Review: <u>Ergodic</u> Markov chain $\equiv$ One that has 1 class of states that is both aperiodic & recurrent. (i.e., cannot have transient states)

<u>Ergodic unichain</u> $\equiv$ 1 ergodic class + possibly some transient states.

<u>Thm</u>: $[P]$ is ergodic, $[P^h] > 0 \quad \forall\, h \geq (M-1)^2+1$.

<u>Lem</u>: For any Markov chain $[P]$,

$$\max_{\ell} P_{\ell j}^{n+1} \leq \max_{\ell} P_{\ell j}^{n} \quad \forall n \in \mathbb{N},\ j \in S$$

<u>Lem</u>: $[P] > 0$, then $\forall j$ $\quad \frac{1}{2} \geq \alpha := \min_{i,j} P_{ij} > 0$.

$$\max_{\ell} P_{\ell j}^{n} - \min_{\ell} P_{\ell j}^{n} \leq (1-2\alpha)^n \to 0.$$

$$\forall j \in S, \quad \underbrace{\lim_{n\to\infty} \max_{\ell} P_{\ell j}^{n}}_{\pi_j} = \lim_{n\to\infty} \min_{\ell} P_{\ell j}^{n} \geq \alpha > 0$$

ending $\qquad\qquad\qquad\qquad$ starting

$$\forall j \in S \qquad P_{ij}^{n} \to \pi_j \text{ as } n \to \infty \quad \forall i.$$

$$[P^n] \longrightarrow \begin{pmatrix} - \pi - \\ - \pi - \\ \vdots \\ - \pi - \end{pmatrix} = e\pi = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (\pi_1, \ldots, \pi_M)$$

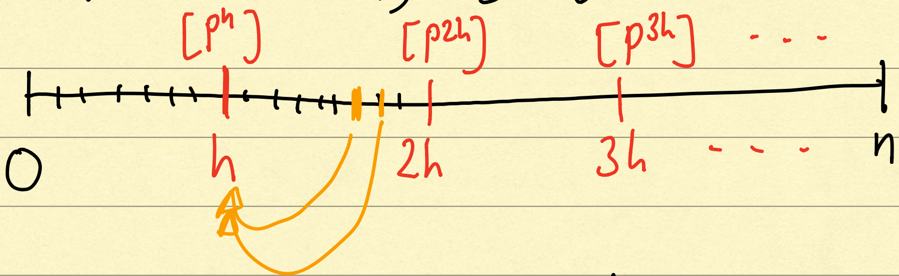
Convergence is geometrically fast.

Thm: $[P]$: ergodic Markov chain

Then $\exists$ unique SS vector $\underline{\pi > 0}$ satisfying $\pi = \pi[P]$,

$\forall j,\ \pi_j > 0$ $\qquad$ $\sum \pi_i = 1$.

Pf: $\quad \forall h \geq (M-1)^2 + 1, \quad [P^h] > 0.$



Applying previous lemma to $[P^h] > 0$,

$$\left| P_{ij}^{hm} - \pi_j \right| \leq (1-2\beta)^m, \quad m \in \mathbb{N}.$$

$$\beta := \min_{i,j} P_{ij}^h$$

$$\left| P_{ij}^{n} - \pi_j \right| \leq (1-2\beta)^{\lfloor n/h \rfloor} \qquad n \in \mathbb{N}.$$

$$\lim_{n\to\infty} P_{ij}^n = \pi_j \qquad \forall i, j \in S.$$

$$\lim_{n\to\infty} [P^n] = e\pi$$

Since $[P^n] \to e\pi$, π satisfies $\pi = \pi[P]$, the SS eqn.

[Uniqueness]

Let μ be any other SS vector, $\mu[P] = \mu$.

Then $\mu = \mu[P^n]$ $\forall n \in \mathbb{N}$.

$$\mu = \lim_{n\to\infty} \mu[P^n] = \mu \lim_{n\to\infty} [P^n] = (\overset{1}{\mu e})\pi = \pi$$

$$\sum_{i=1}^{M} \mu_i = 1, \quad (\mu_1, \dots, \mu_M)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^{M} \mu_i = 1.$$

$$\mu = \pi$$

SS vector is unique.

Ergodic Unichain ($\exists$ transient states)

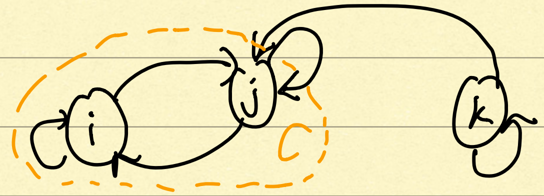
$T = \{1, \dots, t\}$ Transient States

$R = \{t+1, \dots, t+r\}$ Recurrent States

$$[P] = \begin{bmatrix} [P_T] & [P_{TR}] \\ [0] & [P_R] \end{bmatrix} \rightarrow e\pi$$

$$\pi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \pi_{R1} \\ \vdots \\ \pi_{Rr} \end{bmatrix}^T$$

Why? If i, j are recurrent states, there no possibility of leaving recurrent class in getting from i to j.



Suppose from i to j, we can pass through through k (a transient state)

$i \leftrightarrow j$ ($\because$ $i, j \in C$).

$i \to k \to j \Rightarrow k \to i$

$\Rightarrow i \leftrightarrow k \Rightarrow k$ is recurrent $\Rightarrow k$ is in the same class as i.

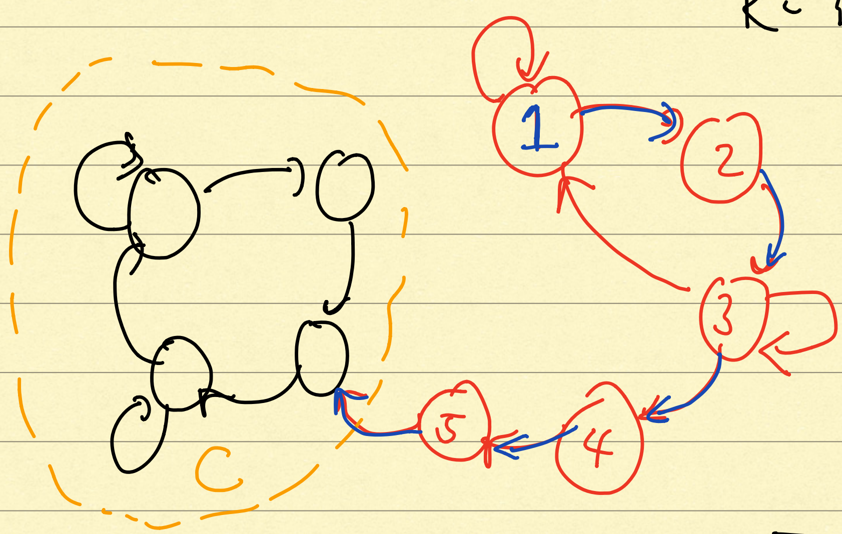
$\Rightarrow$ contradiction.

Objective: If initial state $\in T$, eventually recurrent class is entered & dist gradually converges SS of $[P_R]$; $\pi_i = 0 \ \forall i \in T$.

Fact: $\forall i \in T$, there exist a walk of length $\le t$ to a recurrent state.

($\because$ only t transient states) $T = \{1, \dots, t\}$

$R = \{t+1, \dots, t+r\}$



$\forall i \in T, \quad \sum_{j \in R} P_{ij}^t > 0 \Longleftrightarrow \sum_{j \in T} P_{ij}^t < 1.$

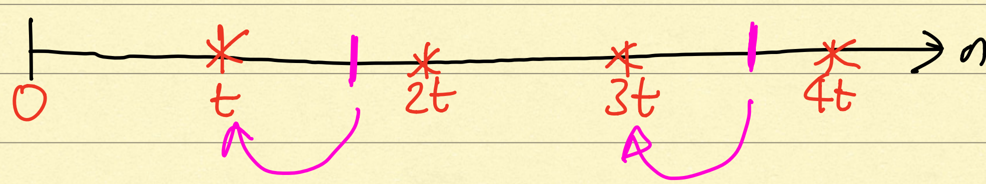
$$\boxed{\gamma := \max_{i\in T} \sum_{j\in T} P_{ij}^{t} < 1.}$$

Lem: $[P]$ unichain (not necessarily ergodic).
T: set of transient states $t = |T|$.

$$\forall n\in\mathbb{N},\ \max_{l\in T} \sum_{j\in T} P_{lj}^{n} \le \gamma^{\lfloor n/t \rfloor}$$

Intuition: If you start in a transient state $l\in T$, the prob. you stay in a transient state $j\in T$ is small ($\gamma<1$, $\gamma^{\lfloor n/t\rfloor} \to 0$).

Pf:



Let mt be an int. multiple of t ($m\in\mathbb{N}$).
$\forall i\in T$,

$$\sum_{j\in T} P_{ij}^{(m+1)t} = \sum_{j\in T}\left(\sum_{k\in T} P_{ik}^{t} P_{kj}^{mt}\right) \quad \text{—}(*)$$

$$= \sum_{k\in T} P_{ik}^{t}\left(\sum_{j\in T} P_{kj}^{mt}\right)$$

$(*)$ $\because$ no possibility of going from recurrent states i,j through transient state k.

$$\leq \left(\sum_{k\in T} p_{ik}^{t}\right)\left(\max_{l\in T} \sum_{j\in T} p_{lj}^{mt}\right)$$

$$\leq \left(\max_{i\in T} \sum_{k\in T} p_{ik}^{t}\right)\left(\max_{l\in T} \sum_{j\in T} p_{lj}^{mt}\right)$$

$$= \gamma \left(\max_{l\in T} \sum_{j\in T} p_{lj}^{mt}\right)$$

$$\underbrace{\max_{i\in T} \sum_{j\in T} p_{ij}^{(m+1)t}}_{g_{m+1}} \leq \gamma \underbrace{\left(\max_{l\in T} \sum_{j\in T} p_{lj}^{mt}\right)}_{g_m .}$$

$$g_1 = \sum_{j\in T} p_{ij}^{t} \leq \gamma$$

$$g_m \leq \gamma^m .$$

$$\max_{l\in T} \sum_{j\in T} p_{lj}^{\overset{=n}{mt}} \leq \gamma^{\overset{=\lfloor n/t \rfloor}{m}} ,$$

Have proved desired result $\forall$ n int. mult. of t.

$$\max_{l\in T} \sum_{j\in T} p_{lj}^{n} \leq \gamma^{\lfloor n/t \rfloor} \quad //$$

$\because$ $n \mapsto \sum_{j\in T} p_{lj}^{n}$ is non-increasing in n.

Consider $i \in T$ and final state $j \in R$.

$$P_{ij}^{2n} = \sum_{k \in T} P_{ik}^{n} P_{kj}^{n} + \sum_{k \in R} P_{ik}^{n} P_{kj}^{n} .$$

Upper Bound

$$P_{ij}^{2n} \leq \sum_{k \in T} P_{ik}^{n} + \sum_{k \in R} P_{ik}^{n} \max_{l \in R} P_{lj}^{n}$$

$$\leq \sum_{k \in T} P_{ik}^{n} + \max_{l \in R} P_{lj}^{n} \quad — (1).$$

Lower Bound

$$P_{ij}^{2n} \geq \sum_{k \in R} P_{ik}^{n} P_{kj}^{n} \geq \sum_{k \in R} P_{ik}^{n} \left(\min_{l \in R} P_{lj}^{n} \right)$$

$$= \left(1 - \sum_{k \in T} P_{ik}^{n} \right) \left(\min_{l \in R} P_{lj}^{n} \right)$$

$$\geq \min_{l \in R} P_{lj}^{n} - \sum_{k \in T} P_{ik}^{n} \quad — (2).$$

The steady-state prob. of recurrent state $j \in R$

$$\min_{l \in R} P_{lj}^{n} \leq \pi_j \leq \max_{l \in R} P_{lj}^{n} \quad — (3)$$

By combining (1) & (2) & (3)

$$\left|P_{ij}^{2n} - \pi_j\right| \le \underbrace{(1-2\beta)^{\lfloor n/h \rfloor}}_{\text{ergodic class}} + \underbrace{\gamma^{\lfloor n/t \rfloor}}_{\text{transience}} .$$

$\forall i \in S$

$$\left|P_{ij}^{n} - \pi_j\right| \le (1-2\beta)^{\lfloor n/2h \rfloor} + \gamma^{\lfloor n/2t \rfloor} \to 0.$$

Thm: $[P]$ ergodic unichain

$$\lim_{n\to\infty} [P^n] = e\pi.$$

π is the SS vector formed from the recurrent class $[P_R]$ and padded by 0's for the transient states

$$\pi = [\overset{1}{0}, \overset{\cdots}{\ldots}, \overset{t}{0}, \underbrace{\pi_{R1}, \ldots, \pi_{Rr}}_{\substack{\text{left-eigenvector corresponding} \\ \text{to e-value } 1 \text{ of } [P_R].}}]$$

$$\boxed{\pi_R [P_R] = \pi_R > 0}.$$

Eigenvalues & Eigenvectors of Stochastic Matrices.

Def: A row vector π is a left-eigenvector of $[P]$ with eigenvalue λ if $\pi \neq 0$, &

$$\pi[P] = \lambda\pi \iff \sum_i \pi_i P_{ij} = \lambda \pi_j \quad \forall j$$

Fact: Ergodic unichain, $\exists!$ SS vector π that is a left-eigenvector with eigenvalue 1.

Fact: For any row-stochastic matrix, $\exists$ SS $\pi \geq 0$. is a left-eigenvector with eigenvalue 1. [Perron-Frobenius Thm].

Def: A column vector v is a <u>right-eigenvector</u> of $[P]$ with e-value λ if $v \neq 0$,

$$[P]v = \lambda v, \text{ i.e., } \sum_j P_{ij} v_j = v_i \quad \forall i.$$

Fact: For any stochastic matrix $\exists$ unique (within a scale factor) right-eigenvector with eigenvalue 1, given by $[1, 1, \dots, 1]^T$.

$$[P]\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad [P]e = e.$$

Focus on the case $M=2$ to get insight.

$$[P] = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} \qquad \boxed{[P^n] \xrightarrow{n\to\infty} (?)}$$

Left eigen-decomposition	Right eigen-decomposition.
$\pi[P] = \lambda\pi$.	$[P]v = \lambda v$
$\pi_1 P_{11} + \pi_2 P_{21} = \lambda \pi_1$	$P_{11}v_1 + P_{12}v_2 = \lambda v_1$
$\pi_1 P_{12} + \pi_2 P_{22} = \lambda \pi_2$	$P_{21}v_1 + P_{22}v_2 = \lambda v_2$.

$$\det(P - \lambda I) = \det\begin{pmatrix} P_{11}-\lambda & P_{12} \\ P_{21} & P_{22}-\lambda \end{pmatrix}$$

$$= (P_{11}-\lambda)(P_{22}-\lambda) - P_{12}P_{21} = 0$$

$$\underline{\lambda_1 = 1}, \quad \underline{\lambda_2 = 1 - P_{12} - P_{21}}.$$

$$\because \ (P_{11}-1)(P_{22}-1) - P_{12}P_{21}$$
$$= (-P_{12})(-P_{21}) - P_{12}P_{21} = 0.$$

Assuming $(P_{12}, P_{21}) \neq (0,0)$

$$\pi^{(1)} = \left[\frac{P_{21}}{P_{12}+P_{21}}, \frac{P_{12}}{P_{12}+P_{21}}\right] \qquad v^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\pi^{(2)} = [1, -1] \qquad v^{(2)} = \begin{bmatrix} \frac{P_{12}}{P_{12}+P_{21}} \\ -\frac{P_{21}}{P_{12}+P_{21}} \end{bmatrix}$$

Rmk: Normalization chosen s.t. $\pi^{(1)}$ is a SS vector

$$\sum_{i=1}^{2} \pi_i^{(1)} = 1 \quad \& \quad \pi^{(1)} v^{(1)} = 1.$$

$$\pi^{(2)} v^{(2)} = 1$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad [U] = \begin{bmatrix} v_1^{(1)} & v_1^{(2)} \\ v_2^{(1)} & v_2^{(2)} \end{bmatrix}$$

$$[P] v^{(i)} = \lambda_i v^{(i)} \qquad = \begin{bmatrix} | & | \\ v^{(1)} & v^{(2)} \\ | & | \end{bmatrix}$$

$$[P][U] = [U][\Lambda]$$

$$\boxed{[P] = [U][\Lambda][U]^{-1}} \quad \text{—— } (*)$$

The inverse $[U]^{-1}$ is the matrix with rows being the left eigenvectors of $[P]$

$$[U]^{-1} = \begin{bmatrix} - \pi^{(1)} - \\ - \pi^{(2)} - \end{bmatrix} = \begin{bmatrix} \pi_1^{(1)} & \pi_2^{(1)} \\ \pi_1^{(2)} & \pi_2^{(2)} \end{bmatrix}$$

Using $(*)$ multiple times

$$= \begin{bmatrix} \frac{P_{21}}{P_{21}+P_{12}} & \frac{P_{12}}{P_{21}+P_{12}} \\ 1 & -1 \end{bmatrix}$$

$$[P^2] = \left([U][\Lambda]\cancel{[U]^{-1}}\right)\left(\cancel{[U]}[\Lambda][U]^{-1}\right)$$

$$= [U][\Lambda^2][U]^{-1}$$

$$\Rightarrow \quad [P^n] = [U][\Lambda^n][U]^{-1}$$

$$= \begin{bmatrix} \pi_1 + \pi_2 \lambda_2^n & \pi_2 - \pi_2 \lambda_2^n \\ \pi_1 - \pi_1 \lambda_2^n & \pi_2 + \pi_1 \lambda_2^n \end{bmatrix}$$

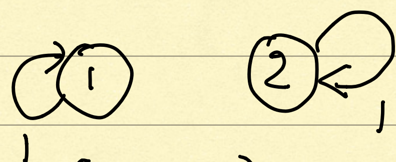
$$[\Lambda] = \begin{pmatrix} 1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad \lambda_2 = 1 - P_{12} - P_{21}$$

$$[\Lambda^n] = \begin{pmatrix} 1 & 0 \\ 0 & \lambda_2^n \end{pmatrix} \qquad [U]^{-1} = \begin{pmatrix} \pi_1 & \pi_2 \\ 1 & -1 \end{pmatrix}$$

$$\lambda_2 = 1 - P_{12} - P_{21} \Rightarrow \quad |\lambda_2| \leq 1$$

Trivial cases: $|\lambda_2| = 1$

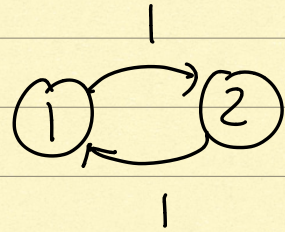
(a) $\lambda_2 = 1$, $P_{12} = P_{21} = 0$



$$[P^n] = \begin{pmatrix} \pi_1 + \pi_2 & \pi_2 - \pi_2 \\ \pi_1 - \pi_1 & \pi_2 + \pi_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ \forall n$$

$[P^n] = I \quad \forall n.$ 2 recurrent classes.

(b) $\lambda_2 = -1$, $P_{12} = P_{21} = 1$



$$[P^n] = \begin{bmatrix} \pi_1 + \pi_2(-1)^n & \pi_2 - \pi_2(-1)^n \\ \pi_1 - \pi_1(-1)^n & \pi_2 + \pi_1(-1)^n \end{bmatrix}$$

$$= \begin{cases} I & n \text{ even.} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = P & n \text{ odd}, \ (-1)^n = -1, \ \pi_1 = \pi_2 = \frac{1}{2}. \end{cases}$$

$[P^n] \not\to$ has no limit.

In all other cases, $|\lambda_2|<1$.

$$[P^n] = \begin{bmatrix} \pi_1 + \cancel{\pi_2 \lambda_2^n} \to 0 & \pi_2 - \cancel{\pi_2 \lambda_2^n} \to 0 \\ \pi_1 - \cancel{\pi_1 \lambda_2^n} \to 0 & \pi_2 + \cancel{\pi_1 \lambda_2^n} \to 0 \end{bmatrix}$$

$$\lim_{n\to\infty} [P^n] = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix}$$

If $|\lambda_2|<1$, $[P^n]$ approach the SS matrix $e\pi$.

$$[P^n] - e\pi = \begin{bmatrix} \pi_2 \lambda_2^n & -\pi_2 \lambda_2^n \\ -\pi_1 \lambda_2^n & \pi_1 \lambda_2^n \end{bmatrix}$$

Geometrically decaying

Geomtrically fast

$$c \in \{\pi_1, \pi_2\}$$

$$\left| [P^n]_{ij} - \pi_j \right| \le c\,|\lambda_2|^n \le |\lambda_2|^n$$

$$\ge \min\{\pi_1, \pi_2\}\,|\lambda_2|^n$$

Rate of convergence is governed by the second-largest eigenvalue of $[P]$, $|\lambda_2|$

$$\lim_{n\to\infty} \frac{1}{n} \log\left| [P^n]_{ij} - \pi_j \right| = \log|\lambda_2|.$$

$$\Uparrow \quad \frac{1}{n} \log \left|[P^n]_{ij} - \pi_j\right| \le \frac{1}{n} \log(|\lambda_2|^n) = \cancel{n} \log|\lambda_2|$$

This result is tight

Case of M distinct eigenvalues. $\lambda_1, \dots, \lambda_M$

$[P] - \lambda_i[I]$ is singular $\forall i \in \{1, \dots, M\}$
so there is a left- & right-eigenvector corr. to λ_i $\forall i$.

Say $\pi^{(i)}$ & $v^{(i)}$, $i \in \{1, \dots, M\}$.

$\pi^{(i)} v^{(i)} = 1$, $i \in \{1, \dots, M\}$.

$[P^n] = [U][\Lambda^n][U]^{-1}$

No need to learn repeated eigenvalues

$$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_M \end{pmatrix}$$

right-eigenvectors

$$[U] = \begin{bmatrix} | & | & & | \\ v^{(1)} & v^{(2)} & \cdots & v^{(M)} \\ | & | & & | \end{bmatrix}$$

$$[U]^{-1} = \begin{bmatrix} \text{---} & \pi^{(1)} & \text{---} \\ & \vdots & \\ \text{---} & \pi^{(M)} & \text{---} \end{bmatrix}$$

left-eigenvectors.

$\pi^{(1)} v^{(1)} = 1$

$\sum_{i=1}^{M} \pi_i^{(1)} = 1$

Rmk: We assumed $\pi^{(i)} v^{(i)} = 1$ $\forall i$, this is the reason why $\pi^{(i)}$ appear as rows of $[U]^{-1}$.

Rmk: $\lambda_1 = 1$, $e = (1, 1, \dots, 1)^T$
Must also have a <u>left e-value</u> of 1.

$|\lambda_i| \leq 1 \quad \forall i \in \{1, \dots, M\}$

$$[P^n] = 1 \cdot e\pi^{(1)} + \lambda_2^n v^{(2)}\pi^{(2)} + \dots + \lambda_M^n v^{(M)}\pi^{(M)}$$

$$[P^n] = \begin{pmatrix} | & & | \\ v^{(1)} & \cdots & v^{(M)} \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_M^n \end{pmatrix} \begin{pmatrix} — \pi^{(1)} — \\ \vdots \\ — \pi^{(M)} — \end{pmatrix}$$

$v^{(1)} = e$

Order $\underset{\lambda_1}{\underset{\parallel}{1}} \geq |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_M|$.

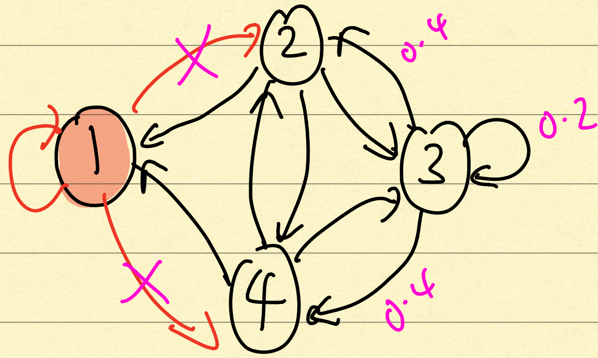
If $|\lambda_2| < 1$, $|\lambda_i| < 1 \quad \forall i > 1$,

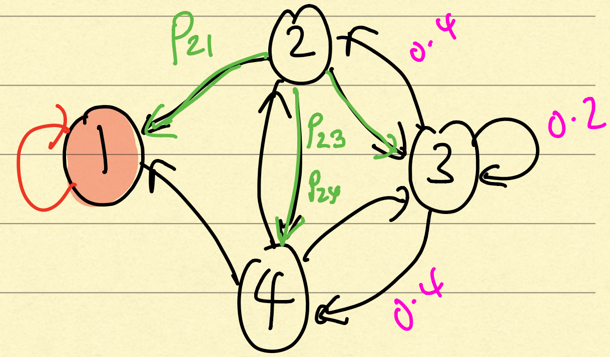
$$[P^n] - e\pi^{(1)} = \lambda_2^n v^{(2)}\pi^{(2)} + \dots + \lambda_M^n v^{(M)}\pi^{(M)}$$

$$\left| [P^n]_{ij} - \pi_j^{(1)} \right| = \Theta\left(\underset{\text{const.}}{\cancel{2}}|\lambda_2|^n\right)$$

If $|\lambda_2| < 1$, $[P^n] \to e\pi^{(1)}$ geometrically fast with rate $|\lambda_2|^n$.

Expected First-Passage Times.





Let v_i be the expected # of steps to first reach state 1 from state $i \neq 1$.

$$\boxed{\begin{aligned} v_2 &= P_{21}\cdot 1 + P_{23}(1+v_3) + P_{24}(1+v_4) \\ v_2 &= 1 + P_{23}v_3 + P_{24}v_4. \end{aligned}}$$

$$v_3 = 1 + P_{32}v_2 + P_{33}v_3 + P_{34}v_4$$

$$v_4 = 1 + P_{42}v_2 + P_{43}v_3.$$

$$\rightarrow \quad v_i = 1 + \sum_{j \neq i} P_{ij}v_j \qquad i \neq 1.$$

$$\underline{v} = \underline{r} + [P]\,\underline{v}$$

$$\rightarrow \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_M \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & \cdots & 0 \\ P_{21} & P_{22} & & P_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ P_{M1} & P_{M2} & \cdots & P_{MM} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_M \end{bmatrix}$$

$M-1$ linearly independent equations if $M-1$ states $2, \dots, M$ are transient & $\{1\}$ is the only recurrent class.