CS70 @ UC Berkeley, Spring 2026

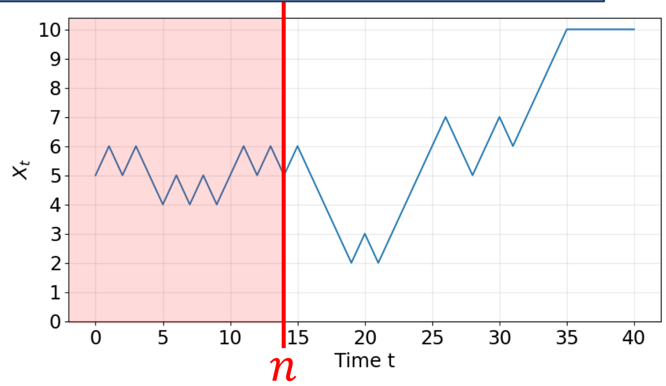
Lecture 27

Markov Chains II

April 30, 2026

Stochastic Process (Lecture 26)

Definition (Random process or stochastic process): A random process is a family $\{X_t, t \in T\}$ of random variables indexed by some set T.
E.g., $T = \mathbb{N}$ (discrete-time) or $T = [0, \infty)$ (continuous-time)



Stochastic processes $\{X_t, t \in T\}$ are distinguished by

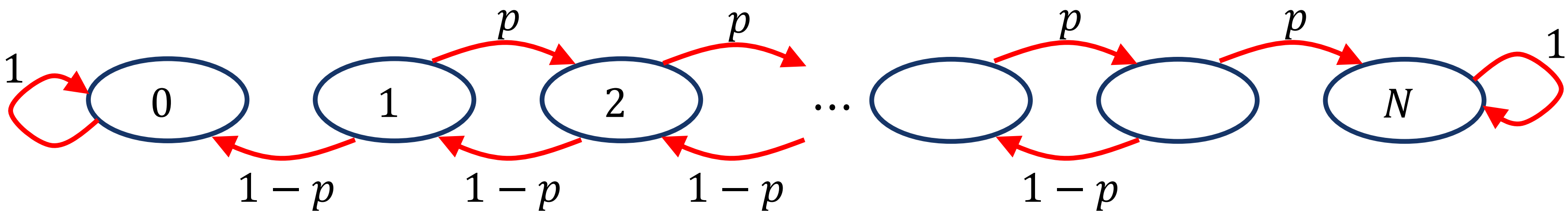
- State space S (the possible values of X_t)
- Time index set T.
- Dependence relation among X_t over time.

Markov property: $\forall n \in T$ and $a_0, \ldots, a_{n-1}, i, j \in S$,

$$\mathbb{P}[X_{n+1} = j \mid X_0 = a_0, X_1 = a_1, \ldots, X_{n-1} = a_{n-1}, X_n = i] = \mathbb{P}[X_{n+1} = j \mid X_n = i]$$

- Time homogeneous transition probability: $P_{ij} := \mathbb{P}[X_{n+1} = j \mid X_n = i]$.
- n-step transition probability: $\mathbb{P}[X_n = j \mid X_0 = i] = [\boldsymbol{P}^n]_{ij}$
- $\mathbb{P}[X_n = j] = [\boldsymbol{\pi}_0 \boldsymbol{P}^n]_j$
- The initial distribution of X_0: $\mathbb{P}[X_0 = i] = [\boldsymbol{\pi}_0]_i, \quad i \in S$

1D Random Walk with Absorbing Boundaries



$\alpha(i) = \mathbb{P}(\text{ Hit 0 before } N \mid X_0 = i\ )$

$$\alpha(i) = \begin{cases} \dfrac{N-i}{N}, & \text{if } p = \dfrac{1}{2} \\ \dfrac{r^i - r^N}{1 - r^N}, & \text{if } p \neq \dfrac{1}{2} \end{cases}$$

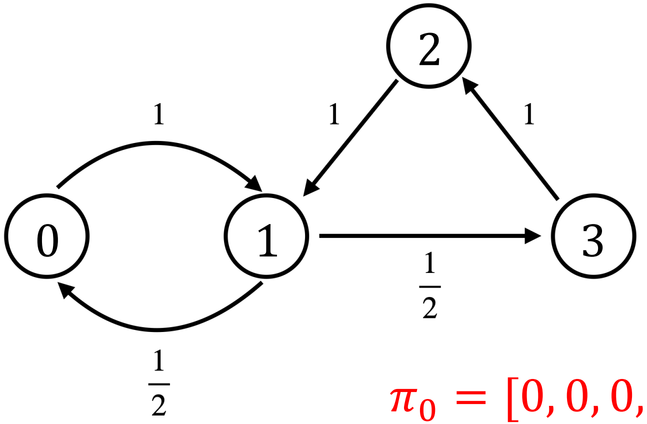
where $r = \frac{1-p}{p}$

$\beta(i)$ = Expected time until hitting either 0 or N, given that $X_0 = i$.

$$\beta(i) = \begin{cases} i(N-i), & \text{if } p = \dfrac{1}{2} \\ \dfrac{1}{1-2p}\left[i - N\dfrac{1-r^i}{1-r^N}\right], & \text{if } p \neq \dfrac{1}{2} \end{cases}$$

where $r = \frac{1-p}{p}$

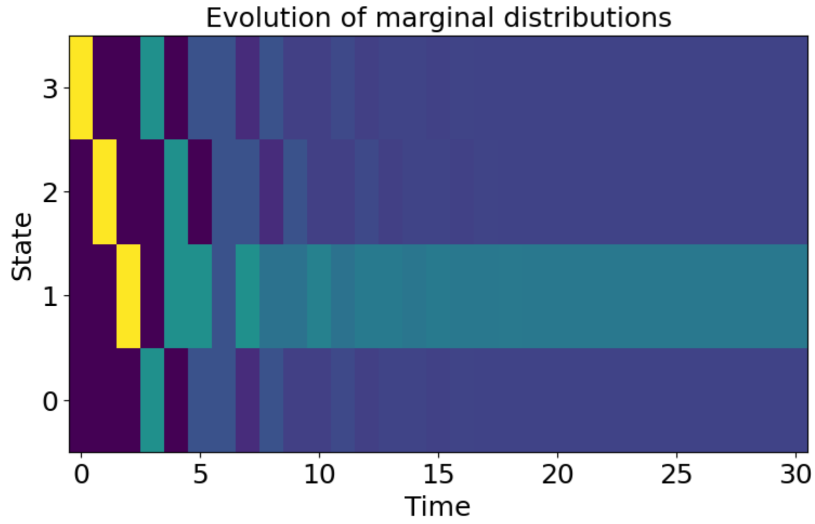
Time Evolution of the Distribution of X_n



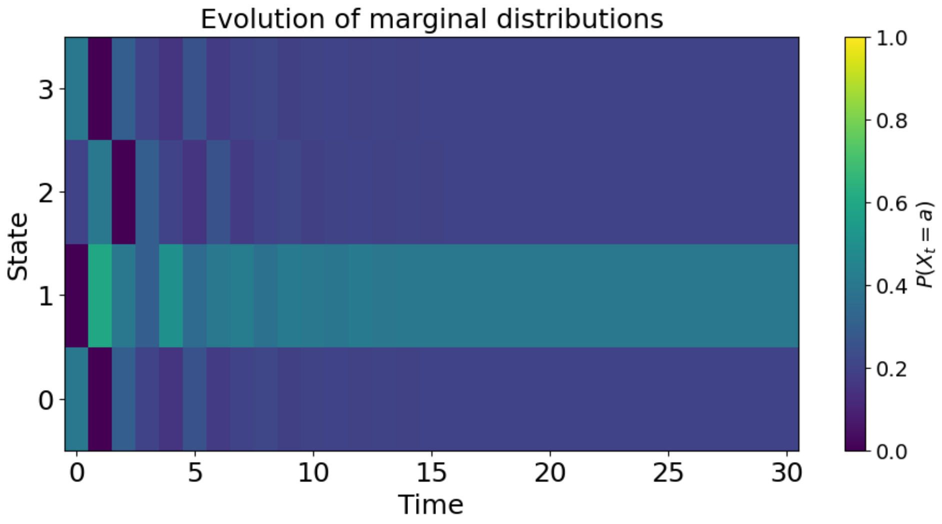
$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbb{P}[X_n = j] = [\pi_0 P^n]_j$$

$\pi_0 = [0, 0, 0, 1]$



$\pi_0 = [0.4, 0, 0.2, 0.4]$

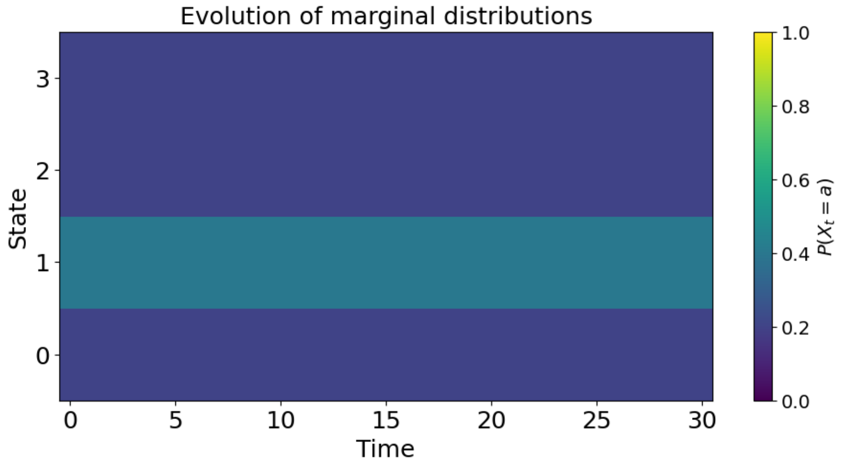


Stationary Distribution (aka Invariant Distribution)

Definition: The row vector $\pi = (\pi_i)_{i \in S}$ is called a **stationary distribution** of the Markov chain with transition probability matrix P if

1. $\pi_i \geq 0, \forall i \in S$ and $\sum_{i \in S} \pi_i = 1$
2. $\pi P = \pi$ (left eigenvector with eigenvalue 1)

Suppose π is a stationary distribution. Then, $X_0 \sim \pi \Rightarrow X_n \sim \pi, \forall n \in \mathbb{N}$



Questions:

1. Does every Markov chain have a stationary distribution?
2. When it exists, is it unique?
3. When does the distribution of X_n converge as $n \to \infty$?
4. When it converges, what does it converge to?

The Long Run Behavior of Markov Chain

- Consider a 2-state Markov chain $\{X_n, n \in \mathbb{N}\}$ with $S = \{1,\ 2\}$ and transition probability matrix
$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}, \text{ where } a, b \in (0,1]$$
- One can show: $P^n = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix} + \frac{(1-a-b)^n}{a+b}\begin{bmatrix} a & -a \\ -b & b \end{bmatrix}$
- For $(a, b) = (1,1)$, $P^n = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{(-1)^n}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{cases} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \text{if } n \text{ is odd,} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{if } n \text{ is even.} \end{cases}$
- For $(a, b) \neq (1,1)$, $|1-a-b| < 1$, so
  - $\lim_{n\to\infty}(1-a-b)^n = 0$ and $\lim_{n\to\infty} P^n = \frac{1}{a+b}\begin{bmatrix} b & a \\ b & a \end{bmatrix}$,
  - the distribution of X_n becomes independent of the initial condition X_0 in the limit as $n \to \infty$,
  - $\pi = \frac{1}{a+b}\begin{bmatrix} b & a \end{bmatrix}$ is the unique stationary distribution of the Markov chain.

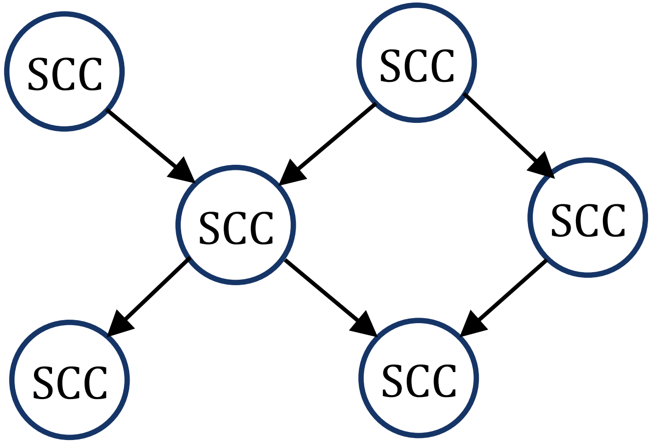
Classification of States

Definition:

- ($i \to j$) State j is accessible from state i if $[P^n]_{ij} > 0$ for some positive $n \in \mathbb{N}$ (i.e., there is a path from i to j in the graphical representation).
- ($i \leftrightarrow j$) States i, j intercommunicate if $i \to j$ and $j \to i$.

- $i \leftrightarrow j$ defines an equivalence relation.
- The state space S can be partitioned into equivalence classes of $i \leftrightarrow j$.
- A subset $C \subset S$ is called irreducible if $i \leftrightarrow j$ for all $i, j \in C$.
- C is a strongly connected component (SCC).

A DAG of SCCs



A Markov Chain is said to be **irreducible** if $i \leftrightarrow j$ for all $i, j \in S$.

Period

Theorem (Perron–Frobenius):

- Every Markov chain with finite S has a stationary distribution.
- If the Markov chain is irreducible, then it has a unique stationary distribution π.

Recall $[P^n]_{ij} = \mathbb{P}[X_n = j \mid X_0 = i]$

Under what condition, is $\lim_{n\to\infty} [P^n]_{ij} = \pi_j, \forall i, j \in S$? Periodicity is an obstruction!

Definition (Period): The Period $d(i)$ of a state $i \in S$ is defined as

$$d(i) = \gcd\{ n \geq 1 \mid [P^n]_{ii} > 0\}.$$

In the 2-state Markov chain example with $a = b = 1$, $d(1) = d(2) = 2$.

Definition: A state $i \in S$ is periodic if $d(i) > 1$, and aperiodic if $d(i) = 1$.

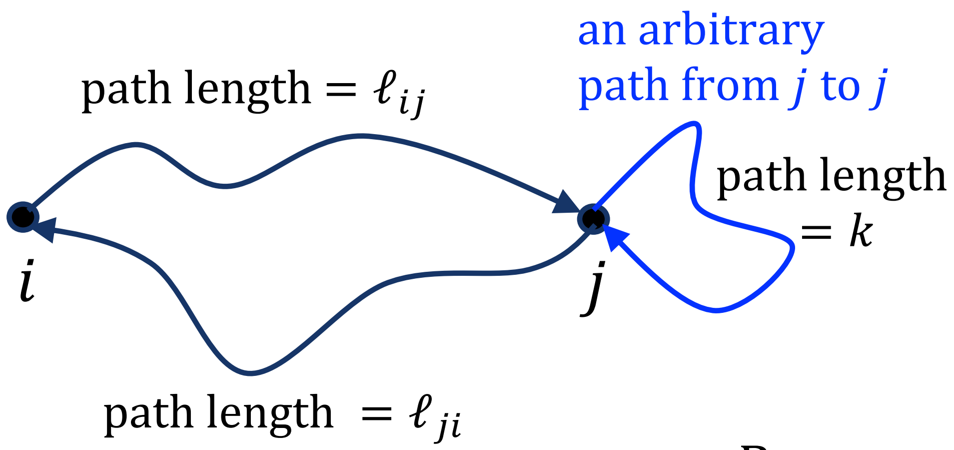
Theorem: If $i \leftrightarrow j$, then $d(i) = d(j)$

Period

Theorem: If $i \leftrightarrow j$, then $d(i) = d(j)$

Proof. Recall $d(i) = \gcd\{ n \geq 1 \mid [P^n]_{ii} > 0\}$.

We will show that $d(i)|d(j)$ and $d(j)|d(i)$, which implies $d(i) = d(j)$.
Since $i \leftrightarrow j$, there is a path from i to j and a path from j to i.



$d(i) |\ \ell_{ij} + \ell_{ji} + k$

$d(i) |\ \ell_{ij} + \ell_{ji}$

$\Rightarrow d(i) |\ k$ for all paths from j to j

Since $d(j)$ is a gcd for all such paths, their path lengths must be of the form $k_a = m_a d(j)$ where $\gcd\{m_a\} = 1$, so $d(i) |\ k_a$ for all a implies $d(i) |\ d(j)$.

By a symmetric argument, $d(j) \mid d(i)$. $\square$

A Directed Acyclic Graph of SCCs

Definition: The **return probability** of $i \in S$ is defined as

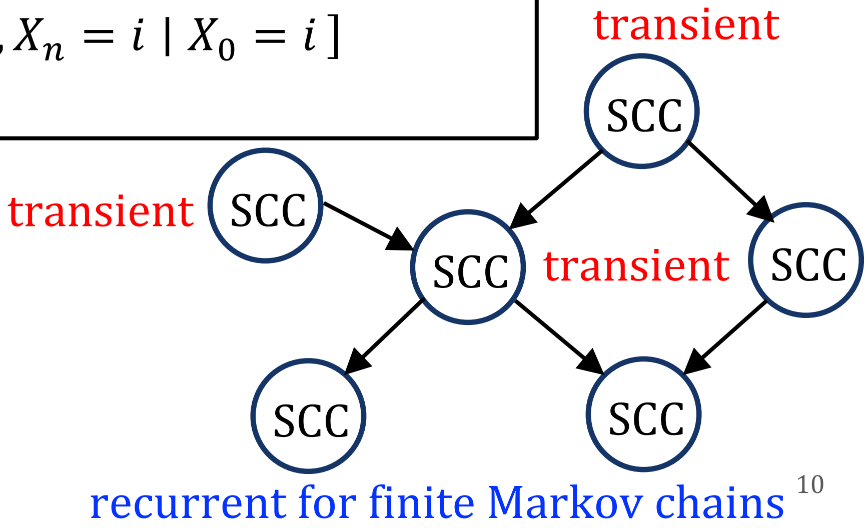
$$f_{ii} = \sum_{n=1}^{\infty} \mathbb{P}[X_1 \neq i, X_2 \neq i, \dots, X_{n-1} \neq i, X_n = i \mid X_0 = i]$$

Definition: The **mean recurrence time** of $i \in S$ is defined as

$$r_i = \sum_{n=1}^{\infty} n\, \mathbb{P}[X_1 \neq i, X_2 \neq i, \dots, X_{n-1} \neq i, X_n = i \mid X_0 = i]$$

A state $i \in S$ is called

- recurrent if $f_{ii} = 1$
    - null recurrent if $r_i = \infty$
    - positive recurrent if $r_i < \infty$
- transient if $f_{ii} < 1$.



Existence and Uniqueness of Stationary Distribution

A more refined version of the Perron–Frobenius Theorem:

Theorem (Finite Markov Chains):

- Every Markov chain with finite S has a stationary distribution.
- If the Markov chain is irreducible, then it has a unique stationary distribution $\pi = (\pi_i)_{i \in S}$.
- Furthermore, $\pi_i = \frac{1}{r_i}, \forall i \in S$, where r_i is the mean recurrence time.

Limit Theorem

For the n-step transition matrix P^n to converge as $n \to \infty$, aperiodicity in crucial.

Theorem (Limiting Distributions):

1. If a Markov chain (finite or not) is irreducible and aperiodic, then
$$\lim_{n\to\infty} [P^n]_{ij} = \frac{1}{r_j}, \forall i,j \in S.$$
2. If a Markov chain (finite or not) is irreducible, aperiodic, and positive recurrent, then
$$\lim_{n\to\infty} [P^n]_{ij} = \frac{1}{r_j} = \pi_j, \forall i,j \in S.$$

- In both cases, note that the limit does not depend on the starting state i.
- For finite Markov chains, irreducibility ⇒ all states are positive recurrent.
- Hence, if a **finite** Markov chain is irreducible and aperiodic, then
$$\lim_{n\to\infty} [P^n]_{ij} = \pi_j, \forall i,j \in S.$$
In Note 21, this is referred to as **"Fundamental Theorem of Markov Chain".**

Additional Results

Theorem (Class Property): Suppose two states $i, j \in S$ of a Markov chain inter-communicate ($i \leftrightarrow j$). Then,

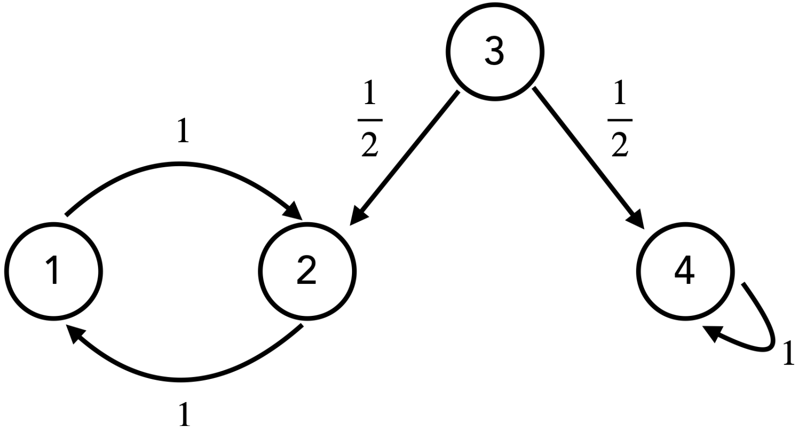
1. i and j have the same period.
2. i is transient if and only if j is transient.
3. i is null recurrent if and only if j is null recurrent.
4. i is positive recurrent if and only if j is positive recurrent.

(Note: Properties 2-4 are not required for CS70)

Theorem: All **finite** Markov chains have the following properties:

1. At least one state is recurrent.
2. All recurrent states are positive recurrent.
3. If the Markov chain is irreducible, then all states are positive recurrent.

Markov Chain 1 (periodic and not irreducible)

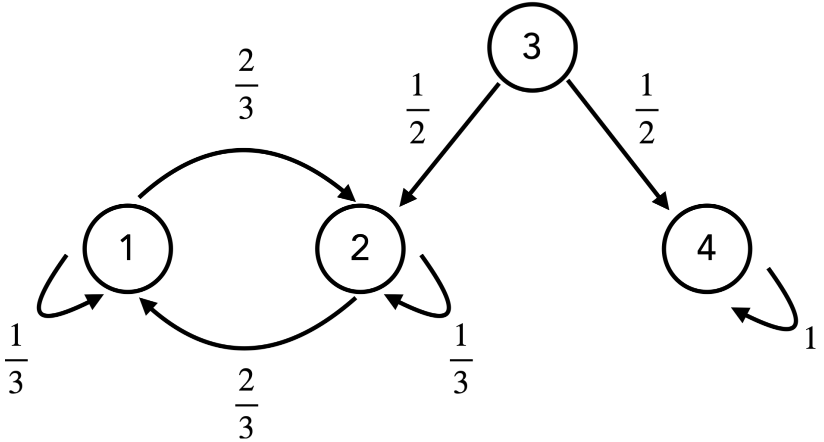


$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & ½ & 0 & ½ \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P^{2k} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & ½ & 0 & ½ \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad P^{2k+1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ ½ & 0 & 0 & ½ \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- States 1 and 2 are periodic, so P^n does **not** converge as $n \to \infty$.
- State 3 is transient, while states 1, 2, and 4 are positive recurrent.
- {1,2} is a terminal strongly connected component (SCC), so the transition matrix restricted to {1,2} is a valid transition matrix for a Markov chain on {1,2}.
- There exists a unique stationary distribution corresponding to this SCC. More precisely, [½, ½, 0, 0] is the unique stationary distribution for this SCC.
- {4} also is a terminal SCC and the unique stationary distribution corresponding to this SCC is [0, 0, 0, 1].

Markov Chain 2 (aperiodic but not irreducible)

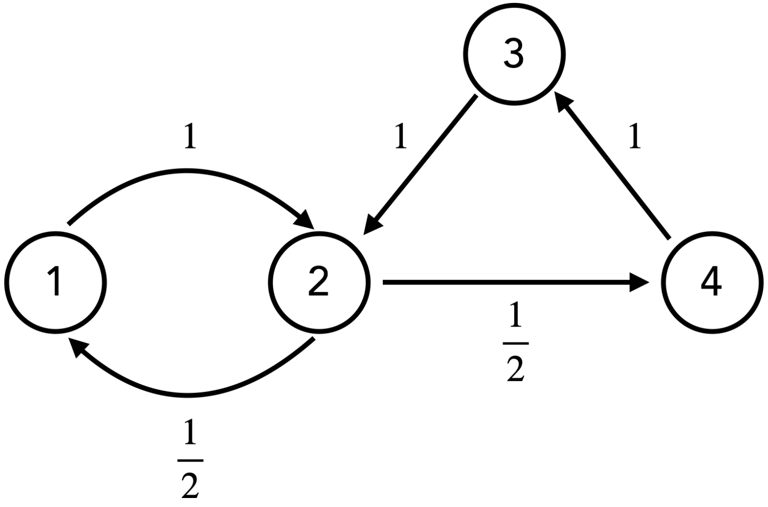


$$P = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 \\ 2/3 & 1/3 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- All states are aperiodic, so P^n converges as $n \to \infty$.
- State 3 is transient, and states 1, 2, and 4 are positive recurrent.
- {1,2} is a terminal SCC, so the transition matrix restricted to {1,2} is a valid transition matrix for a Markov chain on {1,2}.
- There exists a unique stationary distribution corresponding to this SCC. More precisely, [½, ½, 0, 0] is the unique stationary distribution for this SCC.
- {4} also is a terminal SCC and the unique stationary distribution corresponding to this SCC is [0, 0, 0, 1].

Markov Chain 3 (irreducible and aperiodic)



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

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 1/5 & 2/5 & 1/5 & 1/5 \\ 1/5 & 2/5 & 1/5 & 1/5 \\ 1/5 & 2/5 & 1/5 & 1/5 \\ 1/5 & 2/5 & 1/5 & 1/5 \end{bmatrix}$$

All rows are equal to the stationary distribution

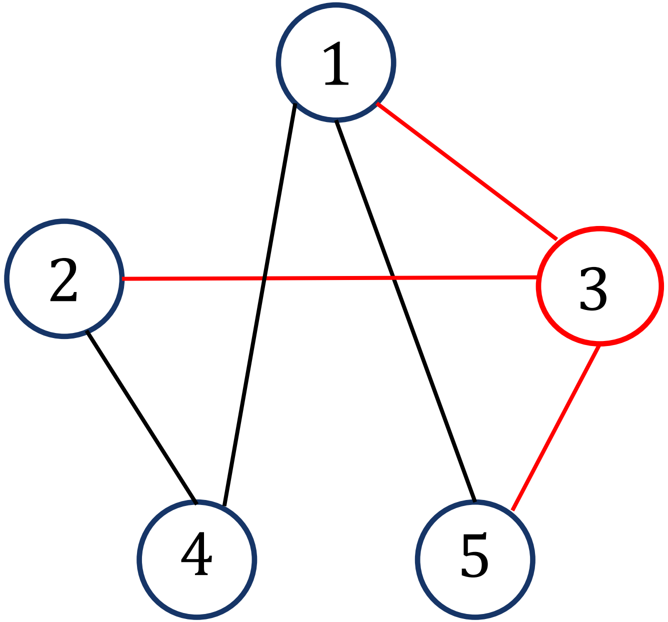
- All states are aperiodic, so P^n converges as $n \to \infty$.
- This Markov chain is irreducible ({1,2,3,4} is a SCC), so all of its states are positive recurrent (by Theorem 2).
- There exists a unique stationary distribution corresponding to this Markov chain. More precisely, $[1/5, 2/5, 1/5, 1/5]$ is the unique stationary distribution.

Random Walk on an Undirected Graph

Let $G = (V, E)$ be a **finite, connected, undirected** graph.

Given that the current state is u, move to a neighbor with probability $\frac{1}{\deg(u)}$.

State space $S = V$



Connectedness implies that the Markov chain is **irreducible**.

$$P_{31} = P_{32} = P_{35} = \frac{1}{3}$$

Random Walk on an Undirected Graph

Let $G = (V, E)$ be a finite, connected, undirected graph.

Given that the current state is u, move to a neighbor with probability $\frac{1}{\deg(u)}$.

State space $S = V$

Theorem: Random walk on a finite, connected, undirected graph has the following unique stationary distribution:

$$\pi_v = \frac{\deg(v)}{2|E|}, \forall v \in V.$$

Proof. For all $v \in V$, $\pi_v = \frac{\deg(v)}{2|E|} > 0$ and $\sum_{v \in V} \pi_v = \sum_{v \in V} \frac{\deg(v)}{2|E|} = 1$. Hence $\pi = (\pi_v)_{v \in V}$ is a valid probability distribution. We now show that it's stationary:

$$[\pi P]_v = \sum_{u:\{u,v\} \in E} \pi_u P_{uv} = \sum_{u \in \text{Neighbor}(v)} \frac{\deg(u)}{2|E|} \frac{1}{\deg(u)} = \sum_{u \in \text{Neighbor}(v)} \frac{1}{2|E|} = \frac{\deg(v)}{2|E|} = \pi_v$$

Further Probability and Statistics Courses at Berkeley

Undergraduate:
EECS126: Probability and Random Processes
CS174: Combinatorics and Discrete Probability
Stat134: Concepts of Probability
Stat135: Concepts of Statistics
Stat150: Stochastic Processes

Graduate:
EE226A: Random Processes in Systems
EE 226B. Applications of Stochastic Process Theory
CS 281A/B: Statistical Learning Theory
STAT 201A: Introduction to Probability at an Advanced Level
STAT 201B: Introduction to Statistics at an Advanced Level
STAT 205A/B: Probability Theory
STAT 210A/B: Theoretical Statistics