MCAA lecture 6

Theorem (to be proven today)

Let $(X_n, n \geq 0)$ be an ergodic Markov chain with finite state space S $(|S| = N)$, transition matrix P and stationary and limiting distribution π satisfying the detailed balance equation. Then

$$\| P_i^n - \pi \|_{TV} \leq \frac{\lambda_*^n}{2\sqrt{\pi_i}} \quad \forall i \in S, \ \forall n \geq 1$$

where $1 = \lambda_0 > \lambda_1 \geq \ldots \geq \lambda_{N-1} > -1$ are the eigenvalues of P

and $\lambda_* = \max\limits_{1 \leq k \leq N-1} |\lambda_k| = \max\{\lambda_1, -\lambda_{N-1}\} < 1$.

Reminder:

- Define $q_{ij} = \sqrt{\pi_i}\, P_{ij} \frac{1}{\sqrt{\pi_j}} \xrightarrow{\text{detailed balance}} Q$ is a symmetric matrix

$\xrightarrow[\text{theorem}]{\text{spectral}}$ $Q u^{(k)} = \lambda_k u^{(k)} \quad k = 0 \ldots N-1$

$$\begin{cases} u^{(0)} \ldots u^{(N-1)} \text{ are orthonormal vectors in } \mathbb{R}^N \left(u^{(k)T} u^{(\ell)} = \delta_{k\ell} \right) \\ \lambda_0 \geq \ldots \geq \lambda_{N-1} \in \mathbb{R} \end{cases}$$

- Define $\phi_j^{(k)} = \frac{u_j^{(k)}}{\sqrt{\pi_j}}$: $P \phi^{(k)} = \lambda_k \phi^{(k)} \quad k = 0 \ldots N-1$

- **Facts**:
  - $\lambda_0 = 1$ and $\phi_j^{(0)} = 1 \quad \forall j \in S \qquad (1)$
  - $|\lambda_k| \leq 1 \quad \forall 1 \leq k \leq N-1 \qquad (2)$
  - $\lambda_1 < 1 \quad \& \quad \lambda_{N-1} > -1 \qquad (3)$

Proof of (1):

To be proven: $P\phi^{(0)} = \overset{\lambda_0}{\overset{\shortparallel}{1}} \cdot \phi^{(0)}$ where $\phi^{(0)}_j = 1 \ \forall j \in S$

$$\left(P\phi^{(0)}\right)_i = \sum_{j\in S} p_{ij} \cdot \underbrace{\phi^{(0)}_j}_{=1} = \sum_{j\in S} p_{ij} \underset{\forall i \in S}{=} 1 = \phi^{(0)}_i \quad \checkmark$$

Proof of (2):

To be proven: Let ϕ be an eigenvector of P

i.e. $\exists \lambda \in \mathbb{R}$ s.t. $P\phi = \lambda\phi$; then $|\lambda| \leq 1$.

ϕ is an eigenvector $\Rightarrow \phi \not\equiv 0$

Let $i \in S$ such that $|\phi_i| \geq |\phi_j| \ \forall j \in S$; $|\phi_i| > 0$

Then $(P\phi)_i = \lambda\phi_i$ i.e. $\lambda\phi_i = \sum_{j\in S} p_{ij}\phi_j$

Take absolute values: $|\lambda \cdot \phi_i| = |\sum_{j \in S} p_{ij} \cdot \phi_j|$

$$\leq \sum_{j \in S} \underbrace{p_{ij}}_{\geq 0} \cdot \underbrace{|\phi_j|}_{\leq |\phi_i|} \leq |\phi_i| \cdot \underbrace{\sum_{j \in S} p_{ij}}_{=1} = |\phi_i|$$

So $|\lambda \cdot \phi_i| \leq |\phi_i|$ i.e. $|\lambda| \cdot \underbrace{|\phi_i|}_{>0} \leq \underbrace{|\phi_i|}_{>0}$ i.e. $|\lambda| \leq 1$. ✓

<u>Proof of (3):</u>

<u>$\lambda_1 < +1$</u>: To be proven: Assume $\phi \in \mathbb{R}^N$ is such that

$P\phi = 1 \cdot \phi$; then ϕ is a multiple of $\phi^{(0)}$

<u>Note</u>: if $P\phi = \phi$, then $P^2\phi = P(P\phi) = P\phi = \phi$

and likewise, $P^n \phi = P^{n-1}\phi \ldots = \phi$ $\forall n \geq 1$.

The chain is irreducible, aperiodic & finite

$\Longrightarrow \exists n_0 \geq 1$ s.t. $p_{ij}(n) > 0 \quad \forall i, j \in S \quad \forall n \geq n_0$

Define $i \in S$ s.t. $|\phi_i| \geq |\phi_j| \quad \forall j \in S$ (so $|\phi_i| > 0$)

$$P^n \phi = \phi \quad : \quad \sum_{j \in S} \underbrace{p_{ij}(n)}_{>0} \phi_j = \phi_i \qquad (n \geq n_0)$$

$$|\phi_i| = \left| \sum_{j \in S} p_{ij}(n) \cdot \phi_j \right| \leq \sum_{j \in S} p_{ij}(n) \cdot \underbrace{|\phi_j|}_{\leq |\phi_i| \ \forall j \in S} \leq |\phi_i| \cdot \underbrace{\sum_{j \in S} p_{ij}(n)}_{=1}$$

$$= |\phi_i|$$

i.e. $\begin{cases} \sum_{j \in S} \underbrace{p_{ij}(n)}_{>0} \cdot |\phi_j| = |\phi_i| \\ \sum_{j \in S} p_{ij}(n) = 1 \end{cases} \Longrightarrow$ By contradiction, all $|\phi_j| = |\phi_i|$ ✓

$$\sum_{j\in S} \underbrace{p_{ij}(n)}_{>0} \phi_j = \phi_i \quad \& \quad \sum_{j\in S} p_{ij}(n) = 1 \quad \& \quad |\phi_j| \equiv |\phi_i|$$

By contradiction, all ϕ_j should have the same sign

So ϕ is a multiple of $\phi^{(0)} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$. ✓

<u>$\lambda_{N-1} > -1$:</u>

<u>To be proven</u>: $\nexists\, \phi \in \mathbb{R}^N, \phi \neq 0$ such that $P\phi = -\phi$

<u>Note</u>: Assume such a ϕ exists. Then

$$P^2\phi = P(P\phi) = P(-\phi) = -P\phi = \phi$$

$$P^3\phi = -\phi, \quad P^4\phi = \phi \rightarrow \begin{cases} n \text{ odd}: & P^n\phi = -\phi \\ n \geq n_0 & \end{cases}$$

Define $i \in S$ s.t. $|\phi_i| \geq |\phi_j| \ \forall j \in S$ $\quad (|\phi_i| > 0)$

$$P^n \phi = -\phi \quad : \quad -\phi_i = \sum_{j \in S} \underbrace{p_{ij}(n)}_{>0} \phi_j$$

$$\text{so} \quad |\phi_i| \leq \sum_{j \in S} p_{ij}(n) \cdot \underbrace{|\phi_j|}_{\leq |\phi_i|} \leq |\phi_i|$$

$$\text{so} \quad |\phi_i| = \sum_{j \in S} \underbrace{p_{ij}(n)}_{>0} \cdot |\phi_j| \ , \quad \sum_{j \in S} p_{ij}(n) = 1$$

$\Longrightarrow$ by contradiction, all $|\phi_j| = |\phi_i|$ ✓

but here $-\phi_i = \sum_{j \in S} \underbrace{p_{ij}(n)}_{>0} \cdot \phi_j$ & $\sum_{j \in S} p_{ij}(n) = 1$

$\Longrightarrow$ by contradiction, it is impossible to satisfy this equality ✓

Proof of the theorem

$$\| P_i^n - \pi \|_{TV} = \frac{1}{2} \sum_{j \in S} | p_{ij}(n) - \pi_j |$$

"dist at time n"   "dist. at +∞"

$$= \frac{1}{2} \sum_{j \in S} \underbrace{\frac{|p_{ij}(n) - \pi_j|}{\sqrt{\pi_j}}}_{a_j} \cdot \underbrace{\sqrt{\pi_j}}_{b_j}$$

$$\leq \frac{1}{2} \sqrt{\sum_{j \in S} \frac{|p_{ij}(n) - \pi_j|^2}{\pi_j} \cdot \underbrace{\sum_{j \in S} \pi_j}_{=1}}$$

$$\leq \frac{1}{2} \sqrt{\sum_{j \in S} \left(\frac{p_{ij}(n)}{\sqrt{\pi_j}} - \sqrt{\pi_j} \right)^2}$$

C-S inequality: $\left| \sum_{j \in S} a_j b_j \right| \leq \sqrt{\sum_{j \in S} a_j^2 \cdot \sum_{j \in S} b_j^2}$

Lemma

$$\frac{p_{ij}(n)}{\sqrt{\pi_j}} - \sqrt{\pi_j} = \sqrt{\pi_j} \sum_{k=1}^{N-1} \lambda_k^n \phi_i^{(k)} \phi_j^{(k)} \quad \forall i,j \in S, \ \forall n \geq 1$$

Proof

$u^{(0)} \dots u^{(N-1)}$ = eigenvectors of Q = orthonormal basis of $\mathbb{R}^N$

for $v \in \mathbb{R}^N$, $\quad v = \sum_{k=0}^{N-1} (v^T u^{(k)}) u^{(k)}$ i.e. $v_j = \sum_{k=0}^{N-1} (v^T u^{(k)}) u_j^{(k)}$ $j \in S$

choose $v_j = \frac{p_{ij}(n)}{\sqrt{\pi_j}}$ (fix $i \in S$) $\qquad v^T u^{(k)} = \sum_{j \in S} \frac{p_{ij}(n)}{\sqrt{\pi_j}} u_j^{(k)} = \sum_{j \in S} p_{ij}(n) \phi_j^{(k)}$

$$= (P^n \phi^{(k)})_i = \lambda_k^n \cdot \phi_i^{(k)}$$

$$\frac{p_{ij}(n)}{\sqrt{\pi_j}} = \sum_{k=0}^{N-1} (\lambda_k^n \phi_i^{(k)}) \cdot u_j^{(k)} = \sqrt{\pi_j} \cdot \sum_{k=0}^{N-1} \lambda_k^n \phi_i^{(k)} \phi_j^{(k)}$$

$$= \sqrt{\pi_j} + \sqrt{\pi_j} \sum_{k=1}^{N-1} \lambda_k^n \phi_i^{(k)} \phi_j^{(k)} \qquad \#$$

$$\| P_i^n - \pi \|_{TV} \leq \frac{1}{2} \sqrt{\sum_{j \in S} \pi_j \underbrace{\left(\sum_{k=1}^{N-1} \lambda_k^n \phi_i^{(k)} \phi_j^{(k)} \right)^2}_{= \sum_{k,\ell=1}^{N-1} \lambda_k^n \lambda_\ell^n \phi_i^{(k)} \phi_i^{(\ell)} \phi_j^{(k)} \phi_j^{(\ell)}}}$$

$$= \frac{1}{2} \sqrt{\sum_{k,\ell=1}^{N-1} \lambda_k^n \lambda_\ell^n \phi_i^{(k)} \phi_i^{(\ell)} \cdot \underbrace{\sum_{j \in S} \pi_j \phi_j^{(k)} \phi_j^{(\ell)}}_{= \sum_{j \in S} u_j^{(k)} u_j^{(\ell)} = (u^{(k)})^T u^{(\ell)} = \delta_{k\ell}}}$$

$$= \frac{1}{2} \sqrt{\sum_{k=1}^{N-1} \underbrace{\lambda_k^{2n}}_{\leq \lambda_*^{2n}} \underbrace{\left(\phi_i^{(k)}\right)^2}_{\geq 0}}$$

$$\leq \frac{\lambda_*^n}{2} \cdot \sqrt{\sum_{k=1}^{N-1} \left(\phi_i^{(k)}\right)^2}$$

Last detail: $\sum_{k=1}^{N-1} \left(\phi_i^{(k)}\right)^2 \leq \frac{1}{\pi_i}$

Remember: $\forall v \in \mathbb{R}^N$: $v_j = \sum_{k=0}^{N-1} \left(v^T u^{(k)}\right) \cdot u_j^{(k)}$

Take $v_j = \delta_{ij} = \begin{cases} 1 & \text{if } j=i \\ 0 & \text{o.w.} \end{cases}$ $\quad v^T u^{(k)} = u_i^{(k)}$

so $\delta_{ij} = \sum_{k=0}^{N-1} u_i^{(k)} u_j^{(k)}$

$$\delta_{ii} = 1 = \sum_{k=0}^{N-1} \left(u_i^{(k)}\right)^2 = \sum_{k=0}^{N-1} \pi_i \left(\phi_i^{(k)}\right)^2$$

$$= \pi_i \left(1 + \sum_{k=1}^{N-1} \left(\phi_i^{(k)}\right)^2\right)$$

$$\Longrightarrow \sum_{k=1}^{N-1} \left(\phi_i^{(k)}\right)^2 = \frac{1}{\pi_i} - 1 \leq \frac{1}{\pi_i} \qquad \#$$

Lazy Markov chains

$$\| P_i^n - \pi \|_{TV} \leq \frac{\lambda_*^n}{2\sqrt{\pi_i}} \leq \frac{e^{-\gamma n}}{2\sqrt{\pi_i}} \qquad \forall i \in S, \ \forall n \geq 1$$

spectral gap: $\gamma = 1 - \lambda_*$

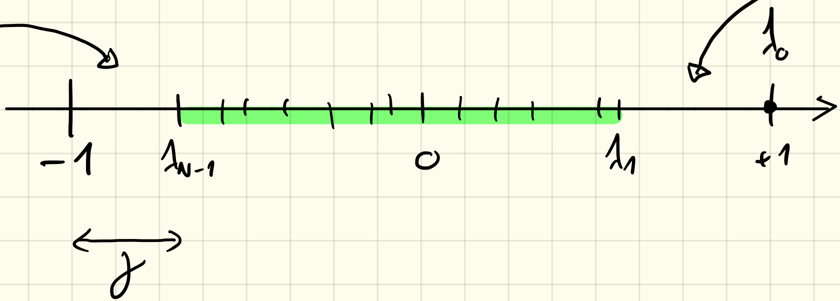
the larger the spectral gap

$\Longleftrightarrow$ the faster the convergence

$\lambda_{N-1} = -1$

iff

the graph of the chain is bipartite (period 2)



$\lambda_1 = 1$

iff

the graph of the chain has two disconnected components (reducible chain)

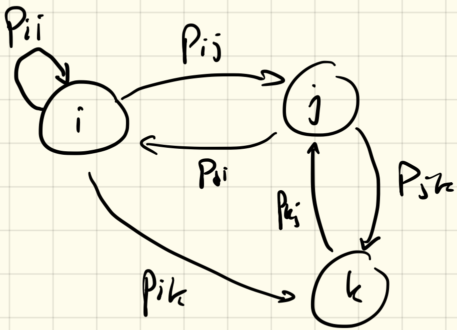
Let P be the transition matrix of a Markov chain

Define $\tilde{P} = \alpha \cdot Id + (1-\alpha) P$ where $0 < \alpha < 1$

<u>Claim</u>: $\tilde{P}$ is also a transition matrix

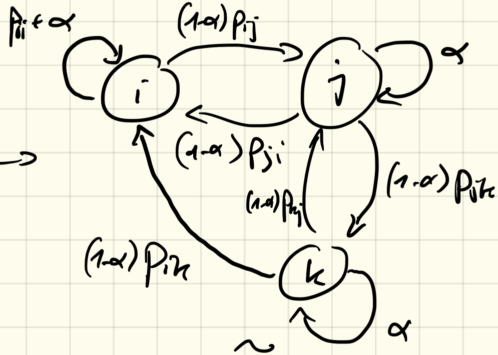
Indeed, $\begin{cases} \tilde{p}_{ij} = \alpha \, \delta_{ij} + (1-\alpha)\, p_{ij} \geq 0 & \forall i,j \in S \\ \sum_{j \in S} \tilde{p}_{ij} = \alpha \cdot 1 + (1-\alpha) \underbrace{\sum_{j \in S} p_{ij}}_{1} = 1 & \forall i \in S \end{cases}$

✓



P

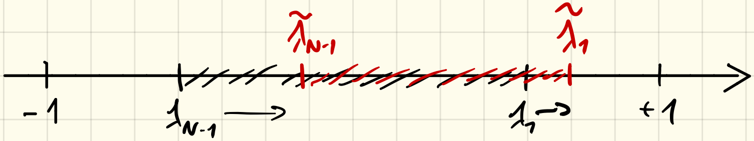
$\longrightarrow$



$\tilde{P}$

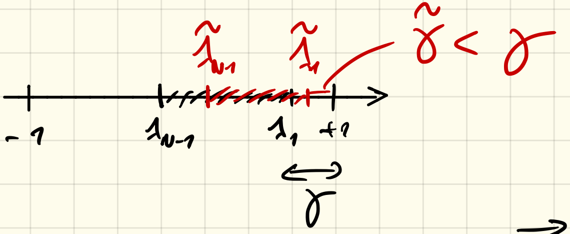
Eigenvalues of $\tilde{P}$: $\tilde\lambda_k = \alpha \cdot 1 + (1-\alpha)\lambda_k \qquad 0 \le k \le N-1$

$$\tilde\lambda_0 = \alpha + (1-\alpha)\underbrace{\lambda_0}_{=1} = 1 \qquad \left(\lambda_{N-1} = -1 \Rightarrow \tilde\lambda_{N-1} = \alpha - (1-\alpha)\right)$$
$$= 2\alpha - 1$$



Two situations:

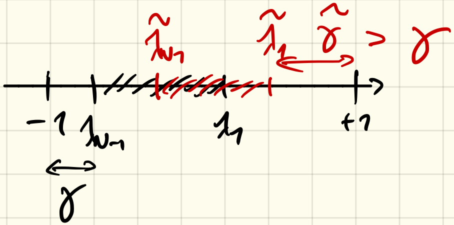
1



the spectral gap is reduced

→ slower convergence

2



the spectral gap is larger

→ faster convergence!