Markov Chains (Michaelmas 2019)

Primary references:

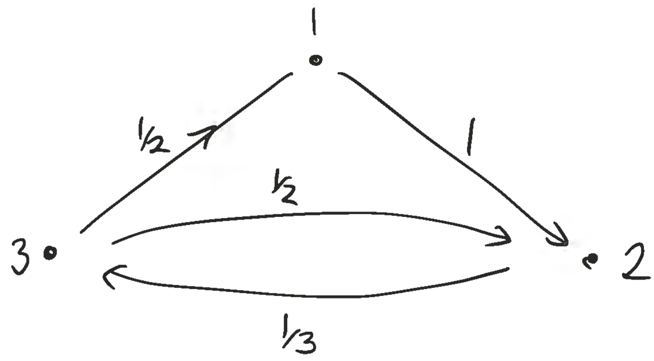
J. Norris, Markov Chains, Cambridge University Press
G. Grimmett and D. Welsh, Probability, An Introduction, Oxford University Press

Please report errors and comments to Roland Bauerschmidt (rb812@cam.ac.uk).

November 19, 2019

0. Examples

(i) States $I = \{1, 2, 3\}$

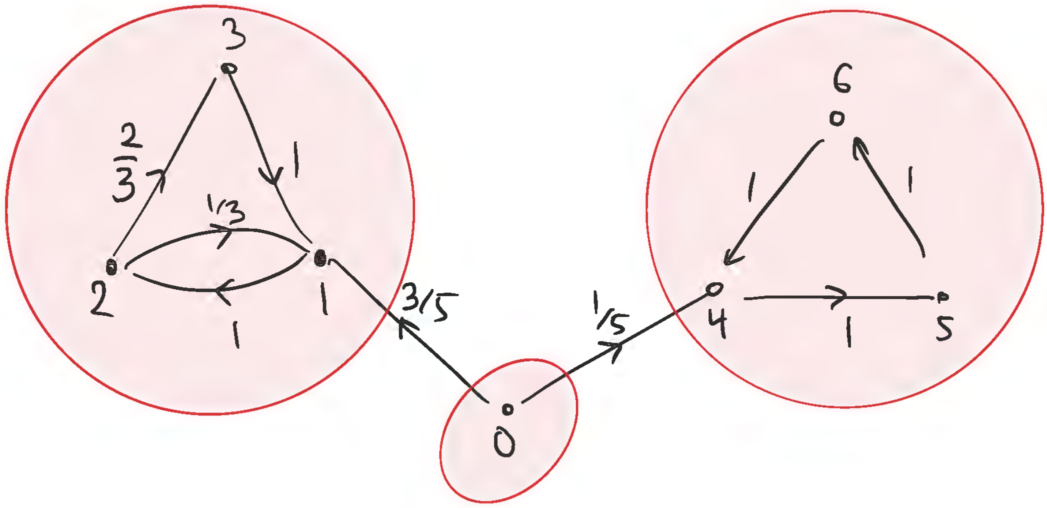


diagram

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

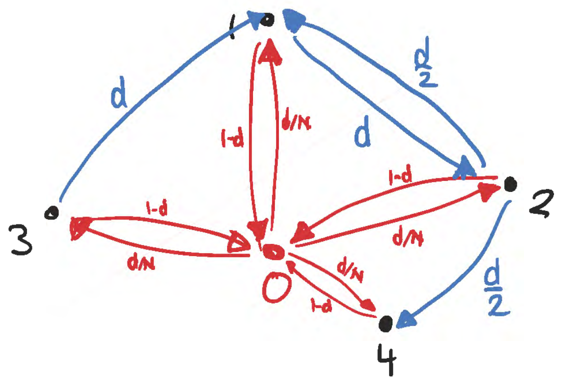
transition matrix

(ii)



- Starting from 0, the probability to hit 6 is 1/4.
- Starting from 1, the probability to hit 3 is 1.
- Starting from 1, it takes on average 3 steps to hit 3.
- Starting from 1, the long-run proportion of time spent in 2 is 3/8.
- ...

(iii) (Page Rank) Consider websites $1, \dots, N$ and a restart state 0.



$(0<d<1)$
parameter

- Blue arrows represent links between websites. The probability associated to every blue arrow is
$$\frac{d}{\#\text{ of outgoing links from site}}$$
The probility to jump to the restart state is $1-d$.

- From the restart state 0, jumps happen to all sites with equal probability d/N.

By a theorem we will prove, the probability to be at a given site $i \in \{1, \dots, N\}$ (or the restart state 0) converges as the number of steps tends to infinity (indep. of start vertex):

$$\mathbb{P}[X_n = i] \longrightarrow \pi_i .$$

The Page Rank is defined as $PR(i) = \frac{N}{d}\pi_i$. It was initially used by Google as a measure how important a webpage is.

1. Definitions and basic properties

We will make the following standing assumptions:

- I is a countable set, the state space; often $I = \{1, 2, \dots\}$.
- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space on which all relevant random variables are defined.

Defn. A sequence of random variables $(X_n)_{n=0,1,\dots}$ is a Markov Chain if, for all $n \geq 0$ and $i_0, \dots, i_{n+1} \in I$,

$$\mathbb{P}[X_{n+1} = i_{n+1} \mid X_0 = i_0, \dots, X_n = i_n] = \mathbb{P}[X_{n+1} = i_{n+1} \mid X_n = i_n].$$

It is homogeneous if, for all $i, j \in I$,

$\mathbb{P}[X_{n+1} = j \mid X_n = i]$ does not depend on n.

From now on, we will assume that all Markov Chains are homogeneous. Then a Markov Chain is characterised by the following data:

(a) the initial distribution $\lambda = (\lambda_i)_{i \in I}$ given by $\lambda_i = \mathbb{P}[X_0 = i]$;

(b) the transition matrix $P = (p_{ij})_{i,j \in I}$ given by $p_{ij} = \mathbb{P}[X_1 = j \mid X_0 = i]$.

Fact.

- λ is a distribution, i.e., $\lambda_i \geq 0$ for all i and $\sum_{i \in I} \lambda_i = 1$.
- P is a stochastic matrix, i.e., $(P_{ij})_j$ is a distribution for every $j \in I$.

Defn. (X_n) is a Markov Chain with initial distribution λ and transition probability P, or (X_n) is Markov(λ, P) for short, if (a) and (b) hold.

Thm. (X_n) is Markov(λ,P) iff for all $n \geq 0$, $i_0, \ldots, i_n \in I$,

$$\mathbb{P}[X_0 = i_0, \ldots, X_n = i_n] = \lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}. \qquad (*)$$

Proof. Suppose (X_n) is Markov(λ, P). Then

$$\begin{aligned}\mathbb{P}[X_0 = i_0, \ldots, X_n = i_n] &= \mathbb{P}[X_n = i_n \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}] \\ &\qquad \times \mathbb{P}[X_0 = i_0, \ldots, X_{n-1} = i_{n-1}] \\ &\overset{\text{(Markov)}}{=} \underbrace{\mathbb{P}[X_n = i_n \mid X_{n-1} = i_{n-1}]}_{p_{i_{n-1} i_n}} \mathbb{P}[X_0 = i_0, \ldots, X_{n-1} = i_{n-1}] \\ &\overset{\text{(Induction)}}{=} \cdots = p_{i_{n-1} i_n} p_{i_{n-2} i_{n-1}} \cdots p_{i_0 i_1} \underbrace{\mathbb{P}[X_0 = i_0]}_{\lambda_{i_0}}.\end{aligned}$$

Conversely, assume $(*)$ holds for all n and $i_0, \ldots, i_n$. Summing over i_n, then i_{n-1}, ..., then i_1 gives

$$\mathbb{P}[X_0 = i_0] = \lambda_{i_0} \underbrace{\sum_{i_1} p_{i_0 i_1}}_{1} \underbrace{\sum_{i_2} p_{i_1 i_2}}_{1} \cdots \underbrace{\sum_{i_n} p_{i_{n-1} i_n}}_{1} = \lambda_{i_0}$$

Also, $(*)$ gives

$$\mathbb{P}[X_n = i_n \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}] = \frac{\mathbb{P}[X_0 = i_0, \ldots, X_n = i_n]}{\mathbb{P}[X_0 = i_0, \ldots, X_{n-1} = i_{n-1}]} = p_{i_{n-1} i_n}.$$

Thus we have shown that (a) and (b) hold, i.e., (X_n) is Markov(λ, P).

Let $\delta_i = (\delta_{ij} : j \in I)$ be the unit mass at $i \in I$:

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Thm (Markov property). Let (X_n) be Markov(λ, P). Then, conditional on $X_m = i$, $(X_{m+n})_{n \geq 0}$ is Markov(δ_i, P) and is independent of $X_0, \ldots, X_m$.

Proof. Let A be any event determined by $X_0, \ldots, X_m$. It suffices to show that then

$$\mathbb{P}[\{X_m = i_m, \ldots, X_{m+n} = i_{m+n}\} \cap A \mid X_m = i]$$

$$= \delta_{i i_m} P_{i_m i_{m+1}} \cdots P_{i_{m+n-1} i_{m+n}} \mathbb{P}[A \mid X_m = i]$$

$\Leftrightarrow$ $\mathbb{P}[\{X_m = i_m, \ldots, X_{m+n} = i_{m+n}\} \cap A] \, \delta_{i i_m}$

$\mathbb{P}[X_m = i] \neq 0$ $\quad = \delta_{i i_m} P_{i_m i_{m+1}} \cdots P_{i_{m+n-1} i_{m+n}} \mathbb{P}[A \cap \{X_m = i\}]$.

For an elementary event $A = \{X_0 = i_0, \ldots, X_m = i_m\}$ this follows from the previous theorem.

As any event A can be written as a countable disjoint union of elementary events,

$$A = \bigcup_{k=1}^{\infty} A_k,$$

the claim follows by summing the identities for A_k.

Notation:

We regard distributions and measures $(\lambda_i)_{i\in I}$ as row vectors.

Here $(\lambda_i)_{i\in I}$ is a measure if $\lambda_i \geq 0$ for all i.

It is called a distribution if in addition $\sum_{i\in I}\lambda_i = 1$.

Matrix multiplication:

$$(\lambda P)_j = \sum_{i\in I}\lambda_i p_{ij}, \quad (P^2)_{ij} = \overbrace{\sum_{k\in I} p_{ik}p_{kj}}^{p_{ij}^{(2)}}, \quad (P^3)_{ij} = \cdots$$

with $P^0 = \mathbb{1}$ the $I\times I$ identity matrix $\mathbb{1}_{ij} = \delta_{ij}$.

When $\lambda_i > 0$, write $\mathbb{P}_i[A] = \mathbb{P}[A \mid X_0 = i]$.

Fact. By the Markov property, $(X_n)_{n\geq 0}$ is Markov(δ_i, P) under $\mathbb{P}_i$. So the behaviour of (X_n) under $\mathbb{P}_i$ does not depend on λ.

Thm. Let $(X_n)_{n\geq 0}$ be Markov(λ, P). Then for all $n, m \geq 0$,

(a) $\mathbb{P}[X_n = j] = (\lambda P^n)_j$

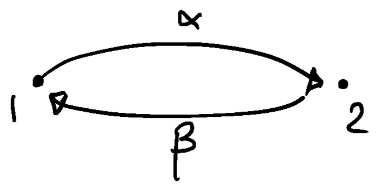
(b) $\mathbb{P}_i[X_n = j] = \mathbb{P}[X_{n+m} = j \mid X_m = i] = p_{ij}^{(n)}$.

Proof. (a)
$$\mathbb{P}[X_n = j] = \sum_{i_0,\dots,i_{n-1}\in I}\mathbb{P}[X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = j]$$
$$= \sum_{i_0,\dots,i_{n-1}}\lambda_{i_0}p_{i_0 i_1}\cdots p_{i_{n-1}j} = (\lambda P^n)_j$$

(b) Use the Markov property and $\lambda = \delta_i$ in (a).

Example. The general two state Markov Chain is given by

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$



for some $\alpha, \beta \in [0,1]$.

$$P^{n+1} = P^n P \Rightarrow p_{11}^{(n+1)} = p_{12}^{(n)}\beta + p_{11}^{(n)}(1-\alpha)$$

$$p_{12}^{(n)} + p_{11}^{(n)} = 1 \Rightarrow p_{11}^{(n+1)} = (1-p_{11}^{(n)})\beta + p_{11}^{(n)}(1-\alpha)$$

$$= p_{11}^{(n)}(1-\alpha-\beta) + \beta$$

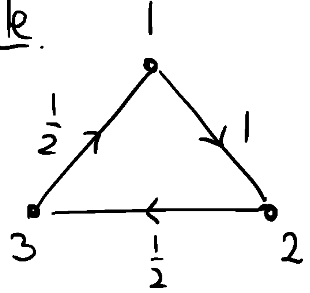
Since $p_{11}^{(0)} = 1$, this recursion relation has unique solution

$$p_{11}^{(n)} = \begin{cases} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \text{if } \alpha+\beta > 0 \\ 1 & \text{if } \alpha+\beta = 0. \end{cases}$$

[Indeed, the formula is true for $n=0$. Assuming it holds for some n, we find $p_{11}^{(n+1)} = \underbrace{\frac{\beta}{\alpha+\beta}(1-\alpha-\beta) + \beta}_{\frac{\beta}{\alpha+\beta}} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^{n+1}$.]

In general, one cannot find a closed equation for $p_{11}^{(n)}$.

Example.



$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

What is $p_{11}^{(n)}$?

General method to find $p_{ij}^{(n)}$ for an N state Markov Chain.

- Find the eigenvalues $\lambda_1, \dots, \lambda_N$ of P, i.e., roots of
$$\det(\lambda - P) = 0.$$
- If all eigenvalues are distinct, then $p_{ij}^{(n)}$ has the form
$$p_{ij}^{(n)} = a_1\lambda_1^n + \cdots + a_n\lambda_N^n \quad \text{for some constants } a_i.$$
If an eigenvalue λ is repeated once then the general form includes a term $(a+bn)\lambda^n$. Similar formulas hold for eigenvalues with higher multiplicity.
- As roots of a polynomial with real coefficients, any complex eigenvalues come in conjugate pairs which are best written in terms of sin and cos.

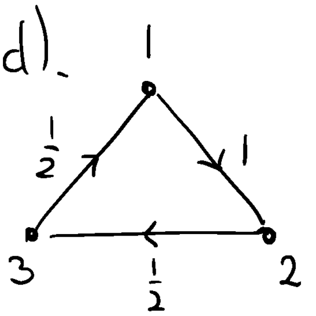
Justification. if P has distinct eigenvalues, it may be diagonalised as

$$P = U\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}U^{-1} \Rightarrow P^n = U\begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_N^n \end{pmatrix}U^{-1}$$

so $p_{ij}^{(n)}$ is of the desired form.

If P has repeated eigenvalues, the more general claim follows from the Jordan normal form of P.

Example (cont'd).



$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

What is $p_{11}^{(n)}$?

Eigenvalues of P: $0 = \det(\lambda - P)$

$$= \lambda(\lambda - \tfrac{1}{2})^2 - \tfrac{1}{4} = \tfrac{1}{4}(\lambda - 1)(4\lambda^2 + 1)$$

$$\Rightarrow \lambda = 1, \tfrac{i}{2}, -\tfrac{i}{2}.$$

$\Rightarrow p_{11}^{(n)} = a + b(\tfrac{i}{2})^n + c(-\tfrac{i}{2})^n$ for some constants a, b, c.

$$(\pm\tfrac{i}{2})^n = (\tfrac{1}{2})^n e^{\pm i\pi n/2} = (\tfrac{1}{2})^n \left(\cos(\tfrac{1}{2}\pi n) \pm i \sin(\tfrac{1}{2}\pi n)\right)$$

$\Rightarrow p_{11}^{(n)} = \alpha + (\tfrac{1}{2})^n \left[\beta \cos(\tfrac{1}{2}\pi n) + \gamma \sin(\tfrac{1}{2}\pi n)\right]$ for some α, β, γ.

Note:
$$\left.\begin{aligned} 1 &= p_{11}^{(0)} = \alpha + \beta \\ 0 &= p_{11}^{(1)} = \alpha + \tfrac{1}{2}\gamma \\ 0 &= p_{11}^{(2)} = \alpha - \tfrac{1}{4}\beta \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} \alpha &= \tfrac{1}{5} \\ \beta &= \tfrac{4}{5} \\ \gamma &= -\tfrac{2}{5} \end{aligned}\right.$$

Thus: $p_{11}^{(n)} = \tfrac{1}{5} + (\tfrac{1}{2})^n \left[\tfrac{4}{5}\cos(\tfrac{1}{2}\pi n) - \tfrac{2}{5}\sin(\tfrac{1}{2}\pi n)\right]$

2. Class structure

Defn. For $i,j \in I$,

- i leads to j, written $i \to j$, if $\mathbb{P}_i[X_n = j \text{ for some } n] > 0$;
- i communicates with j, written $i \leftrightarrow j$, if $i \to j$ and $j \to i$.

Thm. For $i \neq j$, the following are equivalent:

(a) $i \to j$;

(b) $p_{i_1 i_2} \cdots p_{i_{n-1} i_n} > 0$ for some $i_1, \ldots, i_n$ with $i_1 = i$ and $i_n = j$;

(c) $p_{ij}^{(n)} > 0$ for some n.

Proof. Equivalence of (a) and (c) follows from

$$p_{ij}^{(n)} = \mathbb{P}_i[X_n = j] \leq \mathbb{P}_i[X_n = j \text{ for some } n] \leq \sum_{n=0}^{\infty} p_{ij}^{(n)}.$$

Equivalence of (a) and (b) follows from

$$p_{ij}^{(n)} = \sum_{i_2 \cdots i_{n-1}} p_{i i_2} \cdots p_{i_{n-1} j}.$$

Prop. The relation $\leftrightarrow$ is an equivalence relation.

Proof. We must show that $\leftrightarrow$ is reflexive, symmetric, and transitive. That $\leftrightarrow$ is reflexive ($i \leftrightarrow i$) and symmetric ($i \leftrightarrow j$ implies $j \leftrightarrow i$) is clear. That $\leftrightarrow$ is transitive ($i \leftrightarrow j$ and $j \leftrightarrow k$ implies $i \leftrightarrow k$) follows from (b) of the previous theorem.

Defn. The equivalence classes of $\leftrightarrow$ are called <u>communicating classes</u>. The chain is <u>irreducible</u> if there is a single communicating class, i.e., $i \leftrightarrow j$ for all $i, j \in I$.

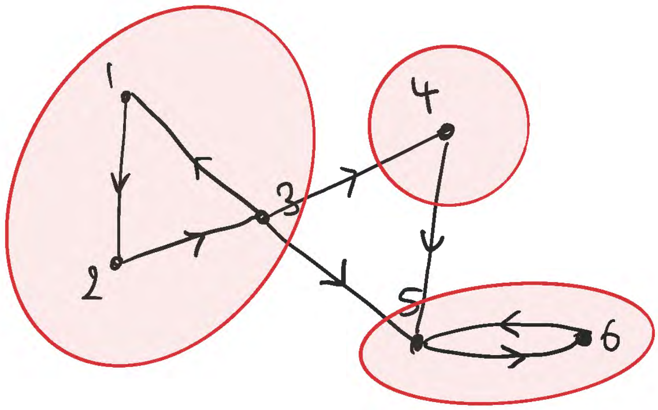
Defn. A subset $C \subset I$ is <u>closed</u> if

$$i \in C,\ i \to j \ \Rightarrow\ j \in C.$$

A state $i \in I$ is <u>absorbing</u> if $\{i\}$ is closed.

<u>Example.</u>

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$



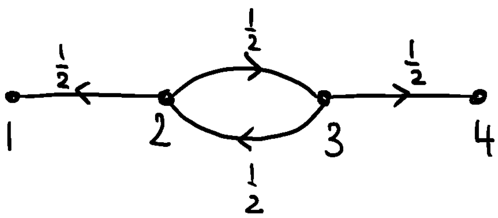
The communicating classes are $\{1,2,3\}$, $\{4\}$, $\{5,6\}$.
Only $\{5,6\}$ is closed.

3. Hitting times and absorption probabilities

Defn. Let (X_n) be a Markov Chain.

- The hitting time of a set $A \subset I$ is the random variable $H^A : \Omega \to \{0,1,2,\ldots\} \cup \{+\infty\}$ given by
$$H^A(\omega) = \inf\{n \geq 0 : X_n(\omega) \in A\}, \quad \inf \emptyset = +\infty.$$
- The hitting probability of A is
$$h_i^A = \mathbb{P}_i[H^A < \infty] = \mathbb{P}_i[\text{hit } A].$$
If A is a closed class, h_i^A is called the absorption time.
- The mean hitting time is the expected time to reach A:
$$k_i^A = \mathbb{E}_i[H^A] = \mathbb{E}_i[\text{time to hit } A].$$

Example.



Starting from 2, what is the probability of absorbtion in 4, and how long does it take until the chain is absorbed in 1 or 4?

Let $h_i = \mathbb{P}_i[\text{hit } 4]$ and $k_i = \mathbb{E}_i[\text{time to hit } \{1,4\}]$.

Note that
$$h_1 = 0, \qquad k_1 = 0,$$
$$h_4 = 1, \qquad k_4 = 0,$$
$$h_2 = \tfrac{1}{2}h_1 + \tfrac{1}{2}h_3, \qquad k_2 = 1 + \tfrac{1}{2}k_1 + \tfrac{1}{2}k_3,$$
count time of first step
$$h_3 = \tfrac{1}{2}h_2 + \tfrac{1}{2}h_4, \qquad k_3 = 1 + \tfrac{1}{2}k_2 + \tfrac{1}{2}k_4.$$

$$\Rightarrow h_2 = \tfrac{1}{2}\left(\tfrac{1}{2}h_2 + \tfrac{1}{2}\right) = \tfrac{1}{4}h_2 + \tfrac{1}{4} = \tfrac{1}{3}$$
$$k_2 = 1 + \tfrac{1}{2}\left(1 + \tfrac{1}{2}k_2\right) = \tfrac{3}{2} + \tfrac{1}{4}k_2 = 2$$

<u>Thm</u>. The vector of hitting probabilities $h^A = (h_i^A)_{i \in I}$ is the minimal non-negative solution to

$$(*) \quad \begin{cases} h_i^A = 1 & (i \in A) \\ h_i^A = \sum_{j \in I} p_{ij} h_j^A & (i \notin A) \end{cases}$$

Minimal means that if $x = (x_i)$ is another solution with $x_i \geq 0$ for all i then $x_i \geq h_i$ for all i.

<u>Proof</u>. <u>Step 1</u>: h^A is a solution to $(*)$.
If $X_0 = i \in A$, then clearly $H^A = 0$, so $h_i^A = 1$.
If $X_0 = i \notin A$, then by the Markov property,

$$\mathbb{P}_i[H^A < \infty \mid X_1 = j] = \mathbb{P}_j[H^A < \infty] = h_j^A$$

and

$$h_i^A = \mathbb{P}_i[H^A < \infty] = \sum_j \mathbb{P}_i[H^A < \infty, X_1 = j]$$
$$= \sum_j \underbrace{\mathbb{P}_i[H^A < \infty \mid X_1 = j]}_{h_j^A} \underbrace{\mathbb{P}_i[X_1 = j]}_{p_{ij}}$$

$\Rightarrow h^A$ is a solution to $(*)$.

<u>Step 2</u>. h^A is minimal.

Let x be any solution to $(*)$. If $i \in A$, clearly $x_i = h_i^A = 1$.

So suppose $i \in A$. Then:

$$x_i = \sum_{j \in I} p_{ij} x_j = \sum_{j \in A} p_{ij} + \sum_{j \notin A} p_{ij} x_j$$

substitute

$$\overset{\text{substitute}}{=} \sum_{j \in A} p_{ij} + \sum_{j \notin A} p_{ij} \left(\sum_{k \in A} p_{jk} + \sum_{k \notin A} p_{jk} x_k \right)$$

$$= \mathbb{P}_i[X_1 \in A] + \mathbb{P}_i[X_1 \notin A, X_2 \in A] + \sum_{j \notin A} \sum_{k \notin A} p_{ij} p_{jk} x_k.$$

By repeated substitution,

$$x_i = \underbrace{\mathbb{P}_i[X_1 \in A] + \mathbb{P}_i[X_1 \notin A, X_2 \in A] + \cdots + \mathbb{P}_i[X_1 \notin A, \dots, X_{n-1} \notin A, X_n \in A]}_{\mathbb{P}_i[H^A \le n]}$$

$$+ \underbrace{\sum_{j_1 \notin A} \cdots \sum_{j_n \notin A} p_{ij_1} p_{j_1 j_2} \cdots p_{j_{n-1} j_n} x_{j_n}}_{\ge 0 \text{ if } x_{j_n} \ge 0}.$$

So if x is nonnegative then

$$x_i \ge \mathbb{P}_i[H^A \le n] \text{ for all } n$$

$$\Rightarrow \quad x_i \ge \lim_{n \to \infty} \mathbb{P}_i[H^A \le n] = \mathbb{P}_i[H^A < \infty] = h_i^A.$$

Thus h^A is minimal.

<u>Example</u> (cont'd). Recall that $h = h^A$ with $A = \{4\}$. (*) is

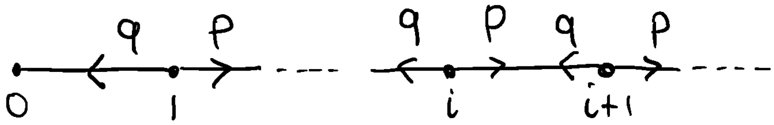
$$\begin{cases} h_4 = 1 & (i \in A) \\ h_1 = h_1 & \\ h_2 = \frac{1}{2} h_1 + \frac{1}{2} h_3 & (i \notin A) \\ h_3 = \frac{1}{2} h_2 + \frac{1}{2} h_4 & \end{cases}$$

$$\Rightarrow h_2 = \tfrac{1}{2}h_1 + \tfrac{1}{4}h_2 + \tfrac{1}{4}h_4 = \tfrac{2}{3}h_1 + \tfrac{1}{3}$$
$$h_3 = \tfrac{1}{2}h_2 + \tfrac{1}{2}h_4 = \tfrac{1}{3}h_1 + \tfrac{2}{3}.$$

The system (*) does not determine h_1, but by the minimality condition we must choose $h_1 = 0$. So again we find

$$h_2 = \tfrac{1}{3}.$$

Example (Gamblers' ruin).



$0 < p = 1-q < 1$

$$P_{00} = 1, \quad P_{i,i-1} = q, \quad P_{i,i+1} = p \quad \text{for } i = 1, 2, \dots$$

Starting with a fortune of i£, what is the probability of leaving broke? I.e., what is $h_i = \mathbb{P}_i[\text{hit } 0]$?

By the theorem $(h_i)_i$ is to minimal solution to

$$\begin{cases} h_0 = 1 \\ h_i = p h_{i+1} + q h_{i-1} \qquad (i = 1, 2, \dots). \end{cases}$$

Assume $p \neq q$. Then the general solution to the recursion is

$$h_i = A + B\left(\frac{q}{p}\right)^i.$$

If $p < q$ (most casinos): $0 \le h_i \le 1$ for all $i \Rightarrow B = 0,\ A = 1.$

$\Rightarrow h_i = 1$ for all i.

If $p > q$: $h_0 = 1 \Rightarrow B = 1 - A \Rightarrow h_i = \left(\frac{q}{p}\right)^i + A\left(1 - \left(\frac{q}{p}\right)^i\right)$

$h_i \ge 0$ for all $i \Rightarrow A \ge 0$

minimality $\Rightarrow A = 0 \Rightarrow h_i = \left(\frac{q}{p}\right)^i.$

If $p=q$ (fair casino), the general solution to the recursion is

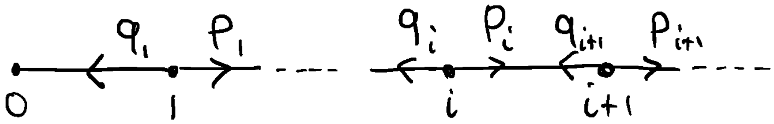
$$h_i = A + Bi$$

$0 \le h_i \le 1 \Rightarrow B = 0$

$h_0 = 1 \Rightarrow A = 1$

Thus even in a fair casino, you end up broke (Gamblers' ruin).

<u>Example</u> (Birth and death chain).



$0 < p_i = 1 - q_i < 1$

$h_i = \mathbb{P}_i[\text{hit } 0]$ is the extinction probability starting from i.

$$(*)\begin{cases} h_0 = 1 \\ h_i = p_i h_{i+1} + q_i h_{i-1} \qquad (i = 1, 2, \dots) \end{cases}$$

Consider $u_i = h_{i-1} - h_i$. Then

$$p_i u_{i+1} - q_i u_i = p_i h_i \underbrace{- p_i h_{i+1} - q_i h_{i-1}}_{-h_i} + q_i h_i = 0$$

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

$$\Rightarrow h_i = 1 - \underbrace{(h_0 - h_i)}_{u_1 + \cdots + u_i} = 1 - \underset{\substack{\uparrow \\ \text{unknown}}}{A}(\gamma_0 + \cdots + \gamma_{i-1}), \quad \text{where } \gamma_0 = 1.$$

$A = u_1$

If $\sum_{i=0}^{\infty} \gamma_i = \infty$: $0 \le h_i \le 1 \Rightarrow A = 0$
$\Rightarrow h_i = 1$ for all i.

If $\sum_{i=0}^{\infty} \gamma_i < \infty$: minimal solution is $A = \left(\sum_{i=0}^{\infty} \gamma_i\right)^{-1} \Rightarrow h_i = \dfrac{\sum_{j=i}^{\infty} \gamma_j}{\sum_{j=0}^{\infty} \gamma_j}$.

Since for any i, we have $h_i < 1$, the population survives with positive probability.

The mean hitting times $k_i^A = \mathbb{E}_i[H^A]$ satisfy a similar system of linear equations as the hitting probabilities.

<u>Thm</u>. The vector of mean hitting times $k^A = (k_i^A)_{i \in I}$ is the minimal solution to

$$(+) \begin{cases} k_i^A = 0 & (i \in A) \\ k_i^A = 1 + \sum_{j \notin A} p_{ij} k_j^A & (i \notin A). \end{cases}$$

<u>Proof</u>. <u>Step 1</u>: (k^A) satisfies $(+)$.
If $X_0 = i \in A$, then $H^A = 0$ so clearly $k_i^A = \mathbb{E}_i[H^A] = 0$.
If $X_0 = i \notin A$, then $H^A \ge 1$ so, by the Markov property,

$$\mathbb{E}_i[H^A \mid X_1 = j] = 1 + \mathbb{E}_j[H^A] = 1 + k_j^A$$

and $$k_i^A = \mathbb{E}_i[H^A] = \sum_{j \in I} \underbrace{\mathbb{E}_i[H^A \mid X_1 = j]}_{1 + k_j^A} \underbrace{\mathbb{P}_i[X_1 = j]}_{p_{ij}} = 1 + \sum_{j \in I} p_{ij} k_j^A = 1 + \sum_{j \notin A} p_{ij} k_j^A$$

Thus k^A is a solution to $(+)$.

Step 2. k^A is minimal.

Suppose x is any solution to (+). Then $x_i = k_i^A = 0$ for $i \in A$.
For $i \notin A$,

$$x_i = 1 + \sum_{j \notin A} p_{ij} x_j = 1 + \sum_{j \notin A} p_{ij} \left(1 + \sum_{k \notin A} p_{jk} x_k \right)$$

$$= \mathbb{P}_i[H^A \geq 1] + \mathbb{P}_i[H^A \geq 2] + \sum_{j \notin A} \sum_{k \notin A} p_{ij} p_{jk} x_k$$

Again, by repeated substitution, for any n,

$$x_i = \mathbb{P}_i[H^A \geq 1] + \cdots + \mathbb{P}_i[H^A \geq n] + \underbrace{\sum_{j_1 \notin A} \cdots \sum_{j_n \notin A} p_{ij_1} p_{j_1 j_2} \cdots p_{j_{n-1} j_n} x_{j_n}}_{\geq 0 \text{ if } x_{j_n} \geq 0}$$

So taking $n \to \infty$,

$$x_i \geq \sum_{n=1}^{\infty} \mathbb{P}_i[H^A \geq n] = \mathbb{E}_i[H^A] = k_i^A.$$

Thus k^A is the minimal solution.

4. Strong Markov Property

Defn. A random variable $T: \Omega \to \{0,1,2,\dots\} \cup \{+\infty\}$ is called a stopping time if the event $\{T=n\}$ only depends on $X_0, X_1, \dots, X_n$ for $n=0,1,2,\dots$.

Example. (a) The first passage time
$$T_j = \inf\{n \geq 1 : X_n = j\}$$
is a stopping time since $\{T_j = n\} = \{X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j\}$.

(b) The hitting time H^A of the last section is a stopping time because $\{H^A = n\} = \{X_1 \notin A, \dots, X_{n-1} \notin A, X_n \in A\}$.

(c) The last exit time
$$L^A = \sup\{n \geq 0 : X_n \in A\}$$
is in general not a stopping time because $\{L^A = n\}$ depends on whether $(X_{n+m})_{m\geq 1}$ visits A or not.

Thm (Strong Markov Property). Let $(X_n)_{n\geq 0}$ be Markov(λ, P), and let T be a stopping time of (X_n). Then, conditional on $T<\infty$ and $X_T = i$, $(X_{T+n})_{n\geq 0}$ is Markov(δ_i, P) and independent of $X_1, \dots, X_T$.

Proof. Let B be an event determined by $X_0, \dots, X_T$. Then $B \cap \{T=m\}$ is determined by $X_0, \dots, X_m$, so by the Markov property,
$$\mathbb{P}[\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B \cap \{T=m\} \cap \{X_T = i\}]$$
$$= \mathbb{P}[X_0 = j_0, X_1 = j_1, \dots, X_n = j_n]\, \mathbb{P}[B \cap \{T=m\} \cap \{X_T = i\}].$$

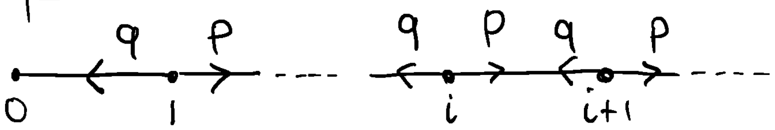
Summing over m gives

$$\mathbb{P}[\{X_T=j_0, X_{T+1}=j_1, \dots, X_{T+n}=j_n\} \cap B \cap \{T<\infty\} \cap \{X_T=i\}]$$
$$= \mathbb{P}[X_0=j_0, X_1=j_1, \dots, X_n=j_n]\, \mathbb{P}[B \cap \{T<\infty\} \cap \{X_T=i\}].$$

Dividing by $\mathbb{P}[T<\infty, X_T=i]$ gives the claim:

$$\mathbb{P}[\{X_T=j_0, X_{T+1}=j_1, \dots, X_{T+n}=j_n\} \cap B \mid T<\infty, X_T=i]$$
$$= \mathbb{P}[X_0=j_0, X_1=j_1, \dots, X_n=j_n]\, \mathbb{P}[B \mid T<\infty, X_T=i].$$

<u>Example</u>. (Gambler's ruin cont'd).



$0<p=1-q<1$

We have previously found $\mathbb{P}_1[\text{hit } 0]$. We now find the distribution of the time to hit 0 from 1. Let

$$H_j = \inf\{n \geq 0 : X_n = j\}$$
$$\phi(s) = \mathbb{E}_1[s^{H_0}] = \mathbb{E}_1[s^{H_0} 1_{H_0<\infty}] = \sum_{n<\infty} s^n\, \mathbb{P}[H_0=n] \qquad (s \in [0,1))$$

<u>Claim</u>: $\mathbb{E}_2[s^{H_0}] = \phi(s)^2$

Conditional on $H_1<\infty$ under $\mathbb{P}_2$, we can write $H_0 = H_1 + \tilde{H}_0$ where $\tilde{H}_0$ is the time after H_1 it takes after H_0 to reach 0.
Since H_1 is a stopping time, by the strong Markov Property at H_1, $\tilde{H}_0$ is independent of H_1 (it only depends on $(X_{H_1+n})_{n\geq 0}$).

$$\Rightarrow \mathbb{E}_2[s^{H_0}] = \mathbb{E}_2[s^{H_1} \mid H_1<\infty]\, \mathbb{E}_2[s^{\tilde{H}_0} \mid H_1<\infty]\, \mathbb{P}[H_1<\infty]$$

($\tilde{H}_0$ is independent from H_1)

$$= \mathbb{E}_2[s^{H_1} 1_{H_1<\infty}]\, \underbrace{\mathbb{E}_2[s^{\tilde{H}_0} \mid H_1<\infty]}_{\mathbb{E}_2[s^{H_1}]} = \mathbb{E}_2[s^{H_1}]^2 = \phi(s)^2.$$

Claim: $ps\,\phi(s)^2 - \phi(s) + qs = 0$

Conditional on $X_1=2$, we have $H_0 = 1 + \bar{H}_0$ where $\bar{H}_0$ is the time it takes after 1 to get to 0. Again, by the strong Markov property, $\bar{H}_0$ under $\mathbb{P}[\cdot \mid X_1=2]$ has the same distribution as H_0 under $\mathbb{P}_2$.

$$\Rightarrow \phi(s) = \mathbb{E}_1[s^{H_0}] = p\,\mathbb{E}_1[s^{H_0} \mid X_1=2] + q\,\mathbb{E}_1[s^{H_0} \mid X_1=0]$$

$$= p\,\mathbb{E}_1[s^{1+\bar{H}_0} \mid X_1=2] + qs$$

$\bar{H}_0$ under $\mathbb{P}_1[\cdot \mid X_1=2]$ has the same distribution as H_0 under $\mathbb{P}_2$.

$$= ps\,\mathbb{E}_2[s^{H_0}] + qs$$

$$= ps\,\phi(s)^2 + qs$$

$$\Rightarrow \phi(0) = 0 \quad \text{and} \quad \phi(s) = \frac{1 \pm \sqrt{1-4pqs^2}}{2ps} \quad \text{for } s>0.$$

Since $\phi(s) \le 1$ and since $\phi(s)$ is continuous, only the negative root is possible for all $s \in [0,1)$.

$$\Rightarrow \phi(s) = \frac{1-\sqrt{1-4pqs^2}}{2ps}$$

$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \cdots$

$$= \frac{1}{2ps}\left[1 - \left(1 + \tfrac{1}{2}(-4pqs^2) - \tfrac{1}{8}(-4pqs^2)^2 + \cdots\right)\right]$$

$$= qs + pq^2 s^2 + \cdots$$

$$= s\,\mathbb{P}[H_0=1] + s^2\,\mathbb{P}[H_0=2] + \cdots$$

So $\mathbb{P}[H_0=1] = q$

$\mathbb{P}[H_0=2] = pq^2$

$\vdots$

As $s \uparrow 1$, we have $\phi(s) \to \mathbb{P}_1[H_0 < \infty]$.

$$\Rightarrow \mathbb{P}_1[H_0 < \infty] = \frac{1 - \sqrt{1-4pq}}{2p} = \begin{cases} 1 & \text{if } p \le q \\ \frac{q}{p} & \text{if } p > q. \end{cases}$$

$$\sqrt{1-4pq} = \sqrt{1-4p(1-p)} = \sqrt{4p^2-4p+1} = |1-2p| = |1-2q|$$

Also, if $p \le q$,

$$\mathbb{E}_1[H_0] = \mathbb{E}_1[H_0 \mathbb{1}_{H_0<\infty}] = \lim_{s \uparrow 1} \phi'(s).$$

Differentiating the quadratic equation for $\phi(s)$ gives

$$2ps\,\phi(s)\phi'(s) + p\,\phi(s)^2 - \phi'(s) + q = 0.$$

$$\Rightarrow \phi'(s) = \frac{p\phi(s)^2 + q}{1 - 2ps\phi(s)} \xrightarrow{s \uparrow 1} \frac{1}{1-2p} = \frac{1}{q-p}.$$

$(p \le q)$

$$\Rightarrow \mathbb{E}_1[H_0] = \frac{1}{q-p}$$

5. Recurrence and transience

Defn. Let (X_n) be a Markov chain. A state $i \in I$ is

- recurrent if $\mathbb{P}_i[X_n = i \text{ for infinitely many } n] = 1$,
- transient if $\mathbb{P}_i[X_n = i \text{ for infinitely many } n] = 0$.

To state the next theorem, recall the first passage time to j:

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

Thm. The following dichotomy holds:

(a) If $\mathbb{P}_i[T_i < \infty] = 1$, then i is recurrent and $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$.

(b) If $\mathbb{P}_i[T_i < \infty] < 1$, then i is transient and $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$.

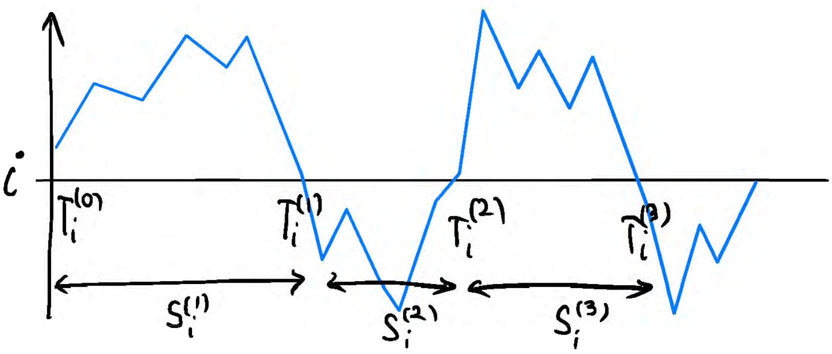
In particular, every state is either recurrent or transient.

In preparation of the proof, inductively define the r^{th} passage time to j:

$$T_j^{(0)} = 0, \quad T_j^{(1)} = T_j, \quad T_j^{(r+1)} = \inf\{n \geq T_j^{(r)} + 1 : X_n = j\}.$$

The length of the rth excursion is defined by

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



Lemma. For $r=2,3,\dots$, conditional on $T_i^{(r-1)}<\infty$, the excursion length $S_i^{(r)}$ is independent of $\{X_m : m < T_i^{(r-1)}\}$ and

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

Proof. By the strong Markov property, conditional on $T_i^{(r-1)}<\infty$, $(X_{T_i^{(r-1)}+n})_{n\geq 0}$ is Markov(δ_i, P) and independent of $X_0, \dots, X_{T_i^{(r-1)}}$. Now

$$S_i^{(r)} = \inf\{n \geq 1 : X_{T_i^{(r-1)}+n} = i\}$$

is the first passage time of $(X_{T_i^{(r-1)}+n})_{n\geq 0}$ to state i.

Let V_i denote the number of visits to i:

$$V_i = \sum_{n=0}^{\infty} 1_{X_n=i}.$$

Then

$$\mathbb{E}_i[V_i] = \mathbb{E}_i\left[\sum_{n=0}^{\infty} 1_{X_n=i}\right] = \sum_{n=0}^{\infty} \mathbb{P}_i[X_n=i] = \sum_{n=0}^{\infty} p_{ii}^{(n)}.$$

Let f_i be the return probability to i:

$$f_i = \mathbb{P}_i[T_i < \infty].$$

Lemma. For $r=0,1,2,\dots$, we have $\mathbb{P}_i[V_i > r] = f_i^r$.

Proof. Note that $\{V_i > r\} = \{T_i^{(r)} < \infty\}$ if $X_0 = i$. Also $\mathbb{P}_i[V_i > 0] = 1$. By induction,

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

Proof of theorem. (a) If $\mathbb{P}_i[T_i < \infty] = 1$, then by the last lemma,

$$\mathbb{P}_i[V_i = \infty] = \lim_{r \to \infty} \mathbb{P}_i[V_i > r] = 1.$$

So i is recurrent and

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i[V_i] = \infty.$$

(b) If $\mathbb{P}_i[T_i < \infty] < 1$, then

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \mathbb{E}_i[V_i] = \sum_{r=0}^{\infty} \mathbb{P}[V_i > r] = \sum_{r=0}^{\infty} f_i^r = \frac{1}{1-f_i} < \infty.$$

So $\mathbb{P}_i[V_i = \infty] = 0$ and i is transient.

Thm. Recurrence and transience are class properties: For any communicating class C, either all states $i \in C$ are recurrent or all are transient.

Proof. Let $i, j \in C$ and assume that i is transient. Since i and j communicate, there exist n, m s.t. $p_{ij}^{(n)} > 0$ and $p_{ji}^{(m)} > 0$. For all $r \geq 0$, then

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

$$\Rightarrow \sum_{r=0}^{\infty} p_{jj}^{(r)} \leq \frac{1}{p_{ij}^{(n)} p_{ji}^{(m)}} \sum_{r=0}^{\infty} p_{ii}^{(n+r+m)} < \infty.$$

So j is transient as well.

<u>Thm</u>. Every recurrent class is closed.

<u>Proof</u>. Let C be a class that is not closed, i.e., there are $i \in C$, $j \notin C$ and $m \geq 1$ s.t.

$$\mathbb{P}_i[X_m = j] > 0.$$

Since C is a communicating class and $j \notin C$,

$$\mathbb{P}_i[\{X_m = j\} \cap \{X_n = i \text{ for infinitely many } n\}] = 0.$$

$$\Rightarrow \mathbb{P}_i[X_n = i \text{ for infinitely many } n]$$
$$= \sum_{j \in I} \mathbb{P}_i[\{X_n = i \text{ for infinitely many } n\} \cap \{X_m = j\}]$$
$$< \sum_{j \in I} \mathbb{P}_i[X_m = j] = 1$$

Thus i is not recurrent and since recurrence is a class property this means that C is not recurrent.

<u>Thm</u>. Every <u>finite</u> closed class is recurrent.

<u>Proof</u>. Let C be a finite closed class and suppose $X_0 \in C$.

$\Rightarrow$ $0 < \mathbb{P}[X_n = i \text{ for infinitely many } n]$ for some $i \in C$

$= \mathbb{P}[X_n = i \text{ for some } n]\, \mathbb{P}_i[X_n = i \text{ for infinitely many } n]$

by the strong Markov property.

$\Rightarrow$ $\mathbb{P}_i[X_n = i \text{ for infinitely many } n] > 0 \Rightarrow i$ is not transient

$\Rightarrow i$ is recurrent.

Cor. Finite classes are recurrent iff closed.

Careful: Infinite closed classes may be transient.

Thm. Suppose P is irreducible and recurrent. Then for all $j \in I$,
$\mathbb{P}[T_j < \infty] = 1$.

Proof. It suffices to show that $\mathbb{P}_i[T_j < \infty]$ for all $i \in I$ since

$$\mathbb{P}[T_j < \infty] = \sum_i \mathbb{P}[X_0 = i]\, \mathbb{P}_i[T_j < \infty] \quad \text{by the Markov prop.}$$

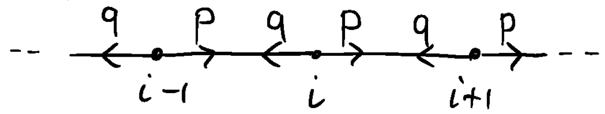
Since P is irreducible, there is m s.t. $p_{ji}^{(m)} > 0$.
Since P is recurrent,

$$\begin{aligned}
1 &= \mathbb{P}_j[X_n = j \text{ for infinitely many } n] \\
&= \mathbb{P}_j[X_n = j \text{ for some } n \geq m+1] \\
&= \sum_{k \in I} \underbrace{\mathbb{P}_j[X_n = j \text{ for some } n \geq m+1 \mid X_m = k]}_{\substack{\mathbb{P}_k[X_n = j \text{ for some } n \geq 1] \\ = \mathbb{P}_k[T_j < \infty]}} \underbrace{\mathbb{P}_j[X_m = k]}_{p_{jk}^{(m)}}
\end{aligned}$$

$\Rightarrow \mathbb{P}_i[T_j < \infty] = 1$ since $\sum_k p_{jk}^{(m)} = 1$ and $p_{ji}^{(m)} > 0$.

6. Recurrence and transience of random walks

<u>Example</u> (Simple random walk on $\mathbb{Z}$).



$0 < p = 1-q < 1$

$p_{00}^{(2n+1)} = 0$ since we cannot return to 0 after an odd number of steps

$$p_{00}^{(2n)} = \binom{2n}{n} p^n q^n = \frac{(2n)!}{(n!)^2} (pq)^n$$

Stirling's formula: $n! \sim \sqrt{2\pi n}\, e^{-n} n^n$

where we write $A_n \sim B_n$ if $\lim_{n\to\infty} \frac{A_n}{B_n} = 1$.

$$\Rightarrow p_{00}^{(2n)} \sim \frac{\sqrt{4\pi n}}{2\pi n} \frac{(2n)^{2n}}{n^{2n}} (pq)^n = \frac{c}{\sqrt{n}} (4pq)^n$$

Case $p = q = \frac{1}{2}$: $p_{00}^{(2n)} \sim \frac{c}{\sqrt{n}} \Rightarrow p_{00}^{(2n)} \geq \frac{c}{2\sqrt{n}}$ for $n \geq n_0$.

$$\Rightarrow \sum_{n=0}^{\infty} p_{00}^{(n)} \geq \sum_{n=n_0}^{\infty} p_{00}^{(2n)} \geq \frac{c}{2} \sum_{n=n_0}^{\infty} n^{-\frac{1}{2}} = \infty$$

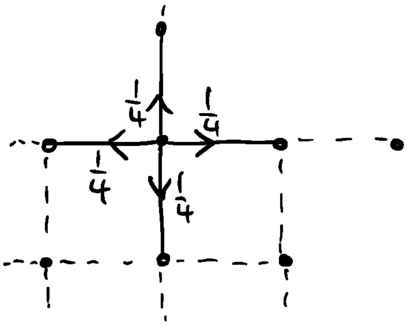
$\Rightarrow$ Random walk is recurrent.

Case $p \neq q$: $r = 4pq < 1 \Rightarrow p_{00}^{(2n)} \leq r^n$ for $n \geq n_0$

$$\Rightarrow \sum_{n=0}^{\infty} p_{00}^{(2n)} < \infty$$

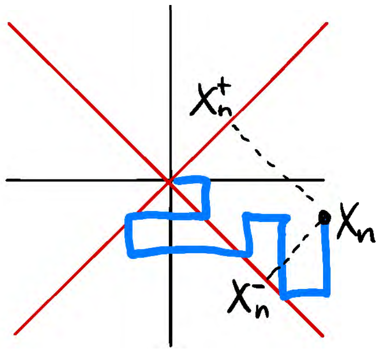
$\Rightarrow$ Random walk is transient.

Example (Simple symmetric random walk on $\mathbb{Z}^2$).



$$p_{ij} = \begin{cases} \frac{1}{4} & \text{if } |i-j| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Suppose $X_0 = 0$ and write $X_n^{\pm}$ for the orthogonal projections on the lines $y = \pm x$:



Observation: $X_n^{\pm}$ are independent simple symmetric random walks on $\frac{1}{\sqrt{2}}\mathbb{Z}$ and $X_n = 0$ iff $X_n^{\pm} = 0$.

$$\Rightarrow p_{00}^{(2n)} = \left(\binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\right)^2 \sim \frac{c}{n}.$$

both X^+ and X^- take $2n$ steps if X does and must return to 0

$$\Rightarrow \sum_{n=0}^{\infty} p_{00}^{(n)} = \infty$$

$\Rightarrow$ The random walk is recurrent.

Example (Simple symmetric random walk on $\mathbb{Z}^3$).

$$p_{ij} = \begin{cases} 1/6 & \text{if } |i-j| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

We will show the random walk is transient.

Again, $p_{00}^{(2n+1)} = 0$ and all walks from 0 to 0 must have have the same number of steps in direction $(1,0,0)$ as in direction $(-1,0,0)$, and analogously for the other two coordinates. Thus

$$p_{00}^{(2n)} = \sum_{\substack{i,j,k\geq 0\\ i+j+k=n}} \underbrace{\frac{(2n)!}{i!i!j!j!k!k!}}_{\binom{2n}{i\,i\,j\,j\,k\,k}} \left(\frac{1}{6}\right)^{2n} = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n} \sum_{\substack{i,j,k\geq 0\\ i+j+k=n}} \underbrace{\left(\frac{n!}{i!j!k!}\right)^2}_{\binom{n}{i\,j\,k}^2} \left(\frac{1}{3}\right)^{2n}$$

<u>Fact 1.</u> If $n=3m$ then $\binom{n}{i\,j\,k} \leq \binom{n}{m\,m\,m}$ for all i,j,k.

(Suppose the maximal $\binom{n}{i\,j\,k}$ has $i>j+1$. Then $i!j! > (i-1)!(j+1)!$. Thus $\binom{n}{i\,j\,k} < \binom{n}{i-1\;j+1\;k}$, so (i,j,k) was not maximal.)

<u>Fact 2.</u> $\displaystyle\sum_{\substack{i,j,k\geq 0\\ i+j+k=n}} \left(\frac{n!}{i!j!k!}\right)\left(\frac{1}{3}\right)^n = 1.$

(The LHS is the total prob. of distributing three balls in three bins.)

$$\Rightarrow p_{00}^{(2n)} \leq \binom{2n}{n}\left(\frac{1}{2}\right)^{2n}\binom{3m}{m\,m\,m}\left(\frac{1}{3}\right)^{3m} \overset{\text{Stirling's formula } n! \sim \sqrt{2\pi n}\, e^{-n} n^n}{\sim} C\frac{\sqrt{n}}{\sqrt{n}^2}\frac{\sqrt{n}}{\sqrt{n}^3} = C n^{-3/2}$$

provided that $n=3m$ for some $m\in\mathbb{N}$

Since $p_{00}^{(2n)} \geq \left(\frac{1}{6}\right)^2 p_{00}^{(2n-2)}$ up to changing C in fact

$p_{00}^{(2n)} \leq C n^{-3/2}$ for all n.

$\Rightarrow \sum_n p_{00}^{(n)} \leq C \sum_n n^{-3/2} < \infty$.

$\Rightarrow$ The random walk is transient.

7. Invariant distributions

Recall: a measure is a row vector $\lambda = (\lambda_i)_{i\in I}$ with $\lambda_i \geq 0\ \forall i$.

Defn. A measure λ is invariant (or stationary or in equilibrium) if $\lambda P = \lambda$.

Thm. Let $(X_n)_{n\geq 0}$ be Markov(λ, P) and suppose that λ is invariant for P. Then $(X_{n+m})_{n\geq 0}$ is also Markov(λ, P).

Proof.

$$\mathbb{P}[X_m = i] = (\lambda P^m)_i = \lambda_i \quad \text{for all } i$$

so the initial distribution of $(X_{n+m})_{n\geq 0}$ is λ

Also, conditional on $X_{m+n} = i$, by the Markov property for (X_n), X_{n+m+1} is independent of $X_m, X_{m+1}, \ldots, X_{n+m}$ and has distribution $(p_{ij})_j$.

Thm. Suppose I is finite. For some $i \in I$, suppose that

$$p_{ij}^{(n)} \longrightarrow \pi_j \quad \text{as } n\to\infty, \quad \text{for all } j \in I.$$

Then $(\pi_j)_j$ is an invariant distribution.

Proof. (π) is a distribution:

$$\sum_{j\in I} \pi_j = \sum_{j\in I} \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{j\in I} p_{ij}^{(n)} = 1.$$

I is finite

(π) is invariant:

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n+1)} = \lim_{n\to\infty} \sum_{k\in I} p_{ik}^{(n)} p_{kj} = \sum_{k\in I} \left(\lim_{n\to\infty} p_{ik}^{(n)}\right) p_{kj} = \sum_k \pi_k p_{kj}$$

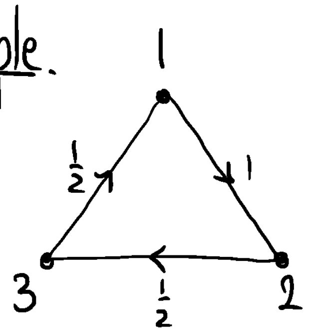
Rk. For the random walks on $\mathbb{Z}^d$, we have $p_{ij}^{(n)} \longrightarrow 0$ as $n \to \infty$, for all $i, j \in \mathbb{Z}^d$. The limit 0 is invariant, but not a distribution.

Example.
$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$
We found earlier that
$$p_{11}^{(n)} = \begin{cases} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \text{if } \alpha+\beta > 0 \\ 1 & \text{if } \alpha+\beta = 0. \end{cases}$$

So if $\alpha+\beta \notin \{0,1\}$, we have $p_{11}^{(n)} \longrightarrow \frac{\beta}{\alpha+\beta}$. Similarly,
$$P^n \longrightarrow \begin{bmatrix} \beta/(\alpha+\beta) & \alpha/(\alpha+\beta) \\ \beta/(\alpha+\beta) & \alpha/(\alpha+\beta) \end{bmatrix}$$

So $(\beta/(\alpha+\beta), \alpha/(\alpha+\beta))$ is an invariant distribution according to the theorem.

Example.



$$\pi P = \pi \quad \Leftrightarrow \quad \begin{cases} \pi_1 = \frac{1}{2}\pi_3 \\ \pi_2 = \pi_1 + \frac{1}{2}\pi_2 \\ \pi_3 = \frac{1}{2}\pi_2 + \frac{1}{2}\pi_3 \end{cases}$$

$$\pi_1 + \pi_2 + \pi_3 = 1 \quad \Rightarrow \quad \pi_3 = \tfrac{1}{2}(1-\pi_1) = \tfrac{1}{2}\left(1 - \tfrac{1}{2}\pi_3\right) = \tfrac{2}{5}$$
$$\pi_1 = \tfrac{1}{5}$$
$$\pi_2 = \tfrac{2}{5}$$

For each state k, let γ_i^k be the expected time spent in the state i between visits to k:

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

Thm. Let P be irreducible and recurrent. Then

(a) $\gamma_k^k = 1$;

(b) $\gamma^k = (\gamma_i^k)_{i\in I}$ is an invariant measure: $\gamma^k P = \gamma^k$

(c) $0 < \gamma_i^k < \infty$ for all $i \in I$.

Proof. (a) obvious from definition

(b) Since P is recurrent, $\mathbb{P}_k[T_k < \infty,\ X_0 = X_{T_k} = k] = 1$.

$$\Rightarrow \quad \gamma_j^k = \mathbb{E}_k \sum_{n=1}^{T_k} 1_{X_n=j}$$

$$= \mathbb{E}_k \sum_{n=1}^{\infty} 1_{X_n=j \text{ and } n\le T_k}$$

$$= \sum_{n=1}^{\infty} \mathbb{P}_k[X_n=j \text{ and } n \le T_k]$$

$$= \sum_{i\in I}\sum_{n=1}^{\infty} \mathbb{P}_k[X_{n-1}=i,\ X_n=j \text{ and } n\le T_k]$$

Since $n \le T_k$ depends only on $X_0, \dots, X_{n-1}$, by the Markov property

$$\mathbb{P}_k[X_{n-1}=i,\ X_n=j \text{ and } n\le T_k] = \mathbb{P}_k[X_{n-1}=i,\ n\le T_k]\,\underbrace{\mathbb{P}[X_n=j \mid X_{n-1}=i]}_{p_{ij}}$$

$$\Rightarrow \quad \gamma_j^k = \sum_{i\in I} p_{ij} \sum_{n=1}^{\infty} \underbrace{\mathbb{P}_k[X_{n-1}=i,\ n\le T_k]}_{\mathbb{E}_k[1_{X_{n-1}=i \text{ and } n\le T_k}]}$$

$$= \sum_{i\in I} p_{ij}\, \mathbb{E}_k \sum_{n=0}^{T_k-1} 1_{X_n=i} = \sum_{i\in I} p_{ij}\gamma_i^k \quad \text{i.e., } \gamma^k = \gamma^k P.$$

(c) P irreducible $\Rightarrow \exists\, n, m \geq 0$ s.t. $p_{ik}^{(n)} > 0$ and $p_{ki}^{(m)} > 0$.

$$\Rightarrow \gamma_i^k \overset{(b)}{\geq} \gamma_k^k p_{ki}^{(m)} \overset{(a)}{=} p_{ki}^{(m)} > 0$$

$$1 \overset{(a)}{=} \gamma_k^k \overset{(b)}{\geq} \gamma_i^k p_{ik}^{(n)} \Rightarrow \gamma_i^k \leq \frac{1}{p_{ik}^{(n)}} < \infty.$$

<u>Thm.</u> Let P be irreducible and λ be an invariant measure for P with $\lambda_k = 1$. Then $\lambda \geq \gamma^k$. If in addition P is recurrent, then $\lambda = \gamma^k$.

<u>Proof.</u> Since λ is invariant,

$$\begin{aligned}
\lambda_j = \sum_{i_1 \in I} \lambda_{i_1} p_{i_1 j} &= \sum_{i_1 \neq k} \lambda_{i_1} p_{i_1 j} + p_{kj} \\
&= \sum_{i_1 \neq k} \Big(\sum_{i_2 \neq k} \lambda_{i_2} p_{i_2 i_1} + p_{k i_1} \Big) p_{i_1 j} + p_{kj} \\
&= \cdots \\
&= \overbrace{\sum_{i_1, \dots, i_n \neq k} \lambda_{i_n} p_{i_n i_{n-1}} \cdots p_{i_1 j}}^{\geq 0} \\
&\quad + \Big(p_{kj} + \sum_{i_1 \neq k} p_{k i_1} p_{i_1 j} + \cdots + \sum_{i_1, \dots, i_{n-1} \neq k} p_{k i_{n-1}} \cdots p_{i_2 i_1} p_{i_1 j} \Big)
\end{aligned}$$

$\Rightarrow$ For $j \neq k$,

$$\begin{aligned}
\lambda_j &\geq \mathbb{P}_k[X_1 = j, T_k \geq 1] + \mathbb{P}_k[X_2 = j, T_k \geq 2] \\
&\quad + \cdots + \mathbb{P}_k[X_n = j, T_k \geq n] \\
&= \mathbb{E}_k\Big[\sum_{m=1}^{\min(n, T_k)} 1_{X_m = j}\Big] = \mathbb{E}_k\Big[\sum_{m=0}^{\min(n, T_k - 1)} 1_{X_m = j}\Big] \xrightarrow{(n \to \infty)} \gamma_j^k.
\end{aligned}$$

$X_m = j \Rightarrow m \neq T_k$ and $m \neq 0$

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

If P is recurrent, γ^k is invariant, so $\mu = \lambda - \gamma^k \geq 0$ is invariant.

P irreducible $\Rightarrow$ $\forall i\ \exists n$ s.t. $p_{ik}^{(n)} > 0$

$$\Rightarrow\ 0 = \mu_k = \sum_{j\in I} \mu_j p_{jk}^{(n)} \geq \mu_i p_{ik}^{(n)} \Rightarrow \mu_i = 0$$

$$\Rightarrow \mu = 0 \Rightarrow \lambda = \gamma^k.$$

Example. The simple symmetric random walk on $\mathbb{Z}$ is clearly irreducible and we have also seen it is recurrent.
The measure (π_i) where

$$\pi_i = 1 \quad \text{for all } i \in \mathbb{Z}$$

is invariant:

$$\pi = P\pi \iff \pi_i = \tfrac{1}{2}\pi_{i-1} + \tfrac{1}{2}\pi_{i+1} \quad \checkmark$$

By the theorem, every invariant measure is a multiple of π.
Since $\sum_{i\in I} \pi_i = \infty$, there is no invariant distribution.

Example. The simple symmetric random walk on $\mathbb{Z}^3$ has an invariant measure, but is not recurrent.

Recall: i is recurrent if $\mathbb{P}_i[X_n = i \text{ for inf. many } n] = 1$

$$\Leftrightarrow \mathbb{P}_i[T_i < \infty] = 1.$$

This does not imply that the expected return time is finite where the expected return time is defined as

$$m_i = \mathbb{E}_i[T_i]$$

Defn. • i is positive recurrent if $m_i < \infty$.
• i is null recurrent if i is recurrent but $m_i = \infty$.

Thm. Let P be irreducible. Then the following are equivalent:

(a) Every state is positive recurrent.

(b) Some state is positive recurrent.

(c) P has an invariant distribution π.

Moreover, when (c) holds then $m_i = 1/\pi_i$.

Proof. (a) $\Rightarrow$ (b): clear.

(b) $\Rightarrow$ (c): If i is positive recurrent, it is certainly recurrent. Therefore γ^i is invariant. Since

$$\sum_{j\in I} \gamma^i_m = m_i < \infty$$

$\pi_j = \gamma^i_j / m_i$ defines an invariant distribution.

(c) $\Rightarrow$ (a). Claim: For every $k \in I$, $\pi_k > 0$.

Indeed, since π is an invariant distribution and P is irreducilde,

$$\pi_k = \sum_{i\in I} \pi_i p_{ik}^{(n)} > 0 \quad \text{for some } n.$$

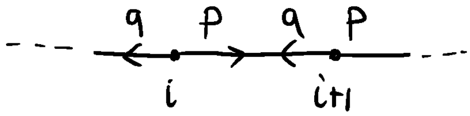
Now set $\lambda_i = \frac{\pi_i}{\pi_k}$. Then λ is an invariant measure with $\lambda_k = 1$. Therefore $\lambda \geq \gamma^k$.

$$\Rightarrow \quad m_k = \sum_{i\in I} \gamma^k_i \leq \sum_{i\in I} \frac{\pi_i}{\pi_k} = \frac{1}{\pi_k} < \infty. \qquad (*)$$

Thus k is positive recurrent.

Finally, knowing that P is recurrent, we have previously also seen that every invariant measure λ with $\lambda_k = 1$ must satisfy $\lambda = \gamma^k$. Thus we have $=$ in $(*)$.

Example.



$q<p$

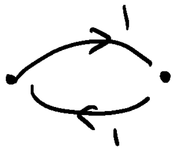
Invariant measure equation: $\pi_i = \sum_j \pi_j P_{ji} = \pi_{i-1} p + \pi_{i+1} q$

This is a recurrence relation with general solution

$$\pi_i = A + B(p/q)^i$$

So there is a two-parameter family of invariant measures. Uniqueness up to multiples does not hold.

8. Convergence to equilibrium

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

$\Rightarrow P^2 = 1 \Rightarrow P^{2n} = 1$ and $P^{2n+1} = P$

$\Rightarrow P^n$ does not converge

But note that P has invariant measure $\pi = (\frac{1}{2}, \frac{1}{2})$.

Defn. A state i is aperiodic if $p_{ii}^{(n)} > 0$ for n sufficiently large. P is aperiodic if all states are aperiodic.

Lemma. Let P be irreducible and have an aperiodic state i. Then for all $j, k \in I$,

$$p_{jk}^{(n)} > 0 \quad \text{for } n \text{ sufficiently large.}$$

In particular, all states are aperiodic.

Proof. P irreducible $\Rightarrow \exists\, r, s$ s.t. $p_{ji}^{(r)} > 0$, $p_{ik}^{(s)} > 0$.

$$\Rightarrow p_{jk}^{(r+n+s)} \geq p_{ji}^{(r)} p_{ii}^{(n)} p_{ik}^{(s)} > 0$$

since i is aperiodic.

Thm. Let P be irreducible and aperiodic and suppose it has invariant distribution π. Let λ be any distribution, and suppose that (X_n) is Markov(λ, P). Then, for all $j \in I$,

$$\mathbb{P}[X_n = j] \longrightarrow \pi_j \quad \text{as } n \to \infty.$$

In particular,

$$p_{ij}^{(n)} \longrightarrow \pi_j \quad \text{for all } i, j \in I.$$

Proof. The proof is by coupling. Let (Y_n) be Markov(π, P) and independent of (X_n). Fix a reference state $b \in I$ and set

$$T = \inf\{n \geq 1 : X_n = Y_n = b\}.$$

Claim: $\mathbb{P}[T < \infty] = 1$

$W_n = (X_n, Y_n)$ is a Markov chain with state space $I \times I$ and

- transition probabilities $\tilde{p}_{(i,k)(j,\ell)} = p_{ij} p_{k\ell}$
- and initial distribution $\tilde{\lambda}_{(i,k)} = \lambda_i \pi_k$.

Since P is aperiodic, the lemma implies that, for all i, j, k, ℓ,

$$\tilde{p}^{(n)}_{(i,k)(j,\ell)} = p_{ij}^{(n)} p_{k\ell}^{(n)} > 0 \quad \text{if } n \text{ is sufficiently large.}$$

$\Rightarrow \tilde{P}$ is irreducible.

$\tilde{P}$ has invariant distribution $\tilde{\pi}_{(i,k)} = \pi_i \pi_k$

$\Rightarrow \tilde{P}$ is positive recurrent.

Since T is the first passage time of the irred. and recurrent Markov chain (W_n) to (b,b), $\mathbb{P}[T < \infty] = 1$.

From the daim, it follows that

$$\mathbb{P}[X_n = i] = \mathbb{P}[X_n=i,\ n<T] + \underbrace{\mathbb{P}[X_n=i,\ n\geq T]}_{\mathbb{P}[Y_n=i,\ n\geq T] \quad \text{(strong Markov)}}$$

$$= \underbrace{\mathbb{P}[Y_n=i]}_{\pi_i} - \mathbb{P}[Y_n=i, n<T]$$

$$\Rightarrow \left|\mathbb{P}[X_n=i] - \pi_i\right| = \left|\mathbb{P}[X_n=i,\ n<T] - \mathbb{P}[Y_n=i,\ n<T]\right|$$

$$\leq \mathbb{P}[n<T] \longrightarrow 0.$$

<u>Example</u> (cont'd). $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\pi = (\frac{1}{2}\ \frac{1}{2})$

If X is Markov(δ_0, P) and Y is Markov(π, P), then with probability $1/2$ one has $Y_0=1$ but $X_0=0$ and X and Y will never meet.

9. Time reversal

Thm. Let P be irreducible and have invariant distribution π. Suppose $(X_n)_{0\le n\le N}$ is Markov(π, P) and set $Y_n = X_{N-n}$. Then $(Y_n)_{0\le n\le N}$ is Markov $(\pi, \hat{P})$ where

$$\pi_j \hat{p}_{ji} = \pi_i p_{ij}$$

and $\hat{P}$ is irreducible with invariant distribution π.

Proof. $\hat{P}$ is a stochastic matrix since

$$\sum_{i\in I} \hat{p}_{ji} \underset{\uparrow}{=} \frac{1}{\pi_j} \sum_{i\in I} \pi_i p_{ij} \underset{\uparrow}{=} \frac{\pi_j}{\pi_j} = 1.$$

$\pi_j > 0$ since P is irreducible and π invariant

π is invariant for P

π is invariant for $\hat{P}$ since

$$\sum_{j\in I} \pi_j \hat{p}_{ji} = \sum_{j\in I} \pi_i p_{ij} \underset{\uparrow}{=} \pi_i$$

P is a stochastic matrix

(Y_n) is Markov$(\pi, \hat{P})$ since

$$\begin{aligned}
\mathbb{P}[Y_0 = i_0, \dots, Y_N = i_N] &= \mathbb{P}[X_0 = i_N, \dots, X_N = i_0] \\
&= \pi_{i_N} p_{i_N i_{N-1}} \cdots p_{i_1 i_0} \\
&= \pi_{i_{N-1}} \hat{p}_{i_{N-1} i_N} p_{i_{N-1} i_{N-2}} \cdots p_{i_1 i_0} \\
&= \cdots = \pi_{i_0} \hat{p}_{i_0 i_1} \cdots \hat{p}_{i_{N-1} i_N}.
\end{aligned}$$

$\hat{P}$ is irreducible since by irreducibility of P, for all $i,j \in I$,

$$p_{i_0 i_1} \cdots p_{i_{n-1} i_n} > 0 \quad \text{for some } i_0, \dots, i_n \text{ with } i_0 = i,\ i_n = j.$$

$$\Rightarrow \hat{p}_{i_n i_{n-1}} \cdots \hat{p}_{i_1 i_0} = \frac{\pi_{i_0}}{\pi_{i_n}} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} > 0.$$

<u>Defn</u>. A stochastic matrix P and a measure λ are in <u>detailed balance</u> if

$$\lambda_i p_{ij} = \lambda_j p_{ji} \quad \text{for all } i,j \in I.$$

<u>Lemma</u>. If P and λ are in detailed balance then λ is invariant for P.

<u>Proof</u>. $(\lambda P)_i = \sum_{j \in I} \lambda_j p_{ji} = \sum_{j \in I} \lambda_i p_{ij} = \lambda_i$

<u>Defn</u>. Let P be irreducible and $(X_n)_{n \geq 0}$ be Markov(λ, P). (X_n) is <u>reversible</u> if, for all N, $(X_{N-n})_{0 \leq n \leq N}$ is also Markov(λ, P).

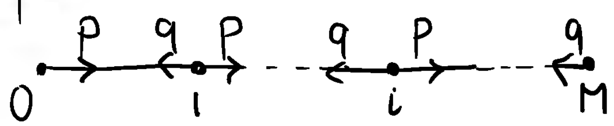
<u>Thm</u>. Let P be irreducible and let λ be a distribution. Suppose (X_n) is Markov(λ, P). Then the following are equivalent:

(a) $(X_n)_{n \geq 0}$ is reversible

(b) P and λ are in detailed balance.

<u>Proof</u>. Both (a) and (b) imply that λ is invariant. By the previous theorem both are thus equivalent to $P = \hat{P}$.

Example.



$0 < p = 1 - q < 1$

λ and P are in detailed balance

$\Leftrightarrow \lambda_i P_{i,i+1} = \lambda_{i+1} P_{i+1,i}$ for $i = 0, \dots, M-1$.

$\Leftrightarrow \lambda_i p = \lambda_{i+1} q$

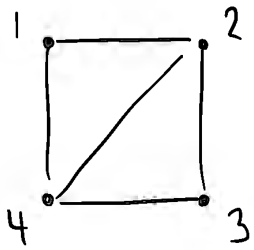
$\Leftrightarrow \lambda_i = C\left(\frac{p}{q}\right)^i$ for some constant C

Thus $\pi_i = \dfrac{\lambda_i}{\sum_j \lambda_j} = \tilde{C}\left(\frac{p}{q}\right)^i$ for suitable $\tilde{C}$ is an invariant distribution.

Hence the chain is reversible if started from π.

Example. (Random walk on a graph)

G =



graph

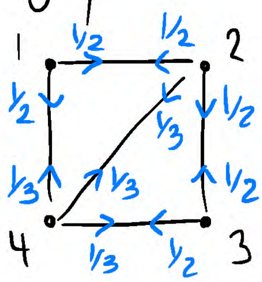


diagram obtained by picking neighbouring edges with equal probability

Let v_i be the valency of vertex i, i.e., the number of edges at i.

$$P_{ij} = \begin{cases} 1/v_i & \text{if } (i,j) \text{ is an edge} \\ 0 & \text{otherwise} \end{cases}$$

G connected $\Rightarrow$ P irreducible

P is in detailed balance with $v = (v_i)$: $v_i P_{ij} = 1 = v_j P_{ji}$.

10. Ergodic theorem

<u>Thm</u> (Strong Law of Large Numbers). Let $(Y_i)_{i=1,2,\dots}$ be a sequence of i.i.d. non-negative random variables with $\mathbb{E}Y_i = \mu$. Then

$$\mathbb{P}\left[\frac{Y_1 + \cdots + Y_n}{n} \longrightarrow \mu \text{ as } n \to \infty\right] = 1.$$

Let

$$V_i(n) = \sum_{k=0}^{n-1} 1_{X_k = i} \quad (= \text{number of visits to } i \text{ before } n).$$

<u>Thm</u> (Ergodic theorem). Let P be irreducible and let λ be a distribution. If (X_n) is Markov(λ, P) then

$$\mathbb{P}\left[\frac{V_i(n)}{n} \longrightarrow \frac{1}{m_i} \text{ as } n \to \infty\right] = 1. \qquad (*)$$

where $m_i = \mathbb{E}_i[T_i]$ is the expected return time to i.

In particular, if P is positive recurrent, then

$$P\left[\frac{V_i(n)}{n} \longrightarrow \pi_i \text{ as } n \to \infty\right] = 1.$$

Proof. Consider first the case that P is transient. Then

$\mathbb{P}[V_i < \infty] = 1$ where $V_i = \sum_{n=0}^{\infty} 1_{X_n = i}$ is the total number of visits to i.

$$\Rightarrow \mathbb{P}\left[\frac{V_i(n)}{n} \leq \frac{V_i}{n} \to 0 = \frac{1}{m_i}\right] = 1$$

as claimed. Thus from on we consider the case that P is recurrent.

Claim: it suffices to consider $\lambda = \delta_i$

Indeed, $\mathbb{P}[T_i < \infty] = 1$ and $(X_{T_i+n})_{n\geq 0}$ is Markov(δ_i, P) and independent of $X_0, \ldots, X_{T_i}$ by the strong Markov property. The claim follows since $\lim_n V_i(n)/n$ remains the same if $(X_n)_{n\geq 0}$ is replaced by $(X_{m+n})_{n\geq 0}$ for any finite m (here $m = T_i$).

Claim: $\mathbb{P}_i\left[\frac{n}{V_i(n)} \longrightarrow m_i \text{ as } n \to \infty\right] = 1$.

Let $S_i^{(r)}$ be the r-th excursion length. We have seen that:

- the $S_i^{(1)}, S_i^{(2)}, \ldots$ are independent
- the $S_i^{(k)}$ are identically distributed with $\mathbb{E}_i[S_i^{(k)}] = m_i$.

$$\underset{\text{strong LLN}}{\Rightarrow} \quad \mathbb{P}_i\left[\frac{S_i^{(1)} + \cdots + S_i^{(n)}}{n} \longrightarrow m_i\right] = 1. \qquad (*)$$

Now:
$$S_i^{(1)} + \cdots + S_i^{(V_i(n))} \geq n \quad \Rightarrow \quad \frac{S_i^{(1)} + \cdots + S_i^{(V_i(n))}}{V_i(n)} \geq \frac{n}{V_i(n)}$$
$$S_i^{(1)} + \cdots + S_i^{(V_i(n)-1)} \leq n-1 \quad \Rightarrow \quad \frac{S_i^{(1)} + \cdots + S_i^{(V_i(n)-1)}}{V_i(n)} \leq \frac{n}{V_i(n)}.$$

By (*) and $\mathbb{P}[V_i(n) \to \infty] = 1$, thus

$$\mathbb{P}\left[\frac{n}{V_i(n)} \longrightarrow m_i\right] = 1.$$