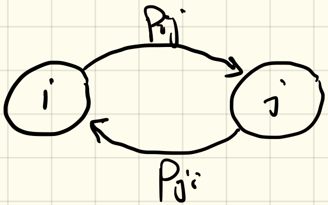
MCAA lecture 5

Reversible chains & detailed balance

Def: Let $(X_n, n \geq 0)$ be an ergodic Markov chain with state space S. It is said to be <u>reversible</u> if its stationary distribution π satisfies the <u>detailed balance equation</u>:

$$\pi_i \, p_{ij} = \pi_j \, p_{ji} \quad \forall i, j \in S \qquad (*)$$



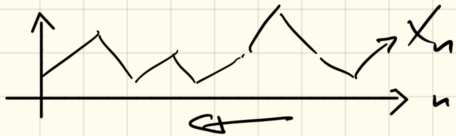
Remark:

- $\pi = \pi P$ does not ensure that eq. (*) is satisfied.
- if eq. (*) is satisfied, then $\pi = \pi P$:

$$(\pi P)_j = \sum_{i \in S} \pi_i p_{ij} \overset{(*)}{=} \sum_{i \in S} \pi_j p_{ji} = \pi_j \underbrace{\sum_{i \in S} p_{ji}}_{=1 \ \forall j \in S} = \pi_j$$

- why "reversible"?

Assume that $\pi^{(0)} = \pi$ and look at the chain backwards in time:

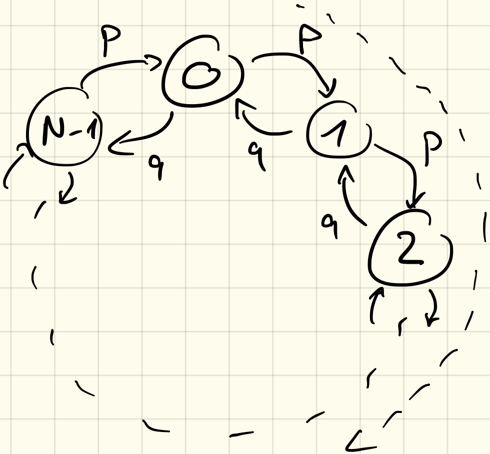


It turns out that if eq. (*) is satisfied, then the backward chain has the same transition probabilities as the forward chain.

Examples & counter-examples

- If $\exists\, i, j \in S$ such that $p_{ij} > 0$ but $p_{ji} = 0$, then the detailed balance equation cannot be satisfied.

- Cyclic random walk on $S = \{0, 1, \dots, N-1\}$, N odd



$0 < p, q < 1$, $p + q = 1$

limiting & stat. dist. $\pi = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)$

(doubly stochastic matrix P)

detailed balance? $\pi_i\, p_{ij} = \pi_j\, p_{ji}$?

$\frac{1}{N} \cdot p = \frac{1}{N} \cdot q$?

only if $p = q = \frac{1}{2}$!

- If S is finite and the matrix P is tridiagonal, i.e.

$$P = \begin{pmatrix} \diagdown\!\diagdown\!\diagdown & & 0 \\ & \diagdown\!\diagdown\!\diagdown & \\ 0 & & \diagdown\!\diagdown\!\diagdown \end{pmatrix}$$

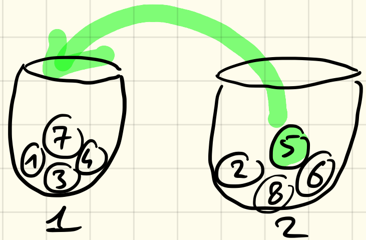
non-zero coefficients

Then the chain is reversible (provided it is also ergodic).

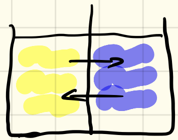
<u>NB:</u> For the cyclic RW:

$$P = \begin{pmatrix} 0 & p & 0 & \cdots & \cdots & q \\ q & 0 & p & 0 & \cdots & 0 \\ 0 & q & 0 & p & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & q & 0 & p \\ p & \cdots & \cdots & \cdots & q & 0 \end{pmatrix}$$

- Consider two urns with N numbered balls.



$N=8$



At each step, pick a number between 1 & N uniformly at random; take the ball with this number and put it in the other urn.

State of the chain: $X_n = \#$ balls in urn 1 at time n

Transition probabilities: $p_{i,i+1} = \frac{N-i}{N}$, $p_{i,i-1} = \frac{i}{N}$ tridiagonal

Detailed balance: $\pi_i \, p_{i,i+1} = \pi_{i+1} \, p_{i+1,i}$ i.e. $\pi_i \frac{N-i}{N} = \pi_{i+1} \cdot \frac{i+1}{N}$

so $\pi_{i+1} = \frac{N-i}{i+1} \cdot \pi_i \implies \pi_i = \frac{(N-i)(N-i+1)\cdots N}{i\cdot(i-1)\cdots 1} \cdot \pi_0 = \binom{N}{i} \cdot \pi_0$

& $\sum_{i=0}^{N} \pi_i = 1 \implies \pi_0 = 1 / \sum_{i=0}^{N} \binom{N}{i} = 1/2^N$

Rate of convergence

Let $(X_n, n \geq 0)$ be an ergodic Markov chain on S with transition matrix P, initial distribution $\pi^{(0)}$ and stationary and limiting distribution π.

Moreover, we assume:

- S is finite, i.e., $|S| = N$.
- detailed balance holds, i.e., $\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j \in S$.

Our aim: to find an upper bound on

$$\| \underbrace{P_i^n}_{\substack{i^{\text{th}} \text{ row of } P^n \\ = \text{distribution at time } n \text{ given } \pi^{(0)} = \delta_i}} - \pi \|_{TV} = \frac{1}{2} \sum_{j \in S} |p_{ij}^{(n)} - \pi_j| \quad \Big(\xrightarrow[n \to \infty]{} 0 \text{ by the ergodic thm} \Big)$$

Eigenvalues and eigenvectors of P

Define a new matrix Q as follows: $q_{ij} = \sqrt{\pi_i} \cdot p_{ij} \frac{1}{\sqrt{\pi_j}} \quad i,j \in S$

Then:
- $q_{ii} = p_{ii} \quad \forall i \in S$
- $q_{ij} \geq 0 \quad \forall i,j \in S$, but $\sum_{j \in S} q_{ij} \neq 1$ in general
- $q_{ij} = q_{ji} \quad \forall i,j \in S$, i.e. Q is symmetric:

  Indeed: $q_{ji} = \sqrt{\pi_j}\, p_{ji} \frac{1}{\sqrt{\pi_i}} = \frac{1}{\sqrt{\pi_j \pi_i}} \cdot \pi_j p_{ji}$

  (detailed balance) $= \frac{1}{\sqrt{\pi_j \pi_i}} \cdot \pi_i p_{ij} = \sqrt{\pi_i}\, p_{ij} \frac{1}{\sqrt{\pi_j}} = q_{ij}$ #

Spectral theorem

As Q is symmetric, there exist real numbers $\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{N-1}$ (= the eigenvalues of Q) and vectors $u^{(0)} \dots u^{(N-1)}$ (= the eigenvectors of Q) such that $Qu^{(k)} = \lambda_k u^{(k)}$ $\forall 0 \leq k \leq N-1$

Moreover, $u^{(0)}, \dots, u^{(N-1)}$ forms an orthonormal basis of $\mathbb{R}^N$.

Proposition

Define $\phi^{(k)} = \left(\frac{u_j^{(k)}}{\sqrt{\pi_j}}, j \in S \right)$. Then $P\phi^{(k)} = \lambda_k \phi^{(k)}$ $\forall 0 \leq k \leq N-1$

Proof: $\left(P\phi^{(k)}\right)_i = \sum_{j\in S} p_{ij} \phi_j^{(k)} = \sum_{j \in S} \left(\frac{1}{\sqrt{\pi_i}} q_{ij} \sqrt{\pi_j} \right) \frac{u_j^{(k)}}{\sqrt{\pi_j}}$

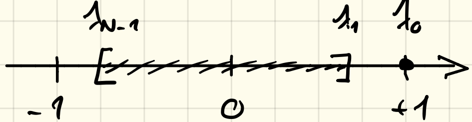
$$= \frac{1}{\sqrt{\pi_i}} \sum_{j \in S} q_{ij} u_j^{(k)} = \frac{1}{\sqrt{\pi_i}} \left(Qu^{(k)}\right)_i = \frac{1}{\sqrt{\pi_i}} \lambda_k u_i^{(k)} = \lambda_k \phi_i^{(k)}$$

$\forall 0 \leq k \leq N-1, \ \forall i \in S$ #

Facts about the eigenvalues of P (to be proven next week):

1. $\lambda_0 = 1$ and $\phi^{(0)} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

2. $|\lambda_k| \leq 1 \quad \forall\, 1 \leq k \leq N-1$



3. $\lambda_1 < 1$ & $\lambda_{N-1} > -1$

Define $\lambda_* = \max\limits_{1 \leq k \leq N-1} |\lambda_k| \overset{\text{exercise!}}{=} \max\{\lambda_1, -\lambda_{N-1}\} < 1$

Theorem: Under all the above assumptions (ergodic chain, finite S, detailed balance), it holds that

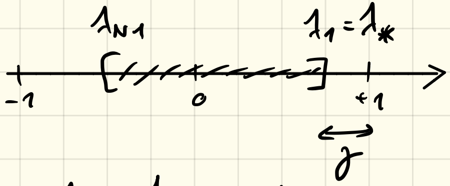
$$\| P_i^n - \pi \|_{TV} \leq \frac{\lambda_*^{\,n}}{2\sqrt{\pi_i}} \qquad \forall i \in S,\ \forall n \geq 1$$

Two more definitions

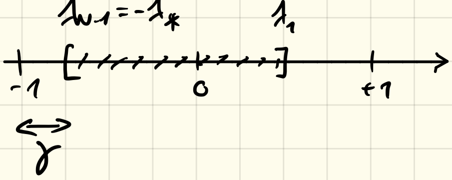
- The <u>spectral gap</u> of the chain is defined as follows:

$$\gamma = 1 - \lambda_* \in \,]0,1]$$

$$= \min\{1-\lambda_1,\ \lambda_{N-1}+1\}$$



$$\ldots \le \frac{\lambda_*^{\,n}}{2\sqrt{\pi_i}} = \frac{(1-\gamma)^n}{2\sqrt{\pi_i}} \le \frac{e^{-\gamma n}}{2\sqrt{\pi_i}}$$



The larger the γ, the faster the convergence

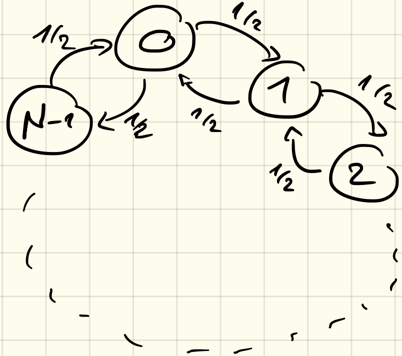
- The <u>mixing time</u> of the chain is defined as follows:

For a given $\varepsilon > 0$, $\quad T_\varepsilon = \inf\{n \ge 1 : \max_{i \in S} \| P_i^n - \pi \|_{TV} \le \varepsilon\}$

How does T_ε behave in terms of $N = |S|$?

Example: cyclic random walk on $S = \{0, \ldots, N-1\}$

with $p = q = \frac{1}{2}$

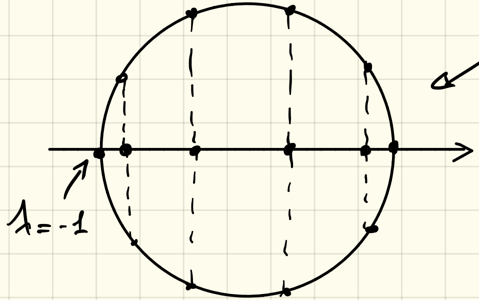


$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \cdots & \cdots & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \cdots & 0 \\ & & \ddots & & \ddots & & \frac{1}{2} \\ \frac{1}{2} & 0 & \cdots & \cdots & \cdots & \frac{1}{2} & 0 \end{pmatrix}$$

$$\pi = \left(\frac{1}{N}, \ldots \ldots, \frac{1}{N} \right)$$

Eigenvalues of P: $\lambda_k = \cos\left(\frac{2\pi k}{N}\right) \quad k = 0 \ldots N-1$ (⚠ no more ordering)

(P = circulant matrix)

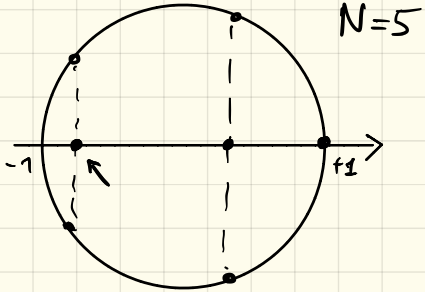


N is even here:

not ergodic

$\lambda_* = 1$

(⇒ no convergence)



$\underline{N \text{ odd}}$: $\lambda_* = \left| \cos\left(\frac{\pi\left(\frac{N-1}{2}\right)}{N} \right) \right|$

$= \left| \cos\left(\pi\left(1-\frac{1}{N}\right)\right) \right| = \cos\left(\frac{\pi}{N}\right)$

$\cong 1 - \frac{\pi^2}{2N^2}$ (N large)

So $\gamma = 1 - \lambda_* \cong \frac{\pi^2}{2N^2}$:

$$\| P_i^n - \pi \|_{TV} \le \frac{\lambda_*^n}{2\sqrt{\pi_i}} = \frac{\sqrt{N}}{2} \cdot \left(1 - \frac{\pi^2}{2N^2}\right)^n \le \frac{\sqrt{N}}{2} \exp\left(-\frac{\pi^2 n}{2N^2}\right)$$

$\le \varepsilon$ when $n >> N^2$

(for example, $n \sim \Theta(N^2 \log N)$)

So $T_\varepsilon = \inf\{ n \ge 1 : \max_{i \in S} \| P_i^n - \pi \|_{TV} \le \varepsilon \} \sim \Theta(N^2 \log N)$