace{\sum_{m=0}^{n} \binom{n}{m}\binom{n}{n-m}} \\
&= \left(\frac{1}{4}\right)^{2n} \binom{2n}{n} \binom{2n}{n} \\
&= \left[\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right]^2 \\
&\sim \frac{1}{\pi n}
\end{aligned}$$

split $2n$ in two equal parts, choose m from first, $(n-m)$ from second.

square of $d=1$ probability.

So $\sum_i p_{0,0}(2n) = \infty$. $\therefore d = 1, 2$ recurrent.

For $d=3$;
$$\begin{aligned}
p_{0,0}(2n) &= \sum_{\substack{i,j,k \geq 0 \\ i+j+k=n}} \frac{(2n)!}{(i!j!k!)^2}\left(\frac{1}{6}\right)^{2n} \\
&= \binom{2n}{n}\left(\frac{1}{2}\right)^{2n} \sum_{\substack{i,j,k \geq 0 \\ i+j+k=n}} \binom{n}{i,j,k}^2 \left(\frac{1}{3}\right)^{2n}
\end{aligned}$$

Now; $\sum_{\substack{i,j,k\geq 0 \\ i+j+k=n}} \binom{n}{i,j,k}\left(\frac{1}{3}\right)^n = 1$

total prob of ways of placing n balls in 3 urns.

For the case $n=3m$, $\binom{n}{i,j,k} = \frac{n!}{i!j!k!} \leq \frac{n!}{m!m!m!}$ $\forall i,j,k$.

So: $p_{0,0}(2n) \leq \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\binom{n}{m,m,m}\left(\frac{1}{3}\right)^n \sim \frac{1}{2A^3}\left(\frac{6}{n}\right)^{3/2}$ as $n\to\infty$

So: $\sum p_{0,0}(6m) < \infty$. But $p_{0,0}(6m) \geq \left(\frac{1}{6}\right)^2 p_{0,0}(6m-2)$ and $p_{0,0}(6m) \geq \left(\frac{1}{6}\right)^4 p_{0,0}(6m-4)$. So $\sum p_{0,0}(n) < \infty$.

Thus $d=3$ is transient, and for $d>3$, we may project $\mathbb{Z}^d$ onto $\mathbb{Z}^3$.

- <u>Hitting Probabilities:</u> Let S be state space and $A\subseteq S$. we wish to know how likely + how long it takes to reach A.

<u>Definition:</u> (Hitting Time) The random variable, $\{H^A = \min\{n\geq 0 : X_n \in A\}\}$

In particular, if we start in A, $H^A = 0$.

we also have : $\{h_i^A = \mathbb{P}_i(H^A<\infty) = \mathbb{P}_i(\text{ever reach } A)\}$

<u>Thm:</u> The vector $(h_i^A : i\in S)$ satisfies:

$$h_i^A = \begin{cases} 1 & i\in A \\ \sum_{j\in S} p_{i,j}h_j^A & i\notin A. \end{cases}$$

and is minimal in that for any non-negative solution $(x_i; i\in S)$, $h_i^A \leq x_i$ $\forall i$.

<u>Pf:</u> By definition, $h_i^A = 1$ if $i\in A$. Otherwise:

$$h_i^A = \mathbb{P}_i(H^A<\infty) = \sum_{j\in S}\mathbb{P}_i(H^A<\infty \mid X_1=j)p_{i,j} = \sum_{j\in S} h_j^A p_{i,j}$$

To show it is the minimal solution, suppose $x=(x_i : i\in S)$ is a non-negative solution i.e:

$$x_i^A = \begin{cases} 1 & i\in A \\ \sum_{j\in S} p_{i,j}x_j^A & i\notin A \end{cases}$$

If $i\in A$, we have $h_i^A = x_i^A = 1$. Otherwise we may write:

$$\begin{aligned} x_i &= \sum_j{}' p_{i,j} x_j \\ &= \sum_{j\in A}{}' p_{i,j} x_j + \sum_{j\notin A}{}' p_{i,j} x_j \\ &= \sum_{j\in A}{}' p_{i,j} + \sum_{j\notin A}{}' p_{i,j} x_j \\ &\geq \sum_{j\in A}{}' p_{i,j} = \mathbb{P}_i(H^A = 1) \end{aligned}$$

By iterating the process:

$$\begin{aligned} x_i &= \sum_{j\in A}{}' p_{i,j} + \sum_{j\notin A}{}' p_{i,j} \left(\sum_k{}' p_{j,k} x_k\right) \\ &= \sum_{j\in A}{}' p_{i,j} + \sum_{j\notin A}{}' p_{i,j} \left(\sum_{k\in A}{}' p_{j,k} x_k + \sum_{k\notin A}{}' p_{j,k} x_k\right) \\ &\geq \mathbb{P}_i(H^A = 1) + \sum_{j\notin A, k\in A}{}' p_{i,j} p_{j,k} \\ &= \mathbb{P}_i(H^A = 1) + \mathbb{P}_i(H^A = 2) = \mathbb{P}_i(H^A \leq 2) \end{aligned}$$

By induction, we obtain:

$x_i \geq \mathbb{P}_i(H^A \leq n)$ for all n. Taking the limit as $n \to \infty$, we get $x_n \geq \mathbb{P}_i(H^A \leq \infty) = h_i^A$.

<u>Definition:</u> $\{k_i^A = \mathbb{E}_i(H^A)\}$

<u>Thm:</u> $(k_i^A : i \in S)$ is the minimal solution to:

$$k_i^A = \begin{cases} 0 & i \in A \\ 1 + \sum_{j\in S}{}' p_{i,j} k_j^A & i \notin A \end{cases}$$

1 move to get to j.

<u>Pf:</u> (of minimal character) Let (y_i^A) be a non-negative solution.

If $i \in A$, $y_i^A = 0 = k_i^A$. Otherwise if $i \notin A$, we have.

$$\begin{aligned} y_i &= 1 + \sum_j{}' p_{i,j} y_j \\ &= 1 + \sum_{j\in A}{}' p_{i,j} y_j + \sum_{j\notin A}{}' p_{i,j} y_j \\ &= 1 + \sum_{j\notin A}{}' p_{i,j} y_j \end{aligned}$$

$$= 1 + \sum_{j \notin A} p_{i,j} \left(1 + \sum_{k \notin A} p_{j,k} y_k\right)$$

$$\geq 1 + \sum_{j \notin A} p_{i,j} = \mathbb{P}_i(H^A \geq 1) + \mathbb{P}_i(H^A \geq 2)$$

By induction, we have that: $y_i \geq \mathbb{P}_i(H^A \geq 1) + \dots + \mathbb{P}_i(H^A \geq n)\ \forall n$.

So: $y_i \geq \sum_{m \geq 1} \mathbb{P}_i(H^A \geq m) = \sum_{m \geq 1} m\, \mathbb{P}_i(H^A = m) = k_i^A$

Example: Birth-death chain. Let $(p_i : i \geq 1)$ be a sequence such that $p_i \in (0,1)$. Let $q_i = 1 - p_i$. Let $\mathbb{N}$ be state space and define: $p_{i,i+1} = p_i$, $p_{i,i-1} = q_i$. This is a general model for population growth. If there are no people, it is impossible to reproduce, so $p_{0,0} = 1$.

Let $A = \{0\}$. Then $h_i^A = h_i^{\{0\}} := h_i$

So, we have: $\quad h_0 = 1, \quad p_i h_{i+1} - h_i + q_i h_{i-1} = 0 \ (i \geq 1)$

We rewrite this as:

$$\begin{aligned} p_i h_{i+1} - h_i + q_i h_{i-1} &= p_i h_{i+1} - (p_i + q_i) h_i + q_i h_{i-1} \\ &= p_i (h_{i+1} - h_i) - q_i (h_i - h_{i-1}) \end{aligned}$$

Let $u_i = h_{i-1} - h_i \ (>0)$. Then we have:

$$u_{i+1} = \frac{q_i}{p_i} u_i = \dots = \left(\frac{q_i}{p_i}\right)\left(\frac{q_{i-1}}{p_{i-1}}\right) \cdots \left(\frac{q_1}{p_1}\right) u_1$$

Let, $\gamma_i = \dfrac{q_i \cdots q_1}{p_i \cdots p_1}$

So: $u_{i+1} = \gamma_i u_1$, and for convenience we define $\gamma_0 = 1$.

Now, summing the equations $u_i = h_{i-1} - h_i$, we get:

$h_0 - h_i = u_1 + \dots + u_i$, $h_0 = 1 \Rightarrow h_i = 1 - u_1(\gamma_0 + \gamma_1 + \dots + \gamma_{i-1})$

Our theorem tells us the value of u_1 minimises h_i, but this depends on $S = \sum_{i=0}^{\infty} \gamma_i$.

i) If S diverges, $u_1 = 0$ since $0 \leq h_i \leq 1$.

ii) If S converges, we take limit as $i \to \infty$, (remembering h_i is non-negative) to see that:

$$0 = 1 - u_1 S \Rightarrow u_1 = \frac{1}{S}$$

So: $h_i = \dfrac{\sum_{k=i}^{\infty} \gamma_k}{\sum_{k=0}^{\infty} \gamma_k}$

• Strong Markov Property + Applications: "Strong" often implies a statement about random variables, whilst "weak" refers to a statement about probabilities. The same is true here.

Definition: {(Stopping Time)} Let X be a Markov Chain. A random variable T (function $\Omega \to \mathbb{N} \cup \{\infty\}$) is a stopping time for the chain $X = (X_n)$ if, for $n \geq 0$, the event $\{T = n\}$ is given in terms of $\{X_0, \dots, X_n\}$.

For example, the hitting time, H^A is a stopping time. This is because, $\{H^A = n\} = \{X_i \notin A \text{ for } i < n\} \cap \{X_n \in A\}$

Thm: (Strong Markov Property): Let X be a Markov Chain with transition matrix P, and let T be a stopping time for X. Given $T < \infty$ and $X_T = i$, the chain $(X_{T+k} ; k \geq 0)$ is a Markov chain with transition matrix P, and initial distribution $X_{T+0} = i$. This chain is independent of $X_0, \dots, X_T$.

Example: (Gambler's Ruin) We want to find the time it takes to get to zero, i.e $H = \inf\{n \geq 0 : X_n = 0\}$. What is the distribution of H?

Define $G_i(s) = \mathbb{E}_i(s^H) = \sum_{n=0}^{\infty} s^n \, \mathbb{P}_i(H = n)$, $|s| < 1$.

We have: $G_1(s) = \mathbb{E}_1(s^H) = p\,\mathbb{E}_1(s^H | X_1 = 2) + q\,\mathbb{E}_1(s^H | X_1 = 0)$

Now, $\mathbb{E}_1(s^H | X_1 = 0) = \mathbb{E}_1(s) = s$ since $H = 1$.

So: $G_1(s) = p\,\mathbb{E}_1(s^H | X_1 = 2) + qs$

Considering $\mathbb{E}_1(s^H | X_1 = 2)$ now, to get to 0, we must pass through 1, so time to get to 0 is time to get from 2 to 1, say H', plus time to get from 1 to 0, say H'' (thus $H = H'' + H' + 1$). By strong Markov property, H', H'' are independent and distributed identically to H. So:

$$G_1(s) = p\,\mathbb{E}_1(s^{H'+H''+1}) + qs = ps\{G_1(s)\}^2 + qs$$

$\Rightarrow G_1(s) = \dfrac{1 \pm \sqrt{1-4pqs^2}}{2ps}$. This says $G_1(s)$ is either $+$ or $-$ but not that this choice is the same $\forall s$. However $G_1(s)$ is a power series $\therefore$ continuous, so one choice will be sufficient.

At $s=0$, denominator is zero $\therefore$ since G_1 converges $\Rightarrow$ numerator $=0$ so $-$ must be correct.

$$\therefore \quad G_1(s) = \frac{1-\sqrt{1-4pqs^2}}{2ps}$$

Then, the probability of ever hitting zero is:

$$\mathbb{P}_1(H<\infty) = \sum_{n=1}^{\infty} \mathbb{P}(H=n) = \lim_{s\to 1} G_1(s) = \frac{1-\sqrt{\cancel{4}\,1-4pq}}{2p}$$

We note $1-4pq = |q-p|^2$

$$\text{So: } \mathbb{P}_1(H<\infty) = \frac{1-|p-q|}{2p} = \begin{cases} 1 & \text{if } q \geq p \\ \frac{q}{p} & \text{if } p>q \end{cases}$$

- Further Classification of States

Thm: Suppose $X_0 = i$. Let $V_i = |\{n \geq 1 : X_n = i\}|$. Let $f_{i,i} = \mathbb{P}_i(T_i < \infty)$. Then: $\mathbb{P}(V_i = r) = f_{i,i}^{\,r}(1-f_{i,i})$. Proof by strong Markov property.

So, we see if $f_{i,i} = 1 \Rightarrow \mathbb{P}(V_i = r) = 0 \quad \forall r \Rightarrow \mathbb{P}(V_i = \infty) = 1$.

Otherwise $\mathbb{P}(V_i = \infty) = 0$

Definition: (Mean Recurrence Time) Let T_i be returning time to a state i. Then mean recurrence time, μ_i, is:

$$\mu_i = \begin{cases} \infty & \text{if } i \text{ transient} \\ \sum_{n=1}^{\infty} n f_{i,i}(n) & \text{if } i \text{ recurrent.} \end{cases}$$

non-null / positive state

SWAP

null-state

Definition: (Period) $d_i = \gcd\{n \geq 1 : p_{i,i}(n) > 0\}$ A state is aperiodic if $d_i = 1$.

Definition: A state is ergodic if it is aperiodic + positive recurrent.

These are states we like to deal with.

Thm: If $i \leftrightarrow j$, then:

1) $d_i = d_j$
2) i is recurrent iff j is recurrent.
3) i is positive recurrent iff j is positive recurrent
4) i is ergodic iff j is ergodic

Pf: 1) Assume $i \leftrightarrow j$, then $\exists m, n \geq 1 : p_{i,j}(m), p_{j,i}(n) > 0$. By Chapman-Kolmogorov equations, we have:

$$p_{i,i}(m+r+n) \geq p_{i,j}(m)\, p_{j,j}(r)\, p_{j,i}(n) \geq \alpha\, p_{j,j}(r)$$

where $\alpha = p_{i,j}(m)\, p_{j,i}(n)$

Let $D_j = \{r \geq 1 : p_{j,j}(r) > 0\}$. Then $d_j = \gcd\{D_j\}$

We have just shown that if $r \in D_j$, $m+r+n \in D_i$. We also have that $n+m \in D_i$ since $p_{i,i}(n+m) \geq p_{i,j}(m)\, p_{j,i}(n) > 0$. So, for any $r \in D_j$, $d_i | m+r+n$ and $d_i | m+n$. So $d_i | r$.

Hence $\gcd\{D_i\} | \gcd\{D_j\}$. By symmetry $\gcd\{D_j\} | \gcd\{D_i\}$.

So $\gcd\{D_i\} = d_i = d_j = \gcd\{D_j\}$.

We prove 2), 3), 4) later.

Proposition: If the chain is irreducible and $j \in S$ is recurrent, then:

$$\mathbb{P}(X_n = j \text{ for some } n \geq 1) = 1.$$

regardless of the initial distribution of X_0.

Pf: Let $f_{i,j} = \mathbb{P}_i(X_n = j \text{ for some } n \geq 1)$

Since $j \to i$, $\exists$ least integer $m \geq 1 : p_{j,i}(m) > 0$. Since m is least we know that:

$$p_{j,i}(m) = \mathbb{P}_j(X_m = i, X_r \neq j \text{ for } r < m).$$

Then, $p_{j,i}(m)\{1 - f_{i,j}\} \leq 1 - f_{j,j}$

Since LHS is probability we first go from j to i in m steps and then never going from i to j again. Whilst right is just probability of never returning to j when starting in j. Hence if $f_{j,j} = 1$, then $f_{i,j} = 1$.

Now, let $\lambda_k = \mathbb{P}(X_0 = k)$ be our initial distribution:

Then $\mathbb{P}(X_n = j \text{ for some } n \geq 1) = \sum_i \lambda_i \mathbb{P}_i(X_n = j \text{ for some } n \geq 1) = 1$.

{Section 3}

• <u>Invariant Distributions</u>: Convergence is considered in terms of probability mass functions. i.e we ask if, for $k \in S$, $\mathbb{P}(X_n = k)$ converges to anything. and whether this is itself a distribution.

We have the following trivial identity for the evolution of a chain with initial distribution λ and transition matrix P:

$$\lambda P^{n+1} = (\lambda P^n) P$$

In the limit, if the distribution converges, then $\lambda P^n, \lambda P^{n+1} \to \pi$.
So, π satisfies: $\{\pi P = \pi\}$ invariant distributions.

<u>Definition:</u> Let (X_n) be a Markov chain with transition probabilities, P. Then, the distribution $\pi = (\pi_k, k \in S)$ is an invariant distribution if:

$$\left\{\begin{array}{l} 1)\ \pi_k \geq 0,\quad \sum_k \pi_k = 1 \\ 2)\ \pi = \pi P. \end{array}\right\}$$

<u>Thm</u>: (*) Consider an irreducible MC. Then:
1) $\exists$ an invariant distribution iff some state is positive recurrent
2) If there is an invariant distribution π, then every state is positive recurrent and:

$$\pi_i = \frac{1}{\mu_i} \leftarrow \text{mean recurrence time.}$$

and this π is unique.

N.B this proves $i \leftrightarrow j$ then i positive recurrent $\leftrightarrow j$ positive recurrent.

We note that it appears sensible that $\pi_i \propto \frac{1}{\mu_i}$ since for small mean recurrence time, π_i increases.

We consider fixing a state k, and looking at the number of times i is visited before returning to k. Let $X_0 = k$ and let W_i be number of visits to i before next visit to k. So:

$$W_i = \sum_{m=1}^{\infty} \mathbb{I}(X_m = i, T_k \geq m)$$

which is the same as: $W_i = \sum_{m=1}^{T_k} \mathbb{I}(X_m = i)$ which is an random variable

Define $e_i = \mathbb{E}_k(w_i)$ which we shall see is almost π_i.

Proposition: For an irreducible recurrent chain and $k \in S$, $e = (e_i : i \in S)$ defined as above by:

$$e_i = \mathbb{E}_k(w_i)\ , \quad w_i = \sum_{m=1}^{T_k} \mathbb{I}(X_m = i)$$

we have:

1) $e_k = 1$
2) $\sum_i e_i = \mu_k$
3) $e = eP$
4) $0 < e_i < \infty \ \forall i \in S$.

Pf: 1) Follows from definition since for $m < T_k$, $X_m \neq k$.

2) Note, $\sum_i w_i = T_k$

We have:

$$\begin{aligned}\sum_i e_i &= \sum_i \mathbb{E}_k(w_i) \\ &= \mathbb{E}_k\Big(\sum_i w_i\Big) \\ &= \mathbb{E}_k(T_k) \\ &= \mu_k .\end{aligned}$$

3)
$$\begin{aligned} e_j &= \mathbb{E}_k(w_j) \\ &= \mathbb{E}_k\Big(\sum_{m\geq 1} \mathbb{I}(X_m = j, T_k \geq m)\Big) \\ &= \sum_{m \geq 1} \mathbb{P}_k(X_m = j, T_k \geq m) \\ &= \sum_{m\geq 1} \sum_{i \in S} \mathbb{P}_k(X_m = j \mid X_{m-1} = i, T_k \geq m)\, \mathbb{P}_k(X_{m-1} = i, T_k \geq m)\end{aligned}$$

Now, using the Markov property, knowing $T_k \geq m$ says $X_0, \ldots, X_{m-1}$ are all not k. So:

$$\begin{aligned} &= \sum_{m\geq 1} \sum_{i \in S} \mathbb{P}_k(X_m = j \mid X_{m-1} = i)\, \mathbb{P}_k(X_{m-1} = i, T_k \geq m) \\ &= \sum_{m \geq 1} \sum_{i \in S} p_{i,j}\, \mathbb{P}_k(X_{m-1} = i, T_k \geq m) \\ &= \sum_{i \in S} p_{i,j} \sum_{m \geq 1} \mathbb{P}_k(X_{m-1} = i, T_k \geq m)\end{aligned}$$

The last term looks very similar to e_i. To show it is in fact e_i, first let $r = m-1$, so

$$\sum_{m \geq 1}' \mathbb{P}_k(X_{m-1} = i, T_k \geq m) = \sum_{r=0}^{\infty}' \mathbb{P}_k(X_r = i, T_k \geq r+1).$$

i) If $i = k$, then $r = 0$ term is 1. since $T_k \geq 1$ is always true. and $X_0 = k$ by construction. So the sum is 1, which is indeed e_k.

ii) In the case $i \neq k$. note $r = 0 \Rightarrow X_0 = k \neq i$. So this term is zero. For $r \geq 1$, we know that if $X_r = i$ and $T_k \geq r$, then we must have $T_k \geq r+1$ since it is not possible for return time to be exactly r if we are at $i \neq k$ at time r. So,
$\mathbb{P}_k(X_r = i, T_k \geq r+1) = \mathbb{P}_k(X_r = i, T_k \geq r)$

So indeed, $\sum_{n \geq 0}' \mathbb{P}_k(X_{m-1} = i, T_k \geq m) = e_i$

So, $e_j = \sum_{i \in S} p_{i,j} e_i$.

4) We show $0 < e_i < \infty$, first fix i and note that $e_k = 1$. We know $e = eP = eP^n$ for $n \geq 1$. So by expanding the matrix sum we know for any m, n:

$e_i \geq e_k p_{k,i}(n)$

$e_k \geq e_i p_{i,k}(m)$

By irreducibility, we now choose n, m: $p_{i,k}(m), p_{k,i}(n) > 0$. So:

$$e_k p_{k,i}(n) \leq e_i \leq \frac{e_k}{p_{i,k}(m)} \Rightarrow 0 < p_{k,i}(n) \leq e_i \leq \frac{1}{p_{i,k}(m)} < \infty$$

We may now prove the theorem (*)

Pf: i) Let k be a recurrent state, then $\pi_i = \frac{e_i}{\mu_k}$ satisfies $\pi_i \geq 0$ and $\sum_i \pi_i = 1$. and is an invariant distribution.

ii) Let π be an invariant distribution, first show all entries are non-zero. $\forall n$ we have:

$\pi = \pi P^n$. Hence, for all $i, j \in S$, $n \in \mathbb{N}$, we have:

$\pi_i \geq \pi_j p_{j,i}(n)$.

Since $\sum_i \pi_i = 1$, $\exists k : \pi_k > 0$.

With $j=k$, we have $\pi_i \geq \pi_i p_{k,i}(n) > 0$ for some n by irreducibility. So $\pi_i > 0 \ \forall i \in S$.

Now, we show all states are positive recurrent.

Assume all states are transient. So $p_{j,i}(n) \to 0 \ \forall i,j \in S$, $n \in \mathbb{N}$. However, we have:

$$\pi_i = \sum_j \pi_j p_{j,i}(n)$$

Consider state space S, split into finite subset F and $S \setminus F$.

Then:
$$\begin{aligned} 0 &\leq \sum_j \pi_j p_{j,i}(n) \\ &= \sum_{j \in F} \pi_j p_{j,i}(n) + \sum_{j \notin F} \pi_j p_{j,i}(n) \\ &\leq \sum_{j \in F} p_{j,i}(n) + \sum_{j \notin F} \pi_j \to \sum_{j \notin F} \pi_j \end{aligned}$$

Now, $\sum_{i \in S} \pi_i = 1$, so $\sum_{j \notin F} \pi_j \to 0$ as $F \to S$. So $\pi_i \to 0$ which is a contradiction.

Ruling out null recurrence requires us to prove μ_i is finite. By definition $\mu_i = \mathbb{E}_i(T_i)$ and we have formula $\mathbb{E}(N) = \sum_n \mathbb{P}(N \geq n)$

So: $\pi_i \mu_i = \sum_{n=1}^{\infty} \pi_i \mathbb{P}_i(T_i \geq n)$

Let π_i be probability of $X_0 = i$ ($\mathbb{P}(X_0 = i)$). Then:

$$\pi_i \mu_i = \sum_{n=1}^{\infty} \mathbb{P}(T_i \geq n, X_0 = i)$$

Now, $\mathbb{P}(T_i \geq 1, X_0 = i) = \mathbb{P}(X_0 = i) = \pi_i$

Also,
$$\begin{aligned} \mathbb{P}(T_i \geq n, X_0 = i) &= \mathbb{P}(X_0 = i, X_m \neq i \text{ for } 1 \leq m \leq n-1) \\ &= \mathbb{P}(X_m \neq i \text{ for } 1 \leq m \leq n-1) - \mathbb{P}(X_m \neq i \text{ for } 0 \leq m \leq n-1) \end{aligned}$$

Now, by invariance $(X_0, \dots, X_{n-2})$ has same distribution as $(X_1, \dots, X_{n-1})$. So:

$$\mathbb{P}(T_i \geq n, X_0 = i) = a_{n-2} - a_{n-1}$$

where $a_r = \mathbb{P}(X_m \neq i, \text{ for } 0 \leq m \leq r)$

So: $\pi_i \mu_i = \pi_i + (a_0 - a_1) + (a_1 - a_2) + \dots$

Then:

$$\pi_i \mu_i = \lim_{N \to \infty} \{\pi_i + a_0 + \dots + (a_{N-2} - a_{N-1})\} = \pi_i + a_0 + \lim_{N \to \infty} a_N$$

$\lim_{N\to\infty} a_N = \mathbb{P}(X_m \neq i \text{ for all } m) = 0$ since state is recurrent.

Also, $a_0 = \mathbb{P}(X_0 \neq i)$, $\pi_i = \mathbb{P}(X_0 = i)$ $\therefore \pi_i + a_0 = 1$

$\Rightarrow \pi_i \mu_i = 1 \Rightarrow \boxed{\pi_i = \frac{1}{\mu_i}}$ and implies that μ_i finite as reqd.

- Convergence to Equilibrium: We have proved a chain has a (unique) invariant distribution iff it is positive recurrent.

Thm: Consider an irreducible MC that is positive recurrent and aperiodic. Then:

$$p_{i,k}(n) \longrightarrow \pi_k$$

as $n \to \infty$.

(Pf is non-examinable, see online notes for coupling ideas).

{Section 4}

- Time Reversal: Suppose we have a Markov Chain $X = (X_0, \dots, X_N)$.

Define a new chain, $Y_k = X_{N-k}$, so $Y = (X_N, \dots, X_0)$. When is Y a Markov chain?

Thm: Let X be positive recurrent, irreducible with invariant distribution π. Suppose X_0 has distribution π. Then Y defined by $Y_k = X_{N-k}$ is a Markov chain with transition matrix $\tilde{P}$ defined by:

$$\tilde{p}_{i,j} = \left(\frac{\pi_j}{\pi_i}\right) p_{j,i} \qquad \text{(just transpose with normalisation)}.$$

Also, π is invariant for $\tilde{P}$.

Pf: First show $\tilde{P}$ is a stochastic matrix.

Clearly $\tilde{p}_{i,j} \geq 0$. Also, for each i:

$$\sum_j{}' \tilde{p}_{i,j} = \frac{1}{\pi_i}\sum_j{}' \pi_j p_{j,i} = \frac{1}{\pi_i} \pi_i = 1.$$

Now show π is invariant for $\tilde{P}$. We have:

$$\sum_i{}' \pi_i \tilde{p}_{i,j} = \sum_i \pi_j p_{j,i} = \pi_j \Big(\sum_i{}' p_{j,i}\Big) = \pi_j \cdot 1 = \pi_j$$

Now, we show that Y is a Markov Chain.

$$\mathbb{P}(Y_0 = i_0, \dots, Y_k = i_k) = \mathbb{P}(X_{N-k} = i_k, \dots, X_N = i_0)$$

$$= \pi_{i_k} p_{i_k, i_{k-1}} \cdots p_{i_1, i_0}$$
$$= (\pi_{i_k} p_{i_k, i_{k-1}}) p_{i_{k-1}, i_{k-2}} \cdots p_{i_1, i_0}$$
$$= \tilde{p}_{i_k, i_{k-1}} (\pi_{i_{k-1}} p_{i_{k-1}, i_{k-2}}) \cdots p_{i_1, i_0}$$
$$= \pi_{i_0} \tilde{p}_{i_0, i_1} \cdots \tilde{p}_{i_{k-1}, i_k}$$

So Y is Markov.

<u>Definition</u>: (Reversible Chain) An irreducible MC, $X = (X_0, \dots, X_N)$ in its invariant distribution, π, is reversible if its reversal has the same transition probabilities as X. i.e

$$\{\pi_i p_{i,j} = \pi_j p_{j,i}\} \quad \text{for all } i,j \in S.$$

Known as {detailed balance equation}. In general if λ is a distribution that satisfies

$$\{\lambda_i p_{i,j} = \lambda_j p_{j,i}\}$$

we say (P, λ) is in detailed balance.

<u>Proposition</u>: Suppose (P, λ) is in detailed balance. Then λ is the (unique) invariant distribution. and the chain is reversible (when X_0 has distribution λ).

Gives us better criterion: we need only solve $\lambda_i p_{i,j} = \lambda_j p_{j,i}$ instead of $\pi_i = \sum_j \pi_j p_{j,i}$.

Pf: $\sum_j \lambda_j p_{j,i} = \sum_j \lambda_i p_{i,j} = \lambda_i \sum_j p_{i,j} = \lambda_i$

So λ is invariant.

<u>Example</u>: (Random walk on finite graph). A graph is a pair $(V,E) = G$ where E contains distinct unordered pairs of distinct vertices (u,v) joined by an edge. (no loops) (no parallel edges).

A graph is <u>connected</u> if $\forall u,v \in V$, there exists a path along the edges from u to v.

Let $G = (V, E)$ be a connected graph with $|V| \leq \infty$. Let $X = (X_n)$ be a random walk on G. i.e if $X_n = x$ then X_{n+1} is chosen uniformly from the neighbours of x i.e $\{y \in V : (x,y) \in E\}$, independently of the past. This is a Markov Chain.

$$P_{i,j} = \begin{cases} 0 & \text{if } j \text{ is not a neighbour of } i. \\ \frac{1}{d_i} & \text{if } j \text{ is a neighbour of } i. \end{cases}$$

where d_i is the (degree of i) no. of neighbours of i.

By connectivity, the MC is irreducible and since it is finite, it is recurrent. In fact it is positive recurrent.

We try and solve the detailed balance equations:

$$\lambda_i P_{i,j} = \lambda_j P_{j,i}$$

$$\Rightarrow \quad \lambda_i \cdot \frac{1}{d_i} = \frac{\lambda_j}{d_j}$$

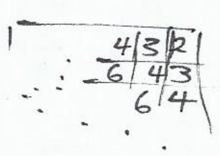
So, let $\lambda_i = c d_i$ then make λ a distribution:

$$1 = \sum_i \lambda_i = c \sum_i d_i \Rightarrow c = \frac{1}{\sum_i d_i}$$

But $\sum_i d_i$ is just $2|E|$. So $c = \frac{1}{2|E|}$

So: $\left\{ \lambda_i = \frac{d_i}{2|E|} \right\}$

e.g May apply this to knight on chessboard. with d_i = no. of possible moves at each square. We see $\sum_i d_i = 336$



So, for example, $\left[\pi_{\text{corner}} = \frac{2}{336}\right]$