Markov chains (Chapter 4)

Each rv $X_n \in S = \{1, \dots, M\}$ has finite support.
X_n: discrete rv

$(X_n)_{n\geq 0}$ n: discrete-time $\{X_n\}_{n\geq 0}$
stochastic process

Def: A discrete-time Markov chain $(X_n)_{n\geq 0}$ is an integer-time stochastic process for which X_n depends on its most recent rv X_{n-1},

i.e., $\forall n \geq 1$, $i, j, k, l \in S$

$$\Pr(X_n = j \mid X_{n-1} = i, X_{n-2} = k, \dots, X_0 = l) = \Pr(X_n = j \mid X_{n-1} = i)$$

for all conditioning events $\{X_{n-2} = k\}, \dots, \{X_0 = l\}$

$$X_0 \to X_1 \to X_2 \to \cdots \to X_{n-1} \to X_n \to X_{n+1}$$

health at age n-1   health at age n

Homogeneous Markov chain is one in which $\Pr(X_n = j \mid X_{n-1} = i)$ does not depend on n (only depends on i & j)

$$P_{ij} = \Pr(X_n = j \mid X_{n-1} = i) \quad \forall n \geq 1$$

X_0 : initial state of the Markov chain
has an arbitrary prob. dist^n over S.

S : state space (finite set).

$$\boxed{p_{X_0}(i) = \Pr(X_0 = i)}$$

Ex: $(Z_n)_{n \geq 0}$ $\quad Z_n$: integer-valued rv $S = \{1, \dots, M\}$
Z_n depends on the past $m \geq 1$ rvs
$Z_{n-1}, Z_{n-2}, \dots, Z_{n-m}$

$$\Pr(Z_n \mid \underbrace{Z_{n-1}, \dots, Z_0}_{n \text{ rvs}}) = \Pr(Z_n \mid \underbrace{Z_{n-1}, \dots, Z_{n-m}}_{m \text{ rvs.}}) \quad \text{— } (*)$$
$n > m$

Is this a first-order Markov chain? No!

Define $(Z_{n-1}, Z_{n-2}, \cdots Z_{n-m})$ as the state of the process at time $n-1$ (state space S^m)

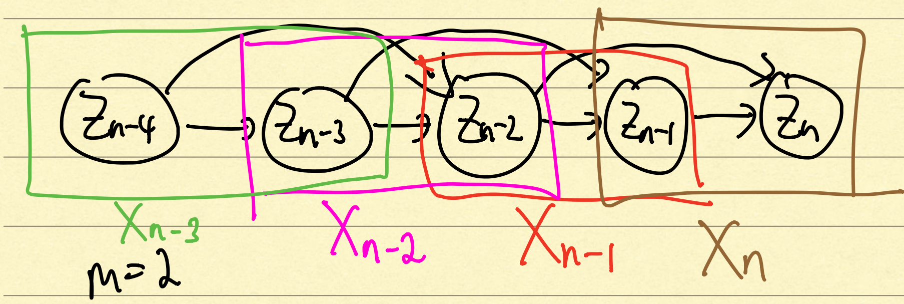
$$\Pr(\overbrace{Z_n, Z_{n-1}, \dots, Z_{n-m+1}}^{X_n} \mid Z_{n-1}, \dots, Z_0)$$
$$= \Pr(Z_n, Z_{n-1}, \dots, Z_{n-m+1} \mid Z_{n-1}, \dots, Z_{n-m+1}, Z_m) \quad (+)$$

If we define $\underline{X_{n-1} = (Z_{n-1}, \dots, Z_{n-m})}$ for each n,
then $(+)$ reduces to

$$\Pr(X_n \mid X_{n-1}, \dots, X_{m-1}) = \Pr(X_n \mid X_{n-1})$$

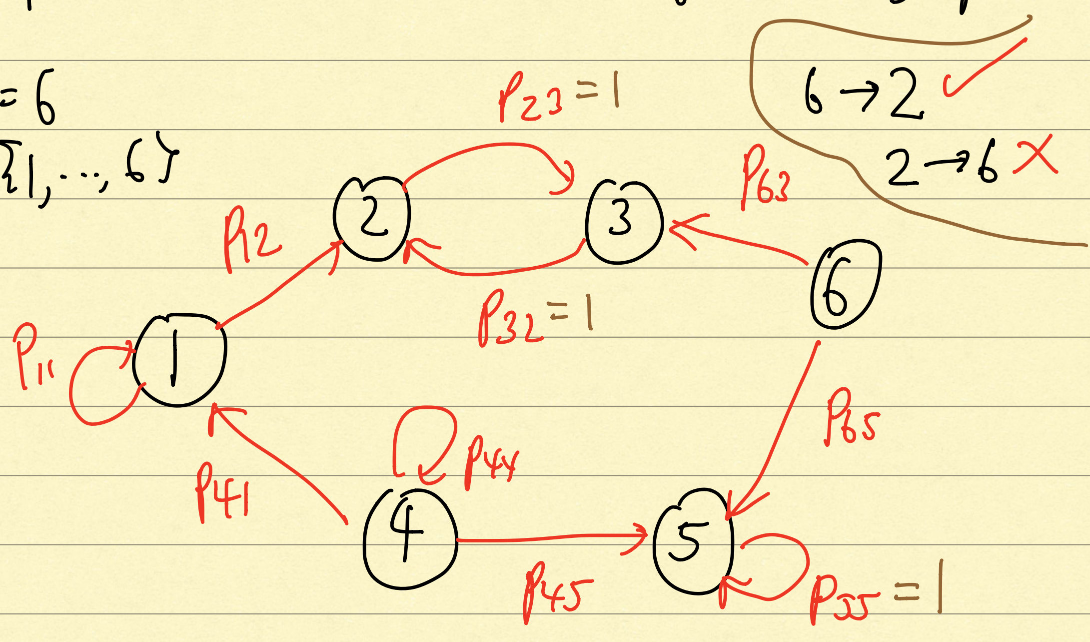
Rmk: By expanding the state space from S to $S^{\times m}$, i.e., considering m-tuples of (Z_n), we convert the dependence on past m rvs to past 1 rv.

If define $X_{n-1} = (Z_{n-1}, \dots, Z_{n-m})$, then $(X_n)_{n \geq m}$ forms a first-order MC.



Description of Markov chains using directed graphs.

$M = 6$
$S = \{1, \dots, 6\}$



$$P_{ij} = Pr(X_n = j \mid X_{n-1} = i) \qquad \sum_{j \in S} P_{ij} = 1$$

Can form a matrix of transition probabilities

$$[P] = \begin{pmatrix} P_{11} & \cdots & P_{16} \\ \vdots & \ddots & \vdots \\ P_{61} & \cdots & P_{66} \end{pmatrix} \qquad \begin{array}{c} \sum_{j=1}^{6} P_{1j} = 1 \\ \vdots \\ \sum_{j=1}^{6} P_{6j} = 1 \end{array}$$

$\exists$ directed arc from $i \in S$ to $j \in S$ iff $P_{ij} > 0$.

Row sums of $[P]$ are all one $\Rightarrow$ Row stochastic matrix

Rmk: If X_0 has PD $P_{X_0}(j) = Pr(X_0 = j)$, we can form an initial prob vector $(P_{X_0}(1), \dots, P_{X_0}(M))$; this is a row vector.

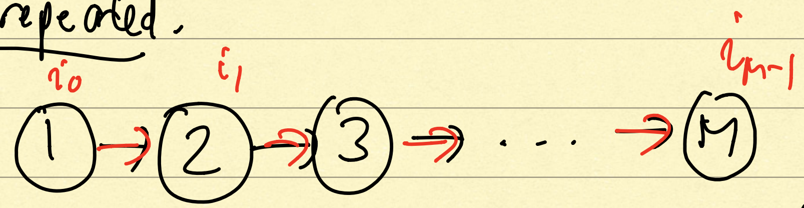
X_1 has distribution?

$$(P_{X_0}(1), \dots, P_{X_0}(M))\,[P] = (P_{X_1}(1) \dots P_{X_1}(M))$$

Classification of States

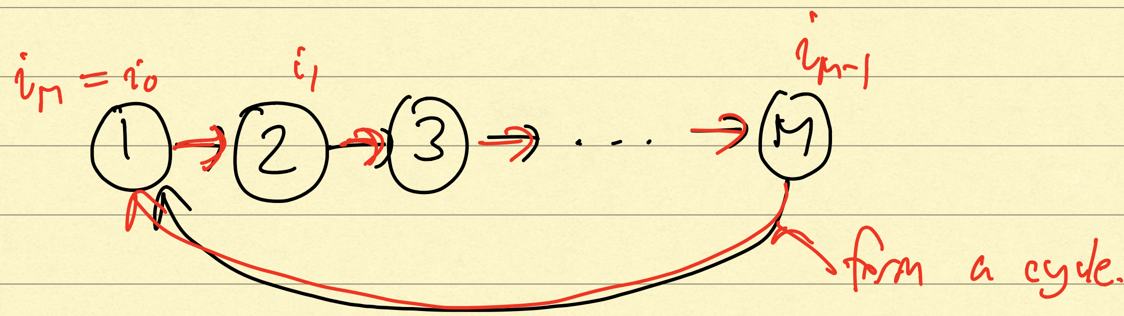
Def: An (n-step) <u>walk</u> is an ordering of a string of nodes $(i_0, i_1, \dots, i_n)$, $n \geq 1$ in which there's a directed arc from i_m to i_{m+1} for each $0 \le m \le n-1$.

Def: A *path* is a walk in which no nodes are repeated.



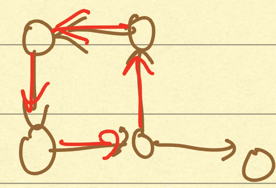
$M-1$ Longest path.

Def: A *cycle* is a walk in which the first and last nodes are equal and no other node is repeated



Rmk: A walk can start and end at the same node; A path cannot.

# of steps in a walk $= \infty$
# of steps in a path $\le M-1$
# of steps in a cycle $\le M$.



Def: A state $j \in S$ is <u>accessible</u> from another state $i \in S$ if $\exists$ a walk from i to j.

We write this as $i \longrightarrow j$.

Ex: $1 \longrightarrow 3$, $5 \not\rightarrow 3$, $2 \rightarrow 2$, $6 \not\rightarrow 6$

Fact: Suppose a walk $(i_0, i_1, \ldots, i_n)$ exists from node i_0 to i_n, then $p_{i_0 i_1}, p_{i_1 i_2}, p_{i_2 i_3}, \ldots, p_{i_{n-1} i_n} > 0$.

Pf: $\Pr(X_n = i_n \mid X_0 = i_0)$

$$= \sum_{(i_1, i_2, \ldots, i_{n-1}) \in S^{n-1}} \Pr(X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1 \mid X_0 = i_0)$$

$$= \sum_{i_1, i_2, \ldots, i_{n-1}} \Pr(X_n = i_n \mid X_{n-1} = i_{n-1}) \cdot \Pr(X_{n-1} = i_{n-1} \mid X_{n-2} = i_{n-2}) \cdots \Pr(X_1 = i_1 \mid X_0 = i_0)$$

$$\left[\begin{aligned} &\Pr(X_2 = i_2, X_1 = i_1 \mid X_0 = i_0) \\ &= \Pr(X_2 = i_2 \mid X_1 = i_1, \cancel{X_0 = i_0}) \Pr(X_1 = i_1 \mid X_0 = i_0) \\ &= \Pr(X_2 = i_2 \mid X_1 = i_1) \Pr(X_1 = i_1 \mid X_0 = i_0) \end{aligned} \right]$$

$$\geq \Pr(X_n = i_n \mid X_{n-1} = i_{n-1}) \cdot \Pr(X_{n-1} = i_{n-1} \mid X_{n-2} = i_{n-2}) \cdots \Pr(X_1 = i_1 \mid X_0 = i_0)$$

$$= p_{i_{n-1} i_n} p_{i_{n-2} i_{n-1}} \cdots p_{i_0 i_1} > 0.$$

All probabilities $p_{i_{m-1} i_m} > 0 \quad \forall\ 1 \le m \le n.$

If $Pr(X_n = i_n \mid X_0 = i_0) > 0$, $\exists$ n-step walk from $i_0 \in S$ to $i_n \in S$.

We write $Pr(X_n = j \mid X_0 = i) = p_{ij}^n = [P^n]_{ij}$

[p_{ij}^n is not the n^{th} power of p_{ij}]

Fact: If $\exists$ n-step walk from i to j & $\exists$ m step walk from j to k, then $\exists$ $(m+n)$-step walk from i to k.

i.e., $\underline{p_{ij}^n > 0, \quad p_{jk}^m > 0} \Rightarrow p_{ik}^{n+m} \ge 0.$

i.e., $i \to j$ & $j \to k \Rightarrow i \to k$.

Pf: $$Pr(X_{m+n} = k \mid X_0 = i) = \sum_{l \in S} Pr(X_{m+n} = k \mid X_n = l, \cancel{X_0 = i}) \cdot Pr(X_n = l \mid X_0 = i)$$

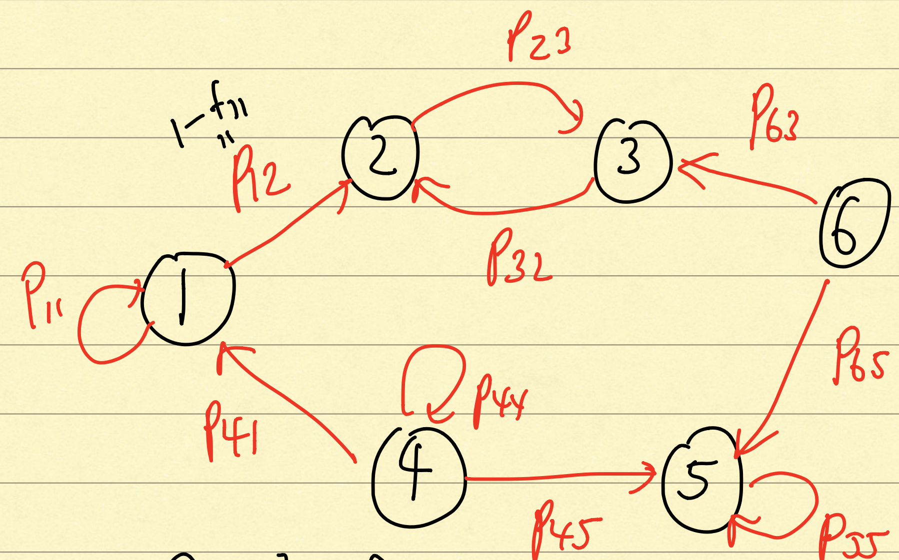
$$\geq Pr(X_{m+n} = k \mid X_n = j)\, Pr(X_n = j \mid X_0 = i)$$

$$= p_{jk}^m \, p_{ij}^n > 0.$$

Def: Two states i & j <u>communicate</u> $(i \leftrightarrow j)$ if $i \to j$ & $j \to i$.

Fact: If $i \leftrightarrow j$ & $j \leftrightarrow k$, then $i \leftrightarrow k$.

Def: A <u>class</u> $C \subseteq \{1, \cdots, M\}$ is a non-empty subset of states s.t. $\forall i \in C$, each state $j \neq i$ satisfies $j \in C$ iff $i \leftrightarrow j$ & $j \notin C$ iff $i \not\leftrightarrow j$.



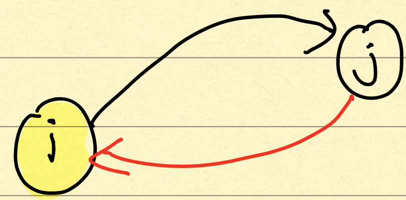
Claim: $C_1 = \{2, 3\}$ forms a class of states.
Claim: $C_2 = \{1\}$ forms a class of states.
Claim: $C_3 = \{4\}$, $C_4 = \{5\}$, $C_5 = \{6\}$.
(You verify)

Def: A <u>recurrent</u> state i is one that is accessible from all other states that are accessible from i.

i.e., i is recurrent $\Leftrightarrow$ $i \to j$ implies that $j \to i$

There is no possibility of going to j from which there's no return.

If one enters a recurrent state i, we return to that state i eventually w.p. 1.

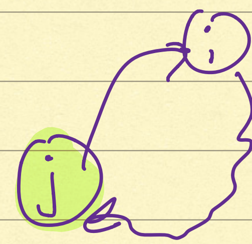


We will visit i infinitely often.

Def: A <u>transient</u> state is one that is no recurrent.

Prop: State j is recurrent iff

$$\sum_{n=1}^{\infty} P_{jj}^{n} = \infty$$



Pf: The number of visits to state j (which is recurrent) if start from j is infinite & thus it has infinite expectation.

$\forall L > 0$

$$E[\# \text{ of visits to } j \mid X_0 = j] \geq L$$

Let $I_n = \begin{cases} 1 & X_n = j \\ 0 & \text{else.} \end{cases}$ conditioned on $X_0 = j$.

$\sum_{n=0}^{\infty} I_n$ refers to the total # of visits to j.

$$E\left[\sum_{n=0}^{\infty} I_n \mid X_0 = j\right] = \sum_{n=0}^{\infty} E[I_n \mid X_0 = j]$$

$$= \sum_{n=0}^{\infty} \Pr(X_n = j \mid X_0 = j)$$

$$= \sum_{n=0}^{\infty} P_{jj}^n$$

Thus result follows if j is recurrent. $\blacksquare$

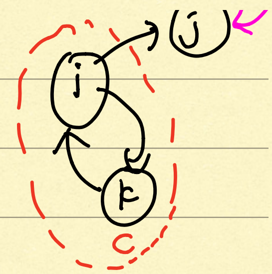
Rmk: If j is transient, every time MC visits j, there's a positive prob say $1-f_{jj}$ ($f_{jj} \in (0,1)$) that it'll never return to j.

The total # of visits to j is geometric with finite mean $\frac{1}{1-f_{jj}} < \infty$.

Thm: In a finite-state Markov chain, either all states in a class are transient or all are recurrent.

Pf: Let i be transient (i.e., $\exists j$ s.t. $i \to j$ but $j \not\to i$). Suppose i & k are in the same class C.

This means that $i \leftrightarrow k$.



We have $k \to i$ & $i \to j$ so $k \to j$.

Now if, to the contrary, if $j \to k$, the walk from j to k can be extended to i ($j \to i$)

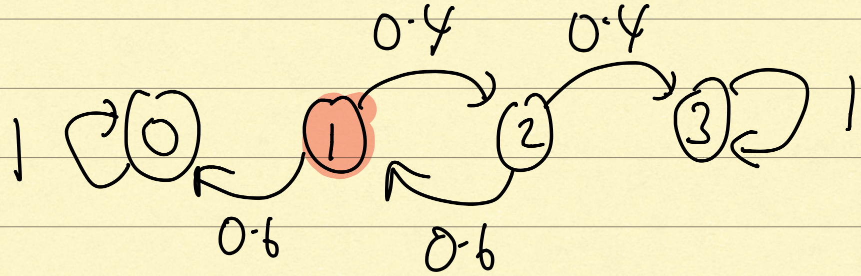
This however, contradicts the fact that $j \not\to i$.

Thus there is no walk from j to k ($j \not\to k$)
Hence k is transient
($\because \exists j$ s.t. $k \to j$ but $j \not\to k$)

Hence all states in a class are transient if any of them are.

All states in a class either all recurrent or all transient.

Ex: Gambler's Ruin Example $S = \{0, 1, 2, 3\}$.



State Transition Matrix.

$$[P] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 \\ 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Classes? Claim: $C_1 = \{0\}$, $C_2 = \{1, 2\}$, $C_3 = \{3\}$.
(You verify)

$C_2 = \{\underset{t}{1}, \underset{t}{2}\}$ is a transient class.

C_1 and C_3 are recurrent classes.

Def: The period of a state $i \in S$, denoted as $d(i)$ is the gcd (greatest common divisor) of those integers $n \in \mathbb{N}$ for which

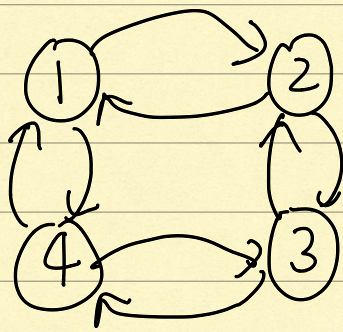
$$P_{ii}^n > 0$$

$$d(i) = \gcd\{n : P_{ii}^n > 0\}.$$

If the period i, $d(i) = 1$, then i is aperiodic.

Otherwise it is periodic with period $d(i)$.

Eg:



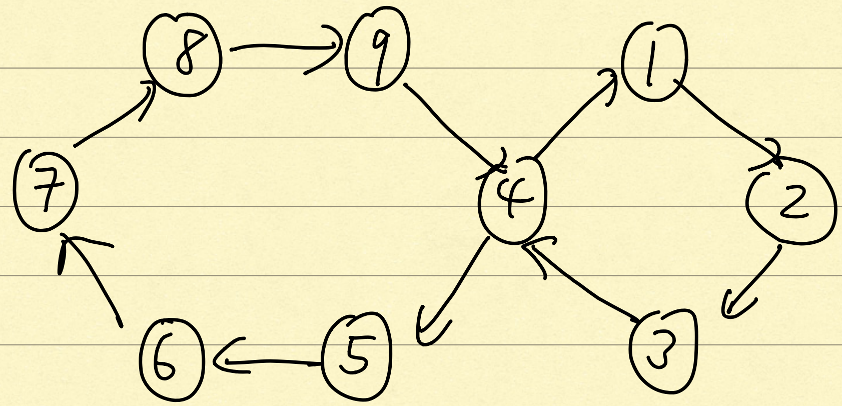
What's the period of i?

$i=1$

$P_{11}^{2} > 0,\ P_{11}^{4} > 0,\ P_{11}^{6} > 0, \dots$

$$d(1) = \gcd\{n : P_{11}^{n} > 0\}$$
$$= \gcd\{n : n \text{ is even}\}$$
$$= 2$$

Ex:



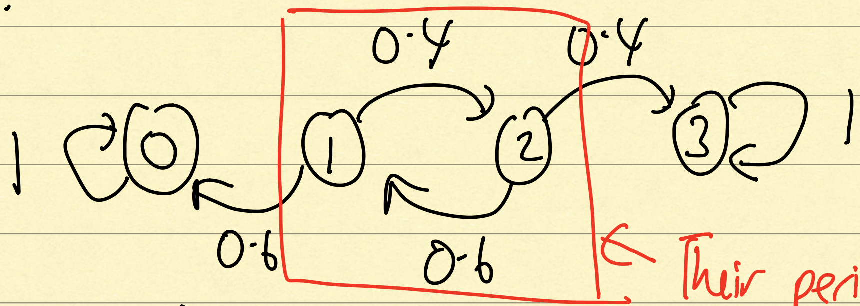
$d(1)$? $\quad P_{11}^{4} > 0,\ P_{11}^{8} > 0,\ P_{11}^{10} > 0,\ P_{11}^{12} > 0, \dots$

$\shortparallel$

$\gcd\{4, 8, 10, 12, \dots\} = 2$

Exercise: $\quad d(i) = 2 \quad \forall i \in \{1, \dots, 9\}$.

Eg:



Their period is 2.

C

$C = \{1, 2\}$

$d(1) = \gcd\{n : P_{11}^n > 0\} = \{2, 4, 6, 8, \dots\}$
$= 2$

$d(2) = 2.$

<u>Thm</u>: For any Markov chain, all states in the same class have the same period.

Pf: Let i & j be a pair of states in the same class say C.

$\Leftrightarrow$ $i \leftrightarrow j \Rightarrow \exists n, m \in \mathbb{N}$ s.t. $P_{ij}^n > 0, P_{ji}^m > 0$.

Since $\exists$ walk of length n from i to j &

$\exists$ walk of length m from j to i, $\exists$ walk of length $n+m$ from i to i (via j).

$\Rightarrow$ $d(i) \mid n+m$ ($n+m$ is divisible by $d(i)$)

$\overset{\parallel}{\gcd\{\ell : P_{ii}^{\ell} > 0\}}$ ($n+m$ is a multiple of $d(i)$)

Let t be any integer such that $P_{jj}^{t} > 0$
(does t exist? Yes.
Such a t exists but it is arbitrary).

There is a walk of length $n+m+t$ from i to i

$d(i) \mid n+m+t$

$kd(i) = n+m$
$\ell d(i) = n+m+t$
$(\ell - k) d(i) = t$

We have $d(i) \mid n+m$ & $d(i) \mid n+m+t$.

$d(i) \mid t$

This is true for every t s.t. $P_{jj}^{t} > 0$.

$\boxed{d(i) \mid d(j)}$

$d(j) := \gcd\{t : P_{jj}^{t} > 0\}$

$\{m, n\} \equiv \{t : P_{jj}^{t} > 0\}$
$a \equiv d(i)$ $\quad d \equiv d(j)$

Fact: m, n s.t. $\gcd(m, n) = d$.
a is any common divisor of m & n, then $a \mid d$.

Flip the roles of i & j

$\Rightarrow$ $\boxed{d(j) \mid d(i)}$ $\Rightarrow$ $d(i) = d(j)$

Rmk: We can say class C is transient

We can say class C is recurrent

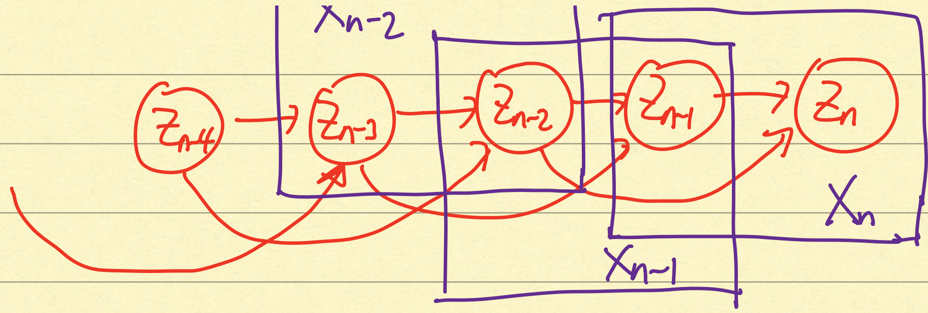
We can say class C has period d.

Quiz 2

- 12th March
- Everyone needs to show up in person LT1
- 6pm – 7pm
- 1 sheet of paper

I will examine you on conditional arrivals

$S_1, \ldots, S_n \mid N(t) = n \overset{d}{=}$



$m=2$

$$p(X_n \mid X_{n-1}, X_{n-2}, \ldots, X_1)$$

$$= p(z_n, z_{n-1} \mid z_{n-1}, z_{n-2}, \ldots, z_0)$$

$$= p(z_n \mid z_{n-1}, z_{n-2}, \ldots, z_0)\, \underbrace{p(z_{n-1} \mid z_n, z_{n-1}, \ldots, z_0)}_{1}$$

$\because$ (if we condition on z_{n-1}, z_{n-1} is deterministic)

$$= p(z_n \mid z_{n-1}, z_{n-2})$$

$$= p(z_n, z_{n-1} \mid z_{n-1}, z_{n-2})$$

$$= p(X_n \mid X_{n-1})$$