L1.1

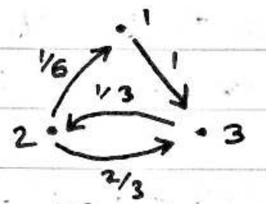
lectured by R Bauerschmidt
statslab → rb812

Markov Chains

Useful books include
- J Norris, Markov Chains
- G Grimmett, D Welsh, Probability

§0. Examples

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

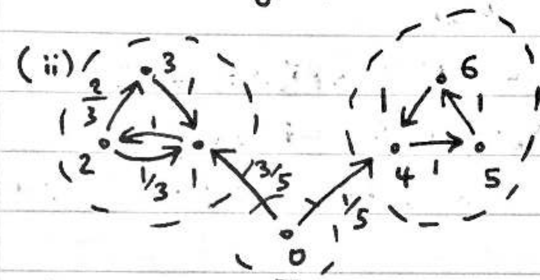


diagram

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1/6 & 1/6 & 2/3 \\ 0 & 1/3 & 2/3 \end{pmatrix}$$

transition matrix

(ii)



"communicating classes"

Starting from 0, what is the probability of hitting 6 ?
This is the probability of going right at some point.

$$\tfrac{1}{5} + \tfrac{2}{5} \cdot \tfrac{1}{5} + \cdots = \tfrac{1}{4}$$

Similarly, starting from 1, the probability of hitting 3 is 1.
Starting from 1, it takes on average 3 steps to hit 3.
Starting from 1, the long-term proportion of time spent at 2 is 3/8.

§1. Definitions and basic properties

We will make the following standing assumptions.
- I is a countable set, the state space; often $I = \mathbb{N}$
- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space on which all relevant random variables are defined.

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

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

It is <u>homogeneous</u> 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 <u>initial distribution</u> $\lambda = (\lambda_i)_{i \in I}$ given by $\lambda_i = \mathbb{P}[X_0 = i]$.

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

<u>Fact</u> · λ is a <u>distribution</u>, i.e., $\lambda_i \geq 0$ for all i and $\sum_{i \in I} \lambda_i = 1$.

· P is a <u>stochastic matrix</u>, i.e. $(P_{ij})_j$ is a distribution for every $i \in I$

<u>Def</u> (X_n) is a Markov Chain with initial distribution λ and transition matrix P, or (X_n) is Markov (λ, P) for short, if (a) and (b) hold.

<u>Thm</u> (X_n) is Markov (λ, P) iff for all $n \geq 0$, $i_0, \dots, i_n \in I$,

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

<u>Proof</u> Suppose (X_n) is Markov (λ, P). Then

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

$$\overset{\text{Markov}}{=} \underbrace{\mathbb{P}[X_n = i_n \mid X_{n-1} = i_{n-1}]}_{p_{i_{n-1} i_n} \, \mathbb{P}[X_1 = i_n \mid X_0 = i_{n-1}]} \cdot \mathbb{P}[X_0 = i_0, \dots, X_{n-1} = i_{n-1}]$$

$$\overset{\text{induc}}{=} p_{i_{n-1} i_n} \cdots p_{i_0 i_1} \mathbb{P}[X_0 = i_0] = p_{i_{n-1} i_n} \cdots p_{i_0 i_n} \lambda_{i_0}.$$

Conversely, assume (*) holds for all n and $i_0, \ldots, i_n \in I$.
Taking $n=0$, we get $\mathbb{P}[X_0 = i_0] = \lambda_{i_0}$.
Also, (*) gives

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

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

$$= \frac{\lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}}{\lambda_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}}} = p_{i_{n-1} i_n}$$

Thus (a) and (b) hold, i.e. (X_n) is Markov. □

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

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

<u>Thm</u> (Markov Property) Let (X_n) be Markov(λ, P).
Then conditional on $X_m = i$, the random variables

$$(X_{m+n})_{n=0,\ldots} \text{ is Markov } (\delta_i, P)$$

and it is independent of $X_0, \ldots, X_m$.

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

$$\mathbb{P}[X_m = i_m, \ldots, X_{m+n} = i_{m+n}, A \mid X_m = i]$$
$$= \delta_{i i_m} p_{i_m i_{m+1}} \cdots p_{i_{m+n-1} i_{m+n}} \mathbb{P}[A \mid X_m = i].$$

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 independent of $X_0, \dots, X_m$.

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

$$\mathbb{P}[\{X_m = i_m, \dots, X_{m+n} = i_{m+n}\} \cap A \mid X_m = i] = \delta_{i i_m} p_{i_m i_{m+1}} \cdots p_{i_{m+n-1} i_{m+n}} \mathbb{P}(A \mid X_m = i) \qquad (*)$$

$$\underset{\mathbb{P}(X_m = i) \neq 0}{\Longleftrightarrow} \quad \mathbb{P}[\{X_m = i_m, \dots, X_{m+n} = i_{m+n}\} \cap A]\, \delta_{i i_m} = \delta_{i i_m} p_{i_m i_{m+1}} \cdots p_{i_{m+n-1} i_{m+n}} \mathbb{P}(A \cap X_m = i)$$

For an elementary event $A = \{X_0 = i_0, \dots, X_m = i_m\}$ the equality $(*)$ follows from the previous theorem.
As any event A can be written as a countable disjoint union of elementary events, i.e.

$$A = \bigsqcup_{k=1}^{\infty} A_k$$

the claim follows by summing over k. yikes □

Notation We regard distributions as row vectors $(\lambda_i)_{i \in I}$
Here (λ_i) is a *measure* if $\lambda_i \geq 0$ for all $i \in I$
" distribution if in addition $\sum \lambda_i = 1$

Matrix multiplication

$$(\lambda P)_j = \sum_i \lambda_i p_{ij}$$

$$(P^2)_{ij} = \sum_k p_{ik} p_{kj} = p_{ij}^{(2)}$$

with $P^0 = 1$, the identity matrix.
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) $\mathbb{P}_i[X_n = j] = \cancel{\xi}\, \mathbb{P}[X_{n+m} = j \mid X_m = i] = p_{ij}^{(n)}$

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

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

Example The general two state Markov chain is given

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} \quad \text{where } 0 \leq \alpha, \beta \leq 1$$

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

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

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

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

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

2. Class structure

Def^n For $i, j \in I$,

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

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

(a) $i \rightarrow j$, (b) $p_{ii_1} p_{i_1 i_2} \cdots p_{i_n j} > 0$ for some $i_1, \cdots i_n, n$

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

Proof (a) $\Leftrightarrow$ (c):

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

Note both dir^{n}s.

($\cancel{a}$ c) $\Leftrightarrow$ (b):

$$p_{ij}^{(n)} = p_{ii_1} \cdots p_{i_{n-1}j}$$

□ yoke

Prop The relation $\leftrightarrow$ is an equivalence relation

Proof reflexive: $(n=0)$ ✓

symmetric: by def^n ✓

transitive: Spose $i \leftrightarrow j$, $j \leftrightarrow k$.

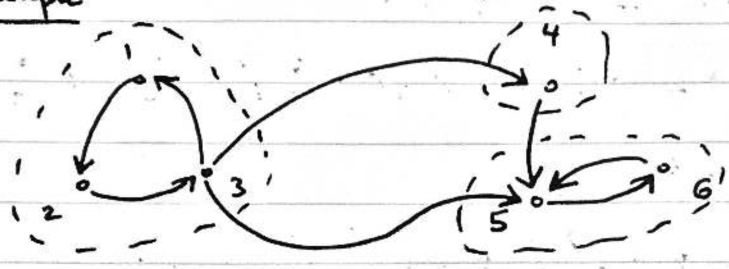
Then by (b) of Thm can go from i to k and back.

Def^n The equivalence classes of $\leftrightarrow$ are called communicating classes.

The chain is irreducible if it has one such class.

i.e. $\forall i, j$, $i \leftrightarrow j$

Example



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

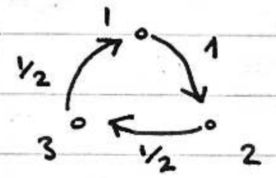
The communicating classes are $\{1,2,3\}, \{4\}, \{5,6\}$.

Def^n A subset $C \subset I$ is <u>closed</u> if
$$i \in C,\ i \to j \Rightarrow j \in C$$

In the example, "only" $\{5,6\}$ is closed.

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

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

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

- Find eigenvalues of P: $\lambda_1, \dots, \lambda_N$
- If all evals are <u>distinct</u>, then $p_{ij}^{(n)}$ has the form
$$p_{ij}^{(n)} = a_1 \lambda_1^n + \dots + a_N \lambda_N^n$$
for some constants $a_1, \dots, a_N$.

Indeed, since the λ_i are distinct, one can diagonalise P as
$$P = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} U^{-1}$$
$$\Rightarrow P^n = U \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_N^n \end{pmatrix} U^{-1}$$

- If an eval λ is repeated once, say, then the general form includes a term $(a + bn)\lambda^n$. Similar formulae hold for more repetitions.

This can be seen by writing P in JNF.

- 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.

<u>Example</u> (cont.) Evals of P:

$$0 = \det(\lambda I - P) = \lambda(\lambda - \tfrac{1}{2})^2 - \tfrac{1}{4} = \tfrac{1}{4}(\lambda - 1)(4\lambda^2 + 1)$$

$$\Rightarrow \lambda = 1, \pm i/2$$

$$\Rightarrow p_{11}^{(n)} = a + b\left(\tfrac{i}{2}\right)^n + c\left(-\tfrac{i}{2}\right)^n \quad \text{for some } \overset{\text{cx}}{a,b,c}$$

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

$$\Rightarrow p_{11}^{(n)} = \alpha + \left(\tfrac{1}{2}\right)^n\left[\beta\cos\left(\tfrac{\pi n}{2}\right) + \gamma\sin\left(\tfrac{\pi n}{2}\right)\right]$$

Now:
$$p_{11}^{(0)} = \alpha + \beta = 1 \qquad\qquad \alpha = 1/5$$
$$p_{11}^{(1)} = \alpha + \tfrac{1}{2}\gamma = 0 \quad \Rightarrow \quad \beta = 4/5$$
$$p_{11}^{(2)} = \alpha - \tfrac{1}{4}\beta = 0 \qquad\qquad \gamma = -2/5$$

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

<u>3. Hitting and absorption probabilities</u>

<u>Def^n</u> Let (X_n) be a Markov Chain.

- The <u>hitting time</u> of a set $A \subset I$ is the random variable

$$H^A : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$$
$$\omega \to \min(n \geq 0 : X_n(\omega) \in A)$$
$$\infty \text{ if } \forall n,\ X_n(\omega) \notin A$$

- The <u>hitting probability</u> of A is

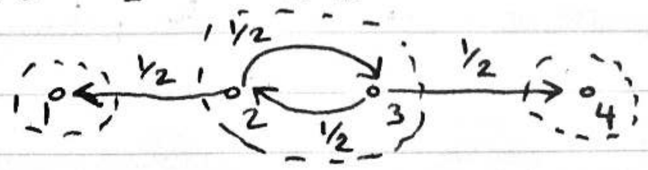
$$h_i^A = \mathbb{P}_i[H^A < \infty] = \mathbb{P}_i[\text{hit } A]$$

If A is a closed set, h_i^A is called the <u>absorption probability</u>.

- The <u>mean hitting time</u> is the expected value of H^A:

$$k_i^A = \mathbb{E}_i[H^A] = \mathbb{E}_i[\text{time to hit } A]$$

<u>Example</u>



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

i.e. what is $h_i = \mathbb{P}_i[\text{hit } 4]$ and $k_i = \mathbb{E}_i[\text{time to absorb}]$

Note that

$$h_1 = 0 \qquad k_1 = 0$$
$$h_4 = 1 \qquad k_4 = 0$$
$$h_2 = \tfrac{1}{2}h_1 + \tfrac{1}{2}h_3 \qquad k_2 = 1 + \tfrac{1}{2}(k_1 + k_3)$$
$$h_3 = \tfrac{1}{2}h_4 + \tfrac{1}{2}h_2 \qquad k_3 = 1 + \tfrac{1}{2}(k_4 + k_2)$$

$$\Rightarrow \quad h_2 = \tfrac{1}{3} \qquad k_2 = 2$$
$$h_3 = \tfrac{2}{3} \qquad k_3 = 2$$

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

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

Minimality means that if $x = (x_i)_{i \in I}$ is another solution with $x_i \geq 0$ for all i, then $x_i \geq h_i$ for all $i \in I$.

$H^A = \inf\{n \geq 0 : X_n \in A\}$, $A \subset I$.

$h_i^A = \mathbb{P}_i[H^A < \infty] = \mathbb{P}_i[\text{hit } A]$

$k_i^A = \mathbb{E}_i[H^A] = \mathbb{E}_i[\text{time to hit } A]$

__Thm__ The vector of hitting probabilities $h^A = (h_i^A)_{i\in I}$ is the __minimal__ non-neg solution to

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

$x_i \geq h_i^A$ for any other sol^n $(x_i)_{i\in I}$

__Proof__ __Step 1__: h^A is a sol^n to $(*)$

If $X_0 = i \in A$, then $H^A = 0$, so $h_i^A = 1$.

If $X_0 = i \notin A$, then conditioning on X_1,

$$h_i^A = \mathbb{P}_i[H^A < \infty] = \sum_{j\in I} \mathbb{P}_i[H^A < \infty, X_1 = j]$$

$$= \sum_{j\in I} \underbrace{\mathbb{P}_i[H^A < \infty \mid X_1 = j]}_{\substack{\mathbb{P}_j[H^A < \infty] \\ \text{by the Markov prop}}} \mathbb{P}_i[X_1 = j]$$

$$= \sum_{j\in I} p_{ij} h_j^A,$$

as desired.

__Step 2__: h^A is minimal

Let $x = (x_i)_{i\in I}$ be any (non-negative) solution to $(*)$.

If $i \in A$, clearly $x_i = h_i^A = 1$. (So $x_i \geq h_i^A$)

So suppose $i \notin A$. Then

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

$$= \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij} \Big(\sum_{k\in A} p_{jk} + \sum_{k\notin A} p_{jk} x_k \Big)$$

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

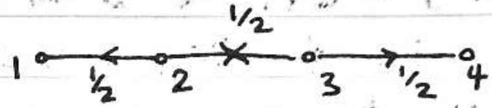
By repeated substitution, for each n

$$x_i \geq \mathbb{P}_i[X_1 \in A] + \cdots + \mathbb{P}_i[X_1 \notin A, \ldots, X_{n-1} \notin A, X_n \in A]$$

So for each n, $x_i \geqslant \mathbb{P}_i[H^A \leqslant n]$.

By continuity of probability, $x_i \geqslant \mathbb{P}_i[H^A < \infty] = h_i^A$. $\square$

<u>Example</u> (cont.) $\qquad A = \{4\}$



(*) is $\begin{cases} h_4^A = 1; & h_2^A = \frac{1}{2}h_1^A + \frac{1}{2}h_3^A \\ h_1^A = h_1^A; & h_3^A = \frac{1}{2}h_4^A + \frac{1}{2}h_2^A \end{cases}$

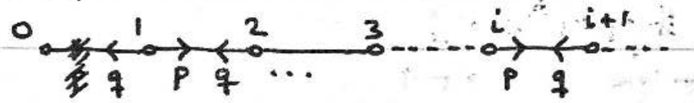
$$\Rightarrow h_2 = \tfrac{1}{2}h_1 + \tfrac{1}{4}h_2 + \tfrac{1}{4}h_4 = \tfrac{2}{3}h_1 + \tfrac{1}{3}$$

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

The system does not determine h_1, but by the minimality condition, we must choose $h_1 = 0$.

So again, $h_2 = \frac{1}{3}$, $h_3 = \frac{2}{3}$.

<u>Example</u> (Gambler's ruin) $\quad I = \{0, 1, 2, \dots\}$



$0 < p = 1-q < 1$.

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

Starting with £i, what is the probability of leaving the casino broke?

i.e. What is $h_i = \mathbb{P}_i[\text{hit } 0]$?

By the theorem, $(h_i)_{i \in I}$ is the minimal sol^n to

$$\begin{cases} h_0 = 1 \\ h_i = p h_{i+1} + q h_{i-1} & \text{for } i = 1, 2, \dots \end{cases}$$

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

$$h_i = A + B(q/p)^i.$$

For $p < q$ (most casinos), $0 \leqslant h_i \leqslant 1$ for all i

$\Rightarrow B = 0$ for otherwise $h_i \to \infty$

$\Rightarrow h_i = 1$ for all i

For $p > q$, $h_0 = 1 \Rightarrow A + B = 1$

$$\Rightarrow h_i = (q/p)^i + A(1 - (q/p)^i)$$

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

minimality $\Rightarrow A = 0 \Rightarrow h_i = (q/p)^i$

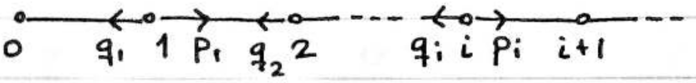
For $p = q = \frac{1}{2}$ (fair casino), the general solution is

$$h_i = A + Bi$$

Since $0 \le h_i \le 1$, $B = 0$ and hence $h_i = 1$ for all i.

Thus, even in a fair casino, the probability to go broke is 1.

<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 from i

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

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

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

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

$$\Rightarrow h_i = 1 - (h_0 - h_i) = 1 - (u_{\cancel{0}\,1} + \cdots + \underset{-1}{u_i})$$

$$= 1 - A(\gamma_0 + \gamma_1 + \cdots + \gamma_{i-1}) \quad \text{where } \gamma_0 = 1, \; A = u_1.$$

If $\sum_i \gamma_i = \infty$, $0 \le h_i \le 1 \Rightarrow A = 0$, $h_i = 1$ for all i

If $\sum_i \gamma_i < \infty$, minimality gives $A = (\sum_i \gamma_i)^{-1}$

$$\therefore h_i = \sum_{j=i}^{\infty} \gamma_j \Big/ \sum_{j=1}^{\infty} \gamma_i < 1$$

This is positive, less than 1 for each i.

$H^A = \inf\{n \geq 0 : X_n \in A\}$, $A \subset I$

$h_i^A = \mathbb{P}_i[H^A < \infty] = \mathbb{P}_i[\text{hit } A]$

$k_i^A = \mathbb{E}_i[H^A] = \mathbb{E}_i[\text{time to hit } A]$

<u>Thm</u> The vector $(k_i^A)_{i \in I}$ is the minimal (non-negative) solution to

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

<u>Proof</u> <u>Step 1</u> (k_i^A) satisfies $(\dagger)$

If $X_0 = i \in A$, then $H^A = 0$, so clearly $k_i^A = 0$.

If $X_0 = i \notin A$, then $H^A \geq 1$, so by the Markov property,

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

$$\Rightarrow k_i^A = \mathbb{E}_i[H^A] = \sum_{j \in I} \mathbb{E}_i[H^A | X_1 = j] \, \mathbb{P}_i[X_1 = j]$$

$$= \sum_{j \in I}(1 + k_j^A) p_{ij} = 1 + \sum_{j \in I} k_j^A p_{ij}$$

$\therefore (k_i^A)$ is a solution to $(\dagger)$

<u>Step 2</u> (k_i^A) is minimal

Suppose $x = (x_i)$ is a non-negative solution to $(\dagger)$.

Then for $i \in A$, $x_i = k_i^A = 0$.

For $i \notin A$, $x_i = \sum_{j \notin A} x_j p_{ij} + 1$

$$= 1 + \sum_{j \notin A} p_{ij} \underbrace{\sum_{k \notin A}}_{\text{swap}} (1 + p_{jk} x_k)$$

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

By repeated substitution, $\forall n \in \mathbb{N}$,

$$x_i \geq \mathbb{P}_i[H^A \geq 1] + \cdots + \mathbb{P}_i[H^A \geq n]$$

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

$\therefore k_i^A$ is minimal, as desired $\square$

4. Strong Markov Property

<u>Def^n</u> A random variable $T : \Omega \to \{0, 1, 2, \dots, \} \cup \{+\infty\}$ is called a <u>stopping time</u> if the event $\{T = n\}$ only depends on $X_0, \dots, X_n$ for all $n = 1, 0, 2, \dots$

<u>Example</u> (a) The <u>first passage time</u>

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

is a stopping time since
$$\{J_j = n\} = \{X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j\}.$$
(b) The <u>hitting times</u> H^A of the previous section are also stopping times.
$$\{H^A = n\} = \{X_0 \notin A, \dots, X_{n-1} \notin A, X_n \in A\}.$$
ponder

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

<u>Thm</u> (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 $X_0, \dots, X_T$.

<u>Proof</u> 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 usual Markov property,
$$\mathbb{P}[\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_T = i\}]$$
$$= \mathbb{P}_i[\{X_0 = j_0, \dots, X_n = j_n\}]\, \mathbb{P}[B \cap \{T = m\} \cap \{X_T = i\}]$$
Summing over m gives
$$\mathbb{P}[\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B \cap \{T < \infty\} \cap \{X_T = i\}]$$
$$= \mathbb{P}_i[\{X_0 = j_0, \dots, X_n = j_n\}]\, \mathbb{P}[B \cap \{T < \infty\} \cap \{X_T = i\}]$$
Dividing by $\mathbb{P}[T < \infty, X_T = i]$ gives the claim:
$$\mathbb{P}[\{X_T = j_0, \dots, X_{T+n} = j_n\} \cap B \mid T < \infty, X_T = i]$$
$$= \mathbb{P}_i[\{X_0 = j_0, \dots, X_n = j_n\}]\, \mathbb{P}[B \mid T < \infty, X_T = i] \quad \square$$

<u>Example</u> (Gambler's ruin)

$0 < p = 1 - q < 1$

We have previously found $\mathbb{P}_i[\text{hit } 0]$. We will now find the distribution of $\{\text{time to hit } 0\}$. Let
$$H_j = \inf\{n \geq 0 : X_n = j\}$$
starting from 1
$$\phi(s) = \mathbb{E}_1[s^{H_0}] = \mathbb{E}_1[s^{H_0} \mathbb{1}_{H_0 < \infty}]$$
$$= \sum_{n=0}^{\infty} s^n\, \mathbb{P}[H_0 = n]$$

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 to reach 0.

By the strong Markov property, as H_1 is a stopping time, $\tilde{H}_0$ is independent from H_1, and its distribution under $\mathbb{P}_1$ is the same as that of H_1 under $\mathbb{P}_2$.

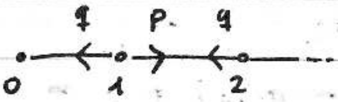
should be a 2?

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

? $\mathbb{E}_{\cancel{1}\,2}[\cancel{1}\, s^{H_1}]$ yes

$$= \mathbb{E}_2[s^{H_1} \mathbb{1}_{H_1 < \infty}]\, \mathbb{E}_2[s^{H_1}]$$

$$= \mathbb{E}_2[s^{H_1}]^2 = \phi(s)^2$$

Last time  $0 < p = 1-q < 1$

$H_j = \inf\{n \geq 0 : X_n = j\}$

$\phi(s) = \mathbb{E}_1[s^{H_0}] = \mathbb{E}_1[s^{H_0} \mathbb{1}_{H_0 < \infty}]$ where $0 \leq s < 1$.

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

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

Conditioning on $X_1 = 2$, we can write $H_0 = 1 + \overline{H}_0$ where $\overline{H}_0$ is the time after the first step to reach 0. By the "strong" Markov property, $\overline{H}_0$ under $\mathbb{P}[\cdot \mid X_1 = 2]$ has the same distribution as H_0 under $\mathbb{P}_2$.

$$\Rightarrow \phi(s) = \mathbb{E}_1[s^{H_0}] = p \underbrace{\mathbb{E}_1[s^{H_0} \mid X_1 = 2]}_{s \underbrace{\mathbb{E}_1[s^{\overline{H}_0} \mid X_1 = 2]}_{\mathbb{E}_2[s^{H_0}] = \phi(s)^2}} + q \underbrace{\mathbb{E}_1[s^{H_0} \mid X_1 = 0]}_{s}$$

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

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

Since $\phi(s) \leq 1$ for $0 \leq s < 1$ and ϕ is continuous, only the negative is possible

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

$$= qs + pq^2 s^3 + \cdots$$

On the other hand,

$$\phi(s) = \sum_{i=0}^{\infty} s^i \,\mathbb{P}[H_0 = i]$$

so we recover the distribution of H_0.

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

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

Also, if $p \leq q$,

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

Differentiating the quadratic equation,

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

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

$$\Rightarrow \mathbb{E}[H_0] = \frac{1}{q-p} \text{ if } q > p, \quad \infty \text{ if } q = p.$$

5. Recurrence & Transience

Def Let (X_n) be a Markov Chain. A state $i \in I$ is recurrent if $\mathbb{P}_i[X_n = i \text{ for infinitely many } n] = 1$, and is transient if $\mathbb{P}_i[X_n = i \text{ for infinitely many } n] = 0$.

For the following, recall the first passage time to $j \in I$, $T_j = \inf\{n \geq 1 : X_n = j\}$.

Thm (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 for the proof, inductively define the r^{th} passage time to j: $T_j^{(0)} = 0$, $T_j^{(1)} = T_j$, $T_j^{(r+1)} = \inf\{n > T_j^{(r)} : X_n = j\}$

The length of the r^{th} excursion is defined by

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

Lemma For $r = 2, 3, \ldots$, conditional on $T_j^{(r-1)} < \infty$, the excursion length $S_j^{(r)}$ is independent of X_m for $m < T_j^{(r-1)}$ and $\mathbb{P}[S_j^{(r)} = n \mid T_j^{(r-1)} < \infty] = \mathbb{P}_j[T_j = n]$.

Proof By the strong Markov property, conditional on $T_j^{(r-1)}$ being finite, $(X_{T_j^{(r-1)}+n})_{n\geq 0}$ is Markov(δ_j, P), and independent of $X_0, \ldots, X_{T_j^{(r-1)}}$.

Now $S_j^{(r)} = \inf\{n \geq 1 : X_{T_j^{(r-1)}+n} = j\}$ is the first passage time of $(X_{T_j^{(r-1)}+n})_{n\geq 0}$ to state j.

L6.3

Let V_i denote the number of visits to i:

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

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

L 7.1

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 $X_0 = i$

$\Rightarrow \mathbb{P}_i[V_i > 0] = 1$

By induction

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

Proof of Theorem (a) If $\mathbb{P}_i[T_i < \infty] = 1$, then by the last Lemma, $\mathbb{P}_i[V_i = \infty] = \lim_{r \to \infty} \mathbb{P}_i[V_i > r] = 1$.

This is exactly what it means for i to be recurrent. Also,

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

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

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

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

Thm Recurrence and transcience 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. It suffices to show that j is transient as well. Since i and j communicate, there exist n, m s.t. $p_{ij}^{(n)} > 0$ and $p_{ji}^{(m)} > 0$. For all $r \geq 0$,

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

$$\Rightarrow \sum_{r=0}^{\infty} p_{jj}^{(r)} \leq \frac{1}{p_{ij}^{(n)} p_{ji}^{(m)}} \underbrace{\sum_{r=0}^{\infty} p_{ii}^{(n+r+m)}}_{<\infty} < \infty \quad \text{by transcience}$$

So j is transient as well.

Thm Every recurrent class is closed

Proof Let C be a class that is not closed, i.e. $\exists\, i \in C, j \notin C$ s.t. $p_{ij}^{(m)} > 0$ for some $m \geqslant 1$.

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

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

Now! $\Rightarrow$ $\mathbb{P}_i[X_n = i \text{ for infinitely many } n]$

$$= \sum_{k \in I} \underbrace{\mathbb{P}_i[X_m = k, X_n = i \text{ for infinitely many } n]}_{\substack{= 0 \text{ if } k = j;\ \text{but } \mathbb{P}_i[X_m = j] > 0 \\ < \mathbb{P}_i[X_m = j] \text{ if } k \neq j}}$$

think it

$$< \sum_{k \in I} \mathbb{P}_i[X_m = k] = 1$$

Thus, i is not recurrent. □

Thm ~~Cor~~ Every Finite closed class~~es~~ is recurrent.

Proof Let C be a finite closed class and $X_0 \in C$.

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

$= \mathbb{P}[X_n = i \text{ for some } n]\ \mathbb{P}_i[X_n = i \text{ for infinetily may } n]$

spell out by Strong Markov?

$\Rightarrow \mathbb{P}_i[X_n = i \text{ for infinitely many } n] > 0$

$\Rightarrow i$ is recurrent, so C is recurrent. □

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

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

$$\mathbb{P}[T_j < \infty] = \sum_{i \in I} \mathbb{P}[X_0 = i]\, \mathbb{P}_i[T_j < \infty]$$

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

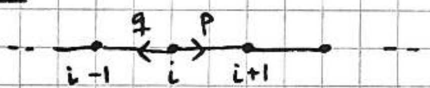
Since P is recurrent,

$$\begin{aligned} 1 = \mathbb{P}_j[X_n = j \text{ for } \infty\ n] &= \mathbb{P}_j[X_n = j \text{ for some } n \geq m+1] \\ &= \sum_{k \in I} \mathbb{P}_j[X_m = k]\, \mathbb{P}_j[X_n = j \text{ for some } n \geq m+1 \mid X_m = k] \\ &= \sum_{k \in I} \mathbb{P}_j[X_m = k]\, \mathbb{P}_k[X_n = j \text{ for some } n \geq 1] \\ &= \sum_{k \in I} \mathbb{P}_j[X_m = k]\, \mathbb{P}_k[T_j < \infty] \end{aligned}$$

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

6. Recurrence and transcience of random walks

Example (Simple random walk on $\mathbb{Z}$)



$0 < p = 1 - q < 1$

$p_{00}^{(2n+1)} = 0$ since chain cannot return in odd steps

$$p_{00}^{(2n)} = \binom{2n}{n} / 2^{2n} \text{ oof } = \binom{2n}{n} (pq)^n = \frac{(2n)!}{(n!)^2} (pq)^n$$

Stirling formula $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$

where $a_n \sim b_n$ meas $\frac{a_n}{b_n} \to 1$.

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

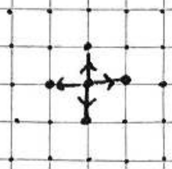
Case $p = q = \frac{1}{2}$: $p_{00}^{(2n)} \sim \frac{C}{\sqrt{n}} \Rightarrow p_{00}^{(2n)} \geq \frac{C}{2\sqrt{n}}$ for $n \geq n_0$

So $\sum_{n=0}^{\infty} p_{00}^{(2n)}$ diverges, and walk is ~~transient~~ " recurrent.

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

So $\sum_{n=0}^{\infty} p_{00}^{(2n)}$ converges, and walk is transient.

<u>Example</u> (Simple symmetric random walk on $\mathbb{Z}^2$)



$$p_{ij} = \begin{cases} 1/4 & \text{if } \|i-j\| = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{for } i,j \in \mathbb{Z}^2$$

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

Observe that $X_n^{\pm}$ are <u>independent</u> simple symmetric random walks on $\frac{1}{\sqrt{2}}\mathbb{Z}$ and $X_n = 0$ iff $X_n^{\pm} = 0$.

$$\Rightarrow \underset{\substack{\uparrow \\ 2D}}{p_{00}^{(2n)}} = \left(\binom{2n}{n} / 4^n\right)^2 \sim \frac{C}{n}$$

$\Rightarrow \sum_{n=0}^{\infty} p_{00}^{(2n)}$ diverges and walk is recurrent $\ddot{\circ}$

<u>Example</u> (Simple symmetric random walk on $\mathbb{Z}^3$)

$$p_{ij} = \begin{cases} 1/6 & \text{if } \|i-j\| = 1, \\ 0 & \text{otherwise} \end{cases} \quad \text{for } i,j \in \mathbb{Z}^3.$$

We will show the walk is transient.

$$p_{00}^{(2n+1)} = 0$$

$$p_{00}^{(2n)} = \sum_{\substack{i,j,k \\ i+j+k=n}} \binom{2n}{i,i,j,j,k,k} 6^{-2n}$$

$$= \binom{2n}{n} 2^{-2n} \sum_{\substack{i,j,k \\ i+j+k=n}} \left(\frac{n!}{i!\,j!\,k!}\right)^2 3^{-2n}$$

<u>Fact</u> if $n = 3m$ then $\binom{n}{i,j,k} \le \binom{n}{m,m,m}$ $\forall i,j,k$

(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 not maximal.)

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

$$p_{00}^{(2n)} = \binom{2n}{n} 2^{-2n} \sum_{i+j+k=n} \underbrace{\binom{n}{i\,j\,k}^2}_{\le \binom{n}{m\,m\,m}\binom{n}{i\,j\,k} \text{ if } n=3m} 3^{-2n}$$

$$\therefore\ p_{00}^{(2n)} \le \binom{2n}{n} 2^{-2n} \binom{n}{m\,m\,m} 3^{-n}$$

Now, apply Stirling

$$\text{RHS} \sim C \frac{\sqrt{n}}{\sqrt{n^2}} \cdot \frac{\sqrt{n}}{\sqrt{n^3}} = C n^{-3/2}$$

Since $p_{00}^{(2n)} \geq \frac{1}{6} p_{00}^{(2n-2)}$ up to changing the constant

$$p_{00}^{(2n)} \le C n^{-3/2} \quad \forall n \geq n_0$$

So $\sum_{n=1}^{\infty} p_{00}^{(n)} < \infty$ and the walk is transient

7. Invariant distributions by Tsuneo

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

Def^n : a measure λ is invariant for a given Markov Chain (or stationary or in equilibrium) if

$$\lambda P = \lambda.$$

Thm Let $(X_n)_{n\geq 0}$ be a Markov chain (λ, P) where λ is invariant under P. Then $(X_{n+m})_{n\geq 0}$ is Markov (λ, P) as well.

Proof $\mathbb{P}[X_m = i] = (\lambda P^m)_i = \lambda_i$ for all $i \in I$.

Also, conditional on $X_{m+n} = i$, we can use the Markov property for the original chain, see that X_{n+m+1} indep of $X_m, \dots, X_{m+n-1}$.

Thm Suppose I is finite. Suppose that for some $i \in I$, π_j,

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

Then $\pi = (\pi_j)_{j \in I}$ is invariant, a distribution.

Proof (π_j) is a distribution:

$$\sum_j \pi_j = \sum_j \lim_{n\to\infty} p_{ij}^{(n)} \underset{\uparrow \text{finite}}{=} \lim_{n\to\infty} \sum_j p_{ij}^{(n)} = \lim_{n\to\infty} 1 = 1$$

(π_j) is invariant:

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} p_{ij}^{(n+1)} = \lim_{n\to\infty} \sum_k p_{ik}^{(n)} p_{kj}$$

L9.2

$$= \sum_k \left(\lim_{n\to\infty} p_{ik}^{(n)} \right) p_{kj} = \sum_k \pi_k p_{kj} = (\pi P)_j$$

<u>Rk</u> For the random walks on $\mathbb{Z}^d$, we have $p_{ij}^{(n)} \longrightarrow 0 \ \forall i,j$

The limit 0 is invariant but not a distribution

<u>Example</u> $P = \begin{pmatrix} -\alpha+1 & \alpha \\ \beta & -\beta+1 \end{pmatrix}$ We found $p_{11}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n$

$(1-\alpha-\beta)^n \downarrow 0$ as $n \to \infty$

$\alpha, \beta > 0$ So $P^{(n)} \longrightarrow \begin{pmatrix} \beta/\alpha+\beta & \alpha/\alpha+\beta \\ \beta/\alpha+\beta & \alpha/\alpha+\beta \end{pmatrix}$,

and $(\beta/\alpha+\beta, \ \alpha/\alpha+\beta)$ is invariant ö "by Theorem".

<u>Example</u>

$$\begin{tikzcd} & 1 \arrow[d, "1"] \\ 3 \arrow[ur, "1/2"] & 2 \arrow[l, "1/2"] \end{tikzcd}$$

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

Using $\pi_1 + \pi_2 + \pi_3 = 1$ gives $(1/5, \ 2/5, \ 2/5)$ ö

For each state k, let $\gamma_i^k = \mathbb{E}_k \sum_{n=0}^{T_k - 1} \mathbb{1}_{X_n = i} =$ expect times through i between k-visits

<u>Thm</u> Let P be <u>irreducible</u> and <u>recurrent</u>. Then

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

(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,k

<u>Proof</u> (a) clear from def^n

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

But also $\mathbb{P}_k[T_k < \infty, X_0 = X_{T_k} = k] = 1$.

$$\Rightarrow \gamma_j^k = \mathbb{E}_k \sum_{n=0}^{T_k-1} \mathbb{1}_{X_n = j} = \mathbb{E}_k \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n = j, \, n \le T_k\}}$$

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

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

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

$$\mathbb{P}_k[X_{n-1} = i, X_n = j, n \le T_k] = \mathbb{P}_k[X_{n-1} = i, n \le T_k] \ \mathbb{P}_i[X_1 = j]$$

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

$$= \sum_{i\in I} p_{ij} \ \mathbb{E}_k \sum_{n=0}^{T_k-1} \mathbb{1}_{X_n = i} = \sum_{i\in I} p_{ij} \gamma_i^k \qquad ö$$

(c) P irreducible

$\Rightarrow \exists n,m$ s.t. $p_{ik}^{(n)} p_{kj}^{(m)} > 0$

$\Rightarrow \gamma_i^k = \sum_j \gamma_j^k p_{ji}^{(m)} \ge \gamma_k^k p_{ki}^{(m)} = p_{ki}^{(m)} > 0$

$1 = \gamma_k^k \ge \gamma_i^k p_{ik}^{(n)}$

$\Leftarrow \ \Rightarrow \gamma_i^k \le \frac{1}{p_{ik}^{(n)}} < \infty$

<u>Thm</u> P irred $\Rightarrow \lambda \geq \gamma^k$ for any invariant λ with $\lambda k = 1$ ← Proof ?

P irred, rec $\Rightarrow \lambda = \gamma^k$

<u>Example</u> The simple symmetric random walk on $\mathbb{Z}$ is recurrent, irreducible

The measure $(\pi_i)_{i\in\mathbb{Z}}$ where

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

is invariant:

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

By the last thm, every invariant measure is a multiple of π.

Since $\sum_i \pi_i = \infty$, there is no invariant distribution

<u>Example</u> The simple symm. random walk on $\mathbb{Z}^3$ has an invariant measure, but is transient.

The <u>expected return time</u> is defined by

$$m_i = \mathbb{E}_i[T_i] \quad (\text{may be } \infty).$$

<u>Def</u> · i is <u>positive recurrent</u> if $m_i < \infty$

· i is <u>null recurrent</u> if i is recurrent but $m_i = \infty$

<u>Rk</u> +ve rec $\not\Leftarrow\!\Rightarrow$ rec

<u>Thm</u> Let P? be an irreducible MC. TFAE

(a) Every state is +ve rec

(b) Some state is +ve rec

(c) P has an invariant distribution π

Moreover, when (c) holds the invariant distro is $\pi_i = \frac{1}{\mu_i} = \frac{1}{m_i}$

<u>Proof</u> (a) $\Rightarrow$ (b) ✓

(b) $\Rightarrow$ (c) Since i is +ve rec (so rec), γ^i is an invariant measure

$$\sum_{j\in I} \gamma^i_j = \mathbb{E}_i \sum_{n=0}^{T_i - 1} 1 = \mathbb{E}_i[T_i] = m_i < \infty$$

So $\pi_j = \dfrac{\gamma^i_j}{m_i}$ is an invariant distribution.

L10. 2

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

Indeed, since π is invariant and P is irreducible,

$$\pi_k = (\pi P^n)_k = \sum_{j \in I} \pi_j p_{jk}^{(n)} \geq \pi_i p_{ik}^{(n)} > 0 \quad \text{for some } n$$

Now set $\lambda_i = \frac{\pi_i}{\pi_k}$.

Then λ is an invariant measure, and $\lambda_k = 1$.

So $\lambda \geq \gamma^k$ by Thm, and hence

$$m_k = \sum_{i \in I} \gamma_i^k \leq \sum_{i \in I} \lambda_i = \frac{1}{\pi_k} \sum_{i \in I} \pi_i = \frac{1}{\pi_k} < \infty$$

Thus k is positive recurrent.

Finally, knowing that P is recurrent, if (c) holds, then

$$m_k = \sum_{i \in I} \gamma_i^k = \sum_{i \in I} \lambda_i = \frac{1}{\pi_k}$$

"Last time" <u>Thm</u> Let P be irred, and λ an invariant measure 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, i.e. $\lambda = \lambda P$,

$$\lambda_j = \sum_{i_1 \in I} \lambda_{i_1} p_{i_1 j} = \sum_{i_1 \neq k} \lambda_{i_1} p_{i_1 j} + \underset{1}{\cancel{\lambda_k}} p_{kj}$$

$$= \sum_{i_1 \neq k} \Big(\underbrace{\sum_{i_2 \neq k} \lambda_{i_2} p_{i_2 i_1}}_{\lambda_k p_{k i_1}} \Big) p_{i_1 j} + p_{kj}$$

$$= \sum_{i_1, i_2 \neq k} \lambda_{i_2} p_{i_2 i_1} p_{i_1 j} + p_{k i_1} p_{i_1 j} + p_{kj}$$

$$\cdots = \sum (\qquad) + \mathbb{P}_k [\ \underbrace{\cancel{\quad\quad}}_{\text{some steps}}\ ] \quad ?$$

For $j \neq k$, $\lambda_j \geq \sum_{i=1}^{n} \mathbb{P}_k [X_i = j, T_k \geq i]$

$$\xrightarrow[n \to \infty]{} \gamma_j^k \quad \text{(makes sense)}$$

So $\lambda_j \geq \gamma_j^k$ as desired.

Now if P is recurrent, γ^k is invariant, so $\mu = \lambda - \gamma^k \geq 0$ is also invariant and $\mu_k = 0$.

So $0 = \mu_k = \sum_{j \in I} \mu_j p_{jk}^{(n)} \geq \mu_i p_{ik}^{(n)} \Rightarrow \mu_i = 0 \ \forall i$ by irred.

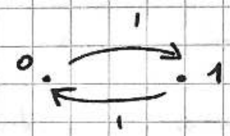
So $\gamma^k = \lambda$. □

<u>Example</u> asym random walk $q < p$

inv measure eqⁿ: $\pi_i = \sum \pi_j p_{ji} = \pi_{i-1} p + \pi_{i+1} q$ (not unique)

has general solⁿ $\pi_i = A + B\left(\frac{p}{q}\right)^i$

8. Convergence to equilibrium

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

$\pi = (\frac{1}{2}, \frac{1}{2})$ is an invariant distribution

But $P^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow P^{2n} = I, \; P^{2n+1} = P$

So P^n does not converge

Def A state i is <u>aperiodic</u> if $p_{ii}^{(n)} > 0$ for all n sufficiently large.

P is aperiodic if all states are aperiodic

Lemma Let P be irreducible. If i is aperiodic, then for all states j, k,

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

In particular, all states are aperiodic.

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

So $p_{jk}^{(r+n+s)} \geq p_{ji}^{(r)} p_{ii}^{(n)} p_{ik}^{(s)} > 0$ for all n sufficiently large. $\square$

Thm Let P be irreducible and aperiodic. Suppose that π is an invariant distribution for P. Let λ be any distribution and consider X_n given by Markov(λ, P). Then for all $j \in I$,

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

In particular, $p_{ij}^{(n)} \longrightarrow \pi_j$ as $n \longrightarrow \infty$.

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

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

<u>Claim</u>: $\mathbb{P}[T < \infty] = 1 \qquad (*)$

Let $W_n = (X_n, Y_n)$. This is a Markov chain with state space $I \times I$ with transition probabilities $\tilde{p}_{ik,jl} = p_{ij} p_{kl}$, and initial distribution $\tilde{\lambda}_{ik} = \lambda_i \pi_k$.

Since P is aperiodic, by the Lemma, for all i, j, k, l,

$$\tilde{p}_{ik,jl}^{(n)} = p_{ij}^{(n)} p_{kl}^{(n)} > 0 \quad \text{for all } n \text{ sufficiently large}$$

So $\tilde{P}$ is irreducible.

$\tilde{P}$ has invariant distribution $\tilde{\pi}_{ik} = \pi_i \pi_k$, so by last time, $\tilde{P}$ is positive recurrent. The <u>Claim</u> follows.

Indeed, T is first passage time of W_n to (b,b).

Now $\mathbb{P}[X_n = j] = \mathbb{P}[X_n = j, T > n] + \underbrace{\mathbb{P}[X_n = j, n \geq T]}_{\mathbb{P}[Y_n = j, n \geq T]}$ MAD

$$= \mathbb{P}[X_n = j, T > n] + \underbrace{\mathbb{P}[Y_n = j]}_{\pi_j} - \mathbb{P}[Y_n = j, T > n]$$

$$\therefore |\mathbb{P}[X_n = j] - \pi_j| = |\mathbb{P}[X_n = j, T > n] - \mathbb{P}[Y_n = j, T > n]|$$
$$\leq \mathbb{P}[T > n] \longrightarrow 0 \quad \text{as } n \to \infty$$

□

<u>Example</u> (cont'd)

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \pi = (\tfrac{1}{2}, \tfrac{1}{2})$$

If X is Markov(δ_0, P) and Y is Markov(π, P), then with probability $\frac{1}{2}$, $W_0 = (0, 1)$ and so W_n never meet. It's irreducible.

9. Time reversal

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

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

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

9. Time Reversal

Thm Let P be irred. with inv. 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

$$\pi_j \hat{p}_{ji} = \pi_i p_{ij} \quad \text{and } \hat{P} \text{ is irreducible with inv. distribution } \pi.$$

Pf $\hat{P}$ is a stochastic matrix since

$$\sum_{i\in I} \hat{p}_{ji} \underset{\pi_j>0}{=} \frac{1}{\pi_j} \sum_{i\in I} \pi_i p_{ij} \underset{\text{invariant}}{=} \frac{1}{\pi_j} \pi_j = 1 .$$

π is invariant for $\hat{P}$ since

$$\sum_{j\in I} \pi_j \hat{p}_{ji} = \sum_{j\in I} \pi_i p_{ij} \underset{P \text{ stochastic}}{=} \pi_i .$$

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

$$\mathbb{P}[Y_0 = i_0, \dots, Y_N = i_N] \underset{\text{def}^n}{=} \mathbb{P}[X_0 = i_N, \dots, X_N = i_0]$$

$$= \underbrace{\pi_{i_N} p_{i_N i_{N-1}}}_{\pi_{i_{N-1}} \hat{p}_{i_{N-1} i_N}} \cdots p_{i_1 i_0} \quad \text{\& repeat}$$

$$= \pi_{0} \hat{p}_{i_0 i_1} \cdots \hat{p}_{i_{N-1} i_N}$$

$\hat{P}$ is irreducible since

$$\forall i,j \in I \ \exists\, i_0,\dots,i_n \text{ with } i_0 = i,\ i_n = j,\quad p_{i_0 i_1} \cdots p_{i_{n-1} i_n} > 0$$

$$\Rightarrow \hat{p}_{i_n i_{n-1}} \cdots \hat{p}_{i_1 i_0} > 0 \quad (\text{use } \pi > 0)$$

□

Def^n A stochastic matrix P and a measure λ are in detailed balance if $\lambda_i p_{ij} = \lambda_j p_{ji} \ \forall i,j \in I$.

Lemma If P, λ are in detailed balance, then λ is invariant for P.

Pf $\sum_i \lambda_i p_{ij} = \sum_i \lambda_j p_{ji} = \lambda_j$ from stochasticity

□

Def^n If P is irreducible and $(X_n)_{n\ge 0}$ is Markov (λ, P), then (X_n) is reversible if, for all N, (X_{N-n}) is also Markov (λ, P).

Thm Let P be irreducible, λ be a distribution. Let (X_n) be Markov (λ, P). Then TFAE

(a) (X_n) is reversible

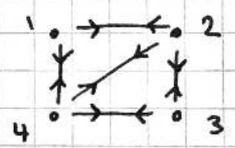
(b) P, λ are in detailed balance

<u>Pf</u> Both (a), (b) imply λ is invariant for P.

Indeed, reversibility implies $\forall N$, X_N has distribution λ.

Now, by the previous theorem, both are equivalent to $P = \hat{P}$. $\square$

<u>Example</u> (Random walk on a graph)

$G =$ 

graph

Produce Markov chain by setting

$$p_{ij} = \frac{1}{v_i} \quad \text{for each neighbour } j$$

(Assume valencies finite)

If G is connected, then P is irreducible.

And P is in detailed balance with v_i.

Indeed, if $p_{ij} = 0$, so is $p_{ji} = 0$. Else $p_i v_{ij} = 1 = p_j v_{ji}$.

10. Ergodic theorem

<u>Thm</u> (*) Let $V_i(n) = \sum_{k=0}^{n-1} \mathbb{1}_{\{X_k = i\}}$ = # visits to i before n.

Let P be irreducible and λ a distribution, and X = Markov(λ, P).

Then $\mathbb{P}\left[\frac{V_i(n)}{n} \to \frac{1}{m_i} \text{ as } n \to \infty\right] = 1$ where $m_i = \mathbb{E}_i[T_i]$.

In particular, if P is +ve recurrent with invariant π_i, tend to π_i.

<u>We need</u>: (Strong Law of Large Numbers) Let (Y_i) be a sequence of independent identically distributed random variables with expectation μ.

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

MAPKOV