NEU Summer Course
Applied Stochastic Modeling and Queues

Topic 2: Discrete Time Markov Chain

Yunan Liu

Department of Industrial and Systems Engineering
NC State University
Outline

- Examples
- Definitions
  - Markov property
  - Transition probabilities
- Transient behavior
  - Chapman-Kolmogorov equation
  - Gambler’s ruin problem
- State classification
  - Communication
  - Periodicity
  - Recurrence and transience
- Steady state of “good” DTMCs
  - Balance equation $\pi P = \pi$
  - Long-run proportion of time in a state $j$
- Reducible DTMCs having absorbing states
  - Canonical form
  - Probability of absorption
- Time reversibility
  - Random walks on weighted graphs
Examples

1. Economics (Asset pricing, market crashes)
2. Chemistry (Enzyme activity, growth of certain compounds)
3. Internet Applications (Google PageRank)
4. Physics
5. Information sciences
6. Sports, etc.

Figure: A DTMC system modeling a remote medical assistance system
An Example: PageRank Algorithm

- Application: Best-known search engine algorithm
- Idea: Counting the number of links to a webpage to estimate its importance
- Developer: Larry Page
Definition and Transient Performance
Markov Processes

Figure: Andrei Markov (1856-1922)

Markov Property:
Given present state, forecast of future is independent of past

\[ P(\text{Future}|\text{Present, Past}) = P(\text{Future}|\text{Present}) \]
Markov Processes

A process that follows the Markov Property is called a Markov Process.

\[
\begin{array}{c|c|c}
\text{Time} & \text{Discrete} & \text{Continuous} \\
\hline
\text{State Space} & \begin{cases}
\text{Discrete} & \text{DTMC} \\
\text{Continuous} & \text{CTMC}
\end{cases} & \begin{cases}
e.g., \text{random walks} & \text{e.g., Brownian motions}
\end{cases}
\end{array}
\]

This class will focus on the discrete space Markov processes. (DTMC and CTMC)

Examples:

Discrete State Space
- Number of customers in queue
- Inventory Level
- Location (with discrete elements)

Continuous State Space
- Temperature Level
- Geographic Location
- Amount of time left in process
Definition

Definition (DTMC)
A discrete time discrete space stochastic process \( \{X_n, n = 0, 1, \ldots\} \) is called a DTMC if

\[
P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1} \ldots X_0 = i_0) = P(X_{n+1} = j|X_n = i) = P_{i,j}
\]

Markov Property - “Given present (today), future (tomorrow) is independent of past (yesterday)”

Example: Tomorrow will be sunny with an 80% chance if today is also sunny or with a 40% chance if today is rainy. Set up a DTMC for weather forecasting.
Transition Probability Matrix (TPM)

Let $\mathbf{P}$ be a transition probability matrix (TPM) with elements defined as

$$P_{i,j} = \mathbb{P}(X_{n+1} = j | X_n = i)$$

Properties of TPM $\mathbf{P}$:

- **Nonnegativity**: $P_{i,j} \geq 0$ for all $i, j \in S$
- **Row sum**: $\sum_{j \in S} P_{i,j} = \sum_{j \in S} \mathbb{P}(X_{n+1} = j | X_n = i) = \sum_{j \in S} \frac{\mathbb{P}(X_{n+1} = j, X_n = i)}{\mathbb{P}(X_n = i)} = 1$

Remark:

- DTMC may in general be time nonhomogeneous:

$$P_{i,j}(n) = \mathbb{P}(X_{n+1} = j | X_n = i).$$

- In this class, we focus on time homogeneous DTMC.
Dynamics of DTMC

\( \{X_n, n \geq 0\} \) evolves in a Markovian manner:

- The chain stays in a state for 1 time unit;
- When in state \( i \), \( X_n \) moves to a state \( j \) w.p. \( P_{i,j} \), \( i, j \in S \).

Sample path of a DTMC with state space \( S = \{1, 2, 3, 4, 5, 6\} \).
Questions of Interest

1. Transient performance (short-term behavior)
   Transition probabilities shortly after the process begins?
   ▶ $P(X_1 = j | X_0 = i) =$
   ▶ $P(X_1 = i, X_4 = j, X_7 = k) =$

2. Steady-state performance (long-run behavior)
   ▶ Asymptotic performance after many transitions: $P(X_{100} = j | X_0 = i) \approx ?$
   ▶ Long-run proportion of time (steps): for $N$ large, $\frac{1}{N} \sum_{n=1}^{N} 1_{\{X_n = j\}} \approx ?$
   ▶ Termination
     * Probability of which state the process will terminate in
     * Expected number of transitions until termination.
Transient Behavior
Chapman-Kolmogorov Equation

Define the $n$-step transition probability from $i \to j$

- $P_{i,j}^{(n)} \equiv \mathbb{P}(X_n = j | X_0 = i)$ for all $i, j \in S$
- $n$-step TPM $P^{(n)}$ has element $P_{i,j}^{(n)}$
- $P_{i,j}^{(1)} = P_{i,j}$

Review:

- $\mathbb{P}(A) = \sum_k \mathbb{P}(A \cap C_k) = \sum_k \mathbb{P}(A | C_k) \mathbb{P}(C_k)$ (LoTP)
- $\mathbb{P}(A | B) = \sum_k \mathbb{P}(A | B \cap C_k) \mathbb{P}(C_k | B)$ (Conditional LoTP)
  (Proved in HW03)
Derivation of C-K Equation

\[(P^{m+n})_{i,j} = P^{(m+n)}_{i,j} = P(X_{n+m} = j|X_0 = i)\]

\[= \sum_k P(X_{n+m} = j|X_m = k, X_0 = i) \cdot P(X_m = k|X_0 = i)\]

\[= \sum_k P(X_{n+m} = j|X_m = k) \cdot P(X_m = k|X_0 = i)\]

\[= \sum_k P^{(m)}_{i,k} \times P^{(n)}_{k,j}\]

\[= \left( P^{(m)} \times P^{(n)} \right)_{i,j}\]

Chapman-Kolmogorov (C-K) equation

\[P^{(m+n)} = P^{(m)} \times P^{(n)} = P^{(n)} \times P^{(m)} \quad \text{(in matrix notation)}\]

The C-K Equation allows us to relate the \((m + n)\) step transition matrix to the product of two smaller step-size \((m)\) and \((n)\) transition matrices.
Derivation of C-K Equation

A quick implication of the C-K Equation:

\[ P^{(n)} = P^{(n-1)} \times P^{(1)} = P^{(n-1)} \times P = \left( P^{(n-2)} \times P \right) \times P \]
\[ = \left( P^{(n-3)} \times P \right) \times P \times P = \ldots = \left( P^{(2)} \times P \right) \times P \times \ldots \times P \]
\[ = P \times P \times \ldots \times P \]

Summary: \( n \)-step TPM is the \( n^{th} \) power of the 1-step TPM

\[ P^{(n)} = P^n \]
Finite-Dimensional Distribution (F.D.D.)

For $0 \leq n_1 \leq n_2 \leq \cdots \leq n_m$, $i_1, i_2, \ldots, i_m \in S$

$$\mathbb{P}(X_{n_1} = i_1, X_{n_2} = i_2, \ldots, X_{n_m} = i_m)$$

$$\mathbb{P}(X_{10} = 1, X_7 = 5, X_3 = 10) = \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A|B \cap C) \cdot \mathbb{P}(B \cap C) = \mathbb{P}(A|B \cap C) \cdot \mathbb{P}(B|C) \cdot \mathbb{P}(C)$$

$$= \mathbb{P}(X_{10} = 1|X_7 = 5, X_3 = 10) \cdot \mathbb{P}(X_7 = 5|X_3 = 10) \cdot \mathbb{P}(X_3 = 10)$$

$$= \mathbb{P}(X_{10} = 1|X_7 = 5) \cdot \mathbb{P}(X_7 = 5|X_3 = 10) \cdot \mathbb{P}(X_3 = 10)$$

$$= P_{5,1}^{(3)} \times P_{10,5}^{(4)} \times \sum_{k \in S} \mathbb{P}(X_3 = 10|X_0 = k) \cdot \mathbb{P}(X_0 = k)$$

$$= P_{5,1}^{(3)} \cdot P_{10,5}^{(4)} \cdot \sum_{k \in S} P_k^{(3)} \cdot \mathbb{P}(X_0 = k) = (P^3)^{5,1} \cdot (P^4)^{10,5} \cdot \sum_{k \in S} (P^3)_{k,10} \cdot \mathbb{P}(X_0 = k)$$

Only need

- Distribution of initial state $\mathbb{P}(X_0 = k)$
- 1-step TPM $P$. 
An Example: Pooh Bear

A bear of little brain named *Pooh* is fond of honey. Bees producing honey are located in three trees: tree A, tree B, and tree C. Tending to be somewhat forgetful, Pooh goes back and form among these three honey trees randomly:

- from A, Pooh goes next to B or C with probability 1/2;
- from B, Pooh goes next to A w.p 3/4 and to C w.p. 1/4;
- from C, Pooh always goes next to A.

Construct a DTMC and specify its TPM.
State Classification
Communication

Two states \( i, j \in S \)

- **Accessibility**: denoted by \( i \rightarrow j \)
  
  State \( j \) is *accessible* from state \( i \) if there exists \( n \geq 0 \) s.t. \( P_{i,j}^{(n)} > 0 \)

- **Communication**: denoted by \( i \leftrightarrow j \)
  
  States \( i \) and \( j \) *communicate* if \( i \rightarrow j \) and \( j \rightarrow i \)
Communication

Two states $i, j \in S$

- **Accessibility:** denoted by $i \rightarrow j$
  
  State $j$ is *accessible* from state $i$ if there exists $n \geq 0$ s.t. $P_{i,j}^{(n)} > 0$

- **Communication:** denoted by $i \longleftrightarrow j$
  
  States $i$ and $j$ *communicate* if $i \rightarrow j$ and $j \rightarrow i$

Properties:

(i) $i \longleftrightarrow i$ (note when $n = 0$, $P^{(0)} = I$)

(ii) if $i \longleftrightarrow j$, then $j \longleftrightarrow i$

(iii) if $i \longleftrightarrow j$, $j \longleftrightarrow k$, then $i \longleftrightarrow k$
Communication

Two states $i, j \in S$

- Accessibility: denoted by $i \rightarrow j$
  State $j$ is accessible from state $i$ if there exists $n \geq 0$ s.t. $P_{i,j}^{(n)} > 0$

- Communication: denoted by $i \leftrightarrow j$
  States $i$ and $j$ communicate if $i \rightarrow j$ and $j \rightarrow i$

Properties:

(i) $i \leftrightarrow i$ (note when $n = 0$, $P^{(0)} = I$)
(ii) if $i \leftrightarrow j$, then $j \leftrightarrow i$
(iii) if $i \leftrightarrow j$, $j \leftrightarrow k$, then $i \leftrightarrow k$

Proof of (iii):

$i \leftrightarrow j \Rightarrow$ there exists $m$ s.t. $P_{i,j}^{(m)} > 0$

$j \leftrightarrow k \Rightarrow$ there exists $n$ s.t. $P_{j,k}^{(n)} > 0$

$P_{i,k}^{(m+n)} = \sum_{j \in S} P_{i,j}^{(m)} \times P_{j,k}^{(n)} > P_{i,j}^{(m)} \times P_{j,k}^{(n)} > 0$

Hence, $i \rightarrow k$. Repeat to show $k \rightarrow i$. 

Yunan Liu (NC State University)
Irreducibility and Periodicity

1. **Class**: We say $i, j$ belong to the same class, if $i \leftrightarrow j$

2. **Irreducible**: We say a DTMC is irreducible if there is only one class.

**Periodicity**:

1. The period of a state $i$, denoted $d(i)$, is the greatest common divisor of all $n \geq 0$ s.t. $P_{i,i}^{(n)} > 0$

2. If $d(i) = 1$ we say $i$ is aperiodic.

**Property of Periodicity**:

Periodicity (aperiodicity) is a class property, that is

$$\text{if } i \leftrightarrow j, \text{ then } d(i) = d(j)$$
Periodicity

Proof:
Assume $P_{i,i}^{(s)} > 0$ for some $s > 0$

- $i \leftrightarrow j \Rightarrow \text{exist } m, n \geq 0 \ P_{i,j}^{(m)} > 0, \ P_{j,i}^{(n)} > 0$

- $P_{j,j}^{(n+m)} \geq P_{j,i}^{(n)} \times P_{i,j}^{(m)} > 0$

- $P_{j,j}^{(n+m+s)} \geq P_{j,i}^{(n)} \times P_{i,i}^{(s)} \times P_{i,j}^{(m)} > 0$

- $d(j)$ is period of of $j \Rightarrow d(j)$ is a divisor of both $(n + m)$ and $(n + m + s)$

- Hence, $d(j)$ is a divisor of $s$

- Because $d(i)$ is the period of $i \Rightarrow d(j)$ divides $d(i)$

- Similarly, we can prove that $d(i)$ divides $d(j)$. Therefore, $d(i) = d(j)$. 
Ever Hitting Probabilities

Define the probability of visiting state $j$ from state $i$ in exactly $n$ steps:

$$f_{i,j}^{(n)} = \mathbb{P}(X_n = j, X_k \neq j, 1 \leq k \leq n - 1 | X_0 = i)$$

$$= \mathbb{P}(1^{st} \text{ transition to } j \text{ occurs at step } n | \text{ starting in } i)$$

Next define the probability that state $j$ will ever be visited after leaving state $i$:

$$f_{i,j} = \sum_{n=1}^{\infty} f_{i,j}^{(n)} = \mathbb{P}(j \text{ is ever visited after leaving } i)$$

Note:
- $f_{i,j} > 0$ if $i \rightarrow j$;
- $f_{i,j} = 0$ if $i \not\rightarrow j$
State Classification

A state $j$ of DTMC is

- **recurrent** if $f_{j,j} = 1$ (Starting in $j$, will always return to $j$ w.p.1.)
- **transient** if $f_{j,j} < 1$
  Starting in $j$, will ever (**never**) return to $j$ w.p. $f_{j,j}$ (w.p. $1 - f_{j,j} > 0$).

$N_{j,j} =$ total number of visits in $j$ after leaving $j$.

$$N_{j,j} = \begin{cases} \infty, & \text{if } j \text{ is recurrent} \\ \text{Geo}(1 - f_{j,j}) - 1, & \text{if } j \text{ is transient} \end{cases}$$

**A necessary and sufficient condition:**

- $j$ is recurrent $\iff \sum_{n=1}^{\infty} P_{j,j}^{(n)} = \infty$
- $j$ is transient $\iff \sum_{n=1}^{\infty} P_{j,j}^{(n)} < \infty$

**Proof:**

$$\mathbb{E}[N_{j,j}] = \mathbb{E} \left[ \sum_{n=1}^{\infty} \mathbb{I}(X_n = j | X_0 = j) \right] = \sum_{n=1}^{\infty} P(X_n = j | X_0 = j) = \sum_{n=1}^{\infty} P_{j,j}^{(n)}$$
State Classification

Both transience & recurrence are class properties.

If \( i \leftrightarrow j \) (\( i, j \) belong to the same class)

- \( i \) is recurrent \( \iff \) \( j \) is recurrent
- \( i \) is transient \( \iff \) \( j \) is transient

Proof:

\[
i \leftrightarrow j \implies P_{i,j}^{(n)} > 0, \ P_{j,i}^{(m)} > 0 \quad \text{for some } m \geq 0, \ n \geq 0
\]

Because \( P_{j,j}^{(m+n+l)} \geq P_{j,i}^{(m)} P_{i,i}^{(l)} P_{i,j}^{(n)} \), we have

\[
\sum_{k=1}^{\infty} P_{j,j}^{(k)} \geq \sum_{k=m+n}^{\infty} P_{j,j}^{(k)} = \sum_{l=1}^{\infty} P_{j,j}^{(m+n+l)} \geq \sum_{l=1}^{\infty} P_{j,i}^{(m)} P_{i,i}^{(l)} P_{i,j}^{(n)} = P_{j,i}^{(m)} P_{i,j}^{(n)} \sum_{l=1}^{\infty} P_{i,i}^{(l)}.
\]
Gambler’s Ruin Problem

A gambler who at each round has probability $p$ of winning one unit and probability $q = 1 - p$ of losing one unit. The game ends either

- the gambler’s total fortune reaches level $N > 0$, or
- his fortune reaches 0 (the gambler is broke).

Let $X_0$ be the initial fortune. Define the $X_i$, $i \geq 1$, as the outcome of $i^{th}$ bet:

$$X_i = \begin{cases} +1, & \text{w.p. } p \\ -1, & \text{w.p. } q = 1 - p \end{cases}$$

The gambler’s fortune after the $n^{th}$ bet:

$$S_n = \sum_{i=0}^{n} X_i.$$

**Question:** Is $\{S_n, n \geq 0\}$ a DTMC? If yes, how many classes? Which are transient and which are recurrent? What are the periods?
Probability of Winning

Define $P_i = \text{prob. gambler’s fortune hits } N \text{ before } 0 \text{ (bankrupt), starting in } i$

$P_i = \mathbb{P}(\text{hits } N \text{ before } 0|X_0 = i)$

$= \mathbb{P}(\text{hits } N \text{ before } 0|X_1 = 1, X_0 = i) \cdot p + \mathbb{P}(\text{hits } N \text{ before } 0|X_1 = -1, X_0 = i) \cdot q$

Therefore, $P_i \cdot p + P_i \cdot q = P_i = P_{i+1} \cdot p + P_{i-1} \cdot q$

$P_{i+1} - P_i = \gamma(P_i - P_{i-1})$, \hspace{1cm} $\gamma = \frac{q}{p}$. 
Probability of Winning

Define $P_i = \text{prob. gambler’s fortune hits } N \text{ before 0 (bankrupt), starting in } i$

\[
P_i = \mathbb{P}(\text{hits } N \text{ before 0}|X_0 = i) = \mathbb{P}(\text{hits } N \text{ before 0}|X_1 = 1, X_0 = i) \cdot p + \mathbb{P}(\text{hits } N \text{ before 0}|X_1 = -1, X_0 = i) \cdot q
\]

Therefore,

\[
P_i \cdot p + P_i \cdot q = P_i = P_{i+1} \cdot p + P_{i-1} \cdot q
\]

\[
P_{i+1} - P_i = \gamma(P_i - P_{i-1}), \quad \gamma = \frac{q}{p}.
\]

A recursion with boundary conditions $P_0 = 0 \text{ and } P_N = 1$

\[
P_2 - P_1 = \gamma(P_1 - P_0) = \gamma P_1
\]

\[
P_3 - P_2 = \gamma(P_2 - P_1) = \gamma^2 P_1
\]

\[\vdots\]

\[
P_i - P_{i-1} = \gamma(P_{i-1} - P_{i-2}) = \gamma^{i-1} P_1
\]

\[\vdots\]

\[
P_N - P_{N-1} = \gamma(P_{N-1} - P_{N-2}) = \gamma^{N-1} P_1
\]
Probability of Winning

Summing up the first $i - 1$ equations yields:

$$P_i = \left[1 + \frac{q}{p} + \cdots + \left(\frac{q}{p}\right)^{i-1}\right] \cdot P_1 = \begin{cases} \frac{1 - \left(\frac{q}{p}\right)^i}{1 - \frac{q}{p}} \cdot P_1, & \text{if } p \neq q \\ i \cdot P_1, & \text{if } p = q. \end{cases}$$
Probability of Winning

Summing up the first \( i - 1 \) equations yields:

\[
P_i = \left[ 1 + \frac{q}{p} + \cdots + \left( \frac{q}{p} \right)^{i-1} \right] \cdot P_1 = \begin{cases} \frac{1 - \left( \frac{q}{p} \right)^i}{1 - \frac{q}{p}} \cdot P_1, & \text{if } p \neq q \\ i \cdot P_1, & \text{if } p = q. \end{cases}
\]

Summing up all \( N \) equations yields:

\[
1 = P_N = \left[ 1 + \frac{q}{p} + \cdots + \left( \frac{q}{p} \right)^{N-1} \right] \cdot P_1 = \begin{cases} \frac{1 - \left( \frac{q}{p} \right)^N}{1 - \frac{q}{p}} \cdot P_1, & \text{if } p \neq q \\ N \cdot P_1, & \text{if } p = q. \end{cases}
\]
Probability of Winning

Summing up the first $i - 1$ equations yields:

$$P_i = \left[1 + \frac{q}{p} + \cdots + \left(\frac{q}{p}\right)^{i-1}\right] \cdot P_1 = \begin{cases} 
\frac{1-(\frac{q}{p})^i}{1-\frac{q}{p}} \cdot P_1, & \text{if } p \neq q \\
i \cdot P_1, & \text{if } p = q.
\end{cases}$$

Summing up all $N$ equations yields:

$$1 = P_N = \left[1 + \frac{q}{p} + \cdots + \left(\frac{q}{p}\right)^{N-1}\right] \cdot P_1 = \begin{cases} 
\frac{1-(\frac{q}{p})^N}{1-\frac{q}{p}} \cdot P_1, & \text{if } p \neq q \\
N \cdot P_1, & \text{if } p = q.
\end{cases}$$

Finally,

$$P_i = \begin{cases} 
\frac{1-(\frac{q}{p})^i}{1-(\frac{q}{p})^N}, & \text{if } p \neq q \text{ (unfair game)} \\
\frac{i}{N}, & \text{if } p = q = 1/2 \text{ (fair game)}.
\end{cases}$$
State Classification (Continued)

Denote

$$T_{j,j} = \text{total \ # \ of \ steps \ until \ returning \ to \ } j \ \text{the 1}\text{st, after leaving } j$$

**Definition:** A recurrent state \( j \) is

- **positive recurrent** if \( \mathbb{E}[T_{j,j}] < \infty \);
- **null recurrent** if \( \mathbb{E}[T_{j,j}] = \infty \).

It is clear that if a class is transient, then w.p. \( 1 - f_{j,j} > 0 \) DTMC never returns.

Summary, for a state \( j \),

$$\mathbb{E}[T_{j,j}] \begin{cases} 
= \infty, & \text{if } j \text{ is transient} \\
< \infty, & \text{if } j \text{ is positive recurrent} \\
= \infty, & \text{if } j \text{ is null recurrent}
\end{cases}$$

Remark:
A null recurrence state will “always” (w.p.1.) return to itself, but it will take nearly “forever” to do so...
An Example: One-Dimensional Simple Random Walks

Consider a simple random walk

- $X_1, X_2, \ldots$ I.I.D. $P(X_1 = 1) = p$, $P(X_1 = -1) = 1 - p$, $0 < p < 1$.
- $S_0 = 0$, $S_n = X_1 + \cdots + X_n$, $n \geq 1$.

The process $\{S_n, n \geq 0\}$ is a simple symmetric (asymmetric) random walk if $p = 1/2$ ($p \neq 1/2$).

Questions:

- Is $\{S_n, n \geq 0\}$ a DTMC? If so, what is the state space? Is it irreducible?
- Classify the states (transient, null-recurrent, positive-recurrent).
Summary of State Classification

A state $i$ is

- transient:
  \[
  \text{if } f_{j,j} = \mathbb{P}(T_{j,j} < \infty) < 1 \Leftrightarrow \sum_{k=1}^{\infty} P_{j,j}^{(k)} < \infty \Leftrightarrow N_{j,j} \sim \text{Geo}(1 - f_{j,j})
  \]

- recurrent:
  \[
  \text{if } f_{j,j} = \mathbb{P}(T_{j,j} < \infty) = 1 \Leftrightarrow \sum_{k=1}^{\infty} P_{j,j}^{(k)} = \infty \Leftrightarrow N_{j,j} = \infty
  \]
  - positive recurrent: if $\mathbb{E}[T_{j,j}] < \infty$
  - null recurrent: if $\mathbb{E}[T_{j,j}] = \infty$
Summary of State Classification

A state \( i \) is

- transient:
  \[
  \text{if } f_{j,j} = \mathbb{P}(T_{j,j} < \infty) < 1 \Leftrightarrow \sum_{k=1}^{\infty} P_{j,j}^{(k)} < \infty \Leftrightarrow N_{j,j} \sim \text{Geo}(1 - f_{j,j})
  \]

- recurrent:
  \[
  \text{if } f_{j,j} = \mathbb{P}(T_{j,j} < \infty) = 1 \Leftrightarrow \sum_{k=1}^{\infty} P_{j,j}^{(k)} = \infty \Leftrightarrow N_{j,j} = \infty
  \]
    - positive recurrent: if \( \mathbb{E}[T_{j,j}] < \infty \)
    - null recurrent: if \( \mathbb{E}[T_{j,j}] = \infty \)

All following features are class properties:

1. Communication
2. Transience
3. Recurrence (null-recurrence, positive-recurrence)
4. Periodicity and aperiodicity
Long-Term Performance
Stationary Distribution

**Definition (Stationary distribution)**

A vector \( \pi = (\pi_1, \pi_2, \pi_3, \ldots) \) s.t. \( \pi_i \geq 0, \sum_{i \in S} \pi_i = 1 \) is called the *stationary distribution* for a DTMC with TPM \( P \), if

\[
\pi P = \pi \quad \text{or} \quad \sum_{i \in S} \pi_i P_{i,j} = \pi_j \quad \text{for all} \quad j \in S.
\]

**Implication:** If \( \mathbb{P}(X_0 = j) = \pi_j, j \in S \)

\[
\mathbb{P}(X_1 = j) = \sum_{i \in S} \mathbb{P}(X_1 = j | X_0 = i) \mathbb{P}(X_0 = i) = \sum_{i \in S} P_{i,j} \pi_i = \pi_j
\]

\[
\mathbb{P}(X_2 = j) = \sum_{i \in S} \mathbb{P}(X_2 = j | X_1 = i) \mathbb{P}(X_1 = i) = \sum_{i \in S} P_{i,j} \pi_i = \pi_j
\]

\[\vdots\]

\[
\mathbb{P}(X_n = j) = \pi_j \quad \text{for all} \quad n \geq 1.
\]

If a DTMC is initially in stationarity, it is in stationarity for all time \( n \).
Theorem 1: Convergence to Steady State

**Theorem (steady-state distribution)**

Consider an (i) **irreducible** and (ii) **aperiodic** DTMC \( \{X_n : n \geq 0\} \).

(i) If all states are transient or null-recurrent,
\[
P^{(n)}_{i,j} \rightarrow 0.
\]

(ii) If all states are positive recurrent,
\[
P^{(n)}_{i,j} \rightarrow \pi_j > 0,
\]

\( \pi = (\pi_1, \pi_2, \ldots) \) is the unique solution to
\[
\begin{cases} 
\pi P = \pi \\
\sum_{i \in S} \pi_i = 1 
\end{cases}
\]

(balance equation)

Equivalently,
\[
(X_n | X_0 = i) \xrightarrow{\mathcal{D}} X_\infty \text{ where } P(X_\infty = j) = \pi_j.
\]

Proof: omitted here, see p.175 – 177 of Blue Ross.
Theorem 1: Convergence to Steady State

Remarks of Theorem 1:
- Steady state is defined through a limit as $n \to \infty$.
- In case (i), there is no steady state;
  In case (ii), convergence occurs (to a limit).
- Definition: ergodicity = aperiodic + irreducibility + positive-recurrent.
- Given convergence in (ii), the limit must be $\pi$.

\[ \pi_j^* = \lim_{n \to \infty} P_{i,j}^{(n+1)} = \lim_{n \to \infty} \sum_k P_{i,k}^{(n)} P_{k,j} = \sum_k \lim_{n \to \infty} P_{i,k}^{(n)} P_{k,j} = \sum_k \pi_k^* P_{k,j}. \]

- All states of a finite state space irreducible DTMC are positive-recurrent.
- If a DTMC is “good” (irreducible and positive-recurrent), then the balance equation has a unique solution (aperiodicity is not necessary)
Corollary (Existence of “good” state in a finite DTMC)

Consider a DTMC with a finite state space. There must exist at least one positive recurrent state.
Theorem 1: Implication

Corollary (Existence of “good” state in a finite DTMC)

Consider a DTMC with a finite state space. There must exist at least one positive recurrent state.

Proof:
Assume no state is positive recurrent - all states are either transient or null recurrent to reach a contradiction.

- By Theorem 1: \( \lim_{n \to \infty} P_{i,j}^{(n)} = 0 \) for all \( j \in S = \{1, 2, \ldots, N\} \).
- Row sum \( \sum_{j=1}^{N} P_{i,j}^{(n)} = 1 \) and take \( n \to \infty \):

\[
1 = \lim_{n \to \infty} \sum_{j=1}^{N} P_{i,j}^{(n)} = \sum_{j=1}^{N} \lim_{n \to \infty} P_{i,j}^{(n)} = \sum_{j=1}^{N} 0 = 0.
\]

Contradiction! There must exist at least one positive recurrent state!
Theorem 1: Implication

Corollary (Existence of “good” state in a finite DTMC)

Consider a DTMC with a finite state space. There must exist at least one positive recurrent state.

Proof:
Assume no state is positive recurrent - all states are either transient or null recurrent to reach a contradiction.

- By Theorem 1: \( \lim_{n \to \infty} P_{i,j}^{(n)} = 0 \) for all \( j \in S = \{1, 2, \ldots, N\} \).
- Row sum \( \sum_{j=1}^{N} P_{i,j}^{(n)} = 1 \) and take \( n \to \infty \):

\[
1 = \lim_{n \to \infty} \sum_{j=1}^{N} P_{i,j}^{(n)} = \sum_{j=1}^{N} \lim_{n \to \infty} P_{i,j}^{(n)} = \sum_{j=1}^{N} 0 = 0.
\]

Contradiction! There must exist at least one positive recurrent state!

Remark:
For an irreducible and finite-state DTMC, all states are positive recurrent.
Theorem 2: Long-Run Proportion of Time

Theorem (long-run proportion of time)

(i) For an irreducible, positive-recurrent DTMC, BE has a unique solution, and the long-run proportion of time DTMC is in a state $k$:

$$
\pi_k = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}(X_n = k)
$$

(ii) Let $r(\cdot)$ be a bounded (cost) function on $S$, then the long-run average cost:

$$
\lim_{n \to \infty} \frac{1}{N} \sum_{n=1}^{N} r(X_n) = \sum_{j \in S} r(j)\pi_j = \sum_{j \in S} r(j)\mathbb{P}(X_\infty = j) = \mathbb{E}[r(X_\infty)]
$$

Remarks:

- (i) is a special case of (ii) when $r(X) = \mathbb{1}(X = k)$
- If $r(X) = X$, then (ii) becomes long-run average of $\{X_n\}$:

$$
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} X_n = \sum_{j \in S} j\pi_j = \mathbb{E}[X_\infty]
$$
Proof of Theorem 2

Define:

1. \( N_n(i) = \sum_{k=1}^{n} \mathbb{1}(X_k = i) \) = time (days) in state \( i \) by \( n \).
2. \( N_n(i, j) = \sum_{k=1}^{n} \mathbb{1}(X_{k-1} = i, X_k = j) \) = \# of direct trans from \( i \rightarrow j \) by \( n \).
3. \( N_n(i) \) = \sum_{\ell=1}^{N_n(i)} \mathbb{1}(\text{next visit is } j \text{ after } \ell^{th} \text{ visit in } i) = \sum_{\ell=1}^{N_n(i)} \mathbb{1} A_{i,j}^{(\ell)} \)

\( A_{i,j}^{(\ell)} \)
Proof of Theorem 2

Define:

- \( N_n(i) = \sum_{k=1}^{n} \mathbb{1}(X_k = i) \) = time (days) in state \( i \) by \( n \).

- \( N_n(i, j) = \sum_{k=1}^{n} \mathbb{1}(X_{k-1} = i, X_k = j) \) = \# of direct trans from \( i \rightarrow j \) by \( n \)

\[
N_n(i) = \sum_{\ell=1}^{N_n(i)} \mathbb{1}(\text{next visit is } j \mid \text{ after } \ell^{th} \text{ visit in } i) = \sum_{\ell=1}^{N_n(i)} \mathbb{1}_{A_{i,j}^{(\ell)}}
\]

Partial proof (i): let \( \pi_j^* = \lim_{n \to \infty} N_n(j)/n \).

\[
\frac{1}{n} N_n(j) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}(X_k = j) = \frac{1}{n} \sum_{i \in S} \sum_{k=1}^{n} \mathbb{1}(X_{k-1} = i, X_k = j)
\]

\[
= \sum_{i \in S} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}(X_{k-1} = i, X_k = j) = \sum_{i \in S} \frac{1}{n} N_n(i, j) = \sum_{i \in S} \frac{1}{n} \sum_{\ell=1}^{N_n(i)} \mathbb{1}_{A_{i,j}^{(\ell)}}
\]

\[
= \sum_{i \in S} \left( \frac{N_n(i)}{n} \right) \left( \frac{1}{N_n(i)} \sum_{\ell=1}^{N_n(i)} \mathbb{1}_{A_{i,j}^{(\ell)}} \right) \rightarrow \sum_{i \in S} \pi_i^* \mathbb{E}[\mathbb{1}_{A_{i,j}}] = \sum_{i \in S} \pi_i^* P_{i,j}.
\]
Proof of Theorem 2 Continued

The total cost by time $n$:

$$\sum_{k=1}^{n} r(X_k) = \sum_{k=1}^{n} \sum_{j \in S} \mathbb{1}(X_k = j) r(X_k) = \sum_{k=1}^{n} \sum_{j \in S} \mathbb{1}(X_k = j) r(j)$$

$$= \sum_{j \in S} \sum_{k=1}^{n} \mathbb{1}(X_k = j) r(j) = \sum_{j \in S} r(j) \sum_{k=1}^{n} \mathbb{1}(X_k = j)$$
Proof of Theorem 2 Continued

The total cost by time $n$:

$$\sum_{k=1}^{n} r(X_k) = \sum_{k=1}^{n} \sum_{j \in S} 1(X_k = j) r(X_k) = \sum_{k=1}^{n} \sum_{j \in S} 1(X_k = j) r(j)$$

$$= \sum_{j \in S} \sum_{k=1}^{n} 1(X_k = j) r(j) = \sum_{j \in S} r(j) \sum_{k=1}^{n} 1(X_k = j)$$

Dividing $n$ on both sides and taking $n \to \infty$:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} r(X_k) = \sum_{j \in S} r(j) \left( \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} 1(X_k = j) \right) = \sum_{j \in S} r(j) \pi_j.$$
Theorem 3: Mean First Passage Time

Consider an irreducible and positive-recurrent DTMC:

(i) If aperiodic: \( P_{i,j}^{(n)} \rightarrow \pi_j = \frac{1}{\mathbb{E}[T_{j,j}]} \)

(ii) If periodic with period \( d > 1 \):

- The sequence \( P_{j,j}^{(n)} \) doesn’t converge;
- The subsequence \( P_{j,j}^{n \cdot d} \rightarrow d \pi_j = \frac{d}{\mathbb{E}[T_{j,j}]} \)
Theorem 3: Mean First Passage Time

Theorem (Mean First Passing Time)

Consider an irreducible and positive-recurrent DTMC:

(i) If aperiodic: \( P_{i,j}^{(n)} \to \pi_j = \frac{1}{\mathbb{E}[T_{j,j}]} \)

(ii) If periodic with period \( d > 1 \):
- The sequence \( P_{j,j}^{(n)} \) doesn’t converge;
- The subsequence \( P_{j,j}^{n \cdot d} \to d \pi_j = \frac{d}{\mathbb{E}[T_{j,j}]} \)

Intuition:
- A cycle: time from \( j \) to \( j \)
- In each cycle: spend exactly 1 time unit (day) in \( j \)
- long-run proportion of time in \( j \): \( 1/\mathbb{E}[\text{cycle length}] = 1/\mathbb{E}[T_{j,j}] \).
Theorem 3: Mean First Passage Time

Proof:
Break up the evolution of CTMC into cycles where a cycle begins every time the chain visits $j$.

- $Y_n \equiv$ number of steps between the $(n - 1)^{\text{th}}$ and $n^{\text{th}}$ visit to $j$.
- DTMC visits state $j$ for the $n^{\text{th}}$ time at time $Y_1 + \cdots + Y_n$.
- Assume $X_0 = j$, $Y_1, \ldots, Y_n$ are I.I.D., $T_{j,j} = Y_1$
- $N_j(n) \equiv \sum_{k=1}^{n} 1\{X_k = j\} = \text{tot. number in } j \text{ at time } n$.

Long-run proportion of time in $j$:

$$
\pi_j = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} 1\{X_k = j\} = \lim_{n \to \infty} \frac{N_j(n)}{\sum_{i=1}^{N_j(n)} Y_i} \\
= \lim_{n \to \infty} \frac{1}{\sum_{i=1}^{N_j(n)} Y_i} \overset{\text{SLLN}}{\Rightarrow} \frac{1}{\mathbb{E}[T_{j,j}]}.
$$
Example: DTMC in Hospitals

When a patient is hospitalized at the NC State hospital, his/her health condition can be described by a DTMC with 3 states: low risk (L), medium risk (M) and high risk (H). Patients are admitted into the hospital in low-, medium- and high-risk conditions with probability $p_1 = 0.6$, $p_2 = 0.3$ and $p_3 = 0.1$, respectively. Their health status are updated once every day according to the TPM

$$
P = \begin{pmatrix}
H & 0.50 & 0.40 & 0.1 \\
M & 0.35 & 0.30 & 0.35 \\
L & 0.05 & 0.15 & 0.8
\end{pmatrix}.
$$

We are interested in:

1. Long-run proportion of days a patient is in low (medium, high) risk.
2. Daily medication costs are independent exponential r.v.’s with rates $\mu_1 = 1/2$, $\mu_2 = 1/4$ and $\mu_3 = 1/10$ for a low-, medium- and high-risk patient. What is long-run average daily cost?
Example: Google PageRank

Develop a DTMC:

1. Each state represents a web page;
2. Create direct transition (an arc) from state $i$ to state $j$ if page $i$ has a link to page $j$;
3. If page $i$ has $k > 0$ outgoing links, then set the probability on each outgoing arc to be $\frac{1}{k}$;
4. Solve the DTMC to determine the stationary distribution. Pages are then ranked based on their $\pi_i$. 
Reducible DTMCs
Canonical Form of Reducible DTMC

Consider a DTMC with both *transient* states and *absorbing* states. Define

- $\mathcal{A}$ = set of absorbing states $= \{i \in S, \text{ s.t. } P_{i,i} = 1\}$
- $\mathcal{T}$ = set of transient states

Suppose $|\mathcal{A}| = n, |\mathcal{T}| = m$
Canonical Form of Reducible DTMC

Consider a DTMC with both *transient* states and *absorbing* states. Define

- \( \mathcal{A} = \) set of absorbing states = \( \{i \in S, \text{ s.t. } P_{i,i} = 1\} \)
- \( \mathcal{T} = \) set of transient states

Suppose \( |\mathcal{A}| = n, |\mathcal{T}| = m \)

Relabel the states and put the TPM \( \mathbf{P} \) in its canonical form:

\[
\mathbf{P} = \begin{bmatrix}
\mathcal{A} & \mathcal{T} \\
\mathcal{T} & \end{bmatrix} = \begin{bmatrix}
I_{n \times n} & 0_{n \times m} \\
R_{m \times n} & Q_{m \times m} \\
\end{bmatrix}
\]

- \( Q_{m \times m} \): transient-to-transient TPM
- \( R_{m \times n} \): transient-to-absorbing TPM
- \( 0_{n \times m} \): 0 matrix
- \( I_{n \times n} \): identity matrix
An Example: Markov Mouse in an Open Maze

Question: Starting in room 1,
- what is the mean number of steps until escaping the maze?
- what is the probability of exiting from door 3?
An Example: Markov Mouse in an Open Maze

Question: Starting in room 1,
- what is the mean number of steps until escaping the maze?
- what is the probability of exiting from door 3?

In general, for a reducible DTMC
- \( \mathbb{E}[\text{number of steps until absorption}|X_0 = i], \text{ for } i \in \mathcal{T} \)
- \( \mathbb{P}(\text{being absorbed by state } j|X_0 = i), \text{ for } j \in \mathcal{A}, i \in \mathcal{T} \)
Mean Number of Steps until Absorption

For transient states $i, j \in \mathcal{T}$

$$N_{i,j} = \text{total \# of steps in } j \text{ starting in } i = \sum_{n=0}^{\infty} 1(X_n = j | X_0 = i).$$
Mean Number of Steps until Absorption

For transient states \( i, j \in \mathcal{T} \)

\[
N_{i,j} = \text{total \# of steps in } j \text{ starting in } i = \sum_{n=0}^{\infty} \mathbb{1}(X_n = j|X_0 = i).
\]

Define the fundamental matrix \( S \in \mathbb{R}^{m \times m} \):

\[
S_{i,j} \equiv \mathbb{E}[N_{i,j}] = \mathbb{E} \left[ \sum_{n=0}^{\infty} \mathbb{1}(X_n = j|X_0 = i) \right] = \sum_{n=0}^{\infty} \mathbb{E} \left[ \mathbb{1}(X_n = j|X_0 = i) \right] = \sum_{n=0}^{\infty} \mathbb{P}(X_n = j|X_0 = i) = \delta_{i,j} + \sum_{n=1}^{\infty} P^{(n)}_{i,j}
\]

\[
= \delta_{i,j} + \sum_{n=1}^{\infty} Q^{(n)}_{i,j}, \quad \text{where} \quad \delta_{i,j} = \mathbb{1}_{\{i=j\}}
\]
Mean Number of Steps until Absorption

For transient states $i, j \in \mathcal{T}$

$$N_{i,j} = \text{total \# of steps in } j \text{ starting in } i = \sum_{n=0}^{\infty} \mathbb{1}(X_n = j|X_0 = i).$$

Define the fundamental matrix $S \in \mathbb{R}^{m \times m}$:

$$S_{i,j} = \mathbb{E}[N_{i,j}] = \mathbb{E} \left[ \sum_{n=0}^{\infty} \mathbb{1}(X_n = j|X_0 = i) \right] = \sum_{n=0}^{\infty} \mathbb{E} [\mathbb{1}(X_n = j|X_0 = i)]$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(X_n = j|X_0 = i) = \delta_{i,j} + \sum_{n=1}^{\infty} P_{i,j}^{(n)}$$

$$= \delta_{i,j} + \sum_{n=1}^{\infty} Q_{i,j}^{(n)}, \quad \text{where} \quad \delta_{i,j} = \mathbb{1}\{i=j\}$$

In matrix notation

$$S = I + Q + Q^2 + Q^3 + \cdots \quad (\text{geometric sequence})$$
Mean Number of Steps until Absorption

Solving the fundamental matrix

\[ S = I + Q + Q^2 + Q^3 + \cdots \]
\[ QS = SQ = Q + Q^2 + Q^3 + Q^4 + \cdots \]
\[ S(I - Q) = (I - Q)S = I. \]

What’s missing? \( \lim_{n \to \infty} Q^n = 0. \)
Mean Number of Steps until Absorption

Solving the fundamental matrix

\[ S = I + Q + Q^2 + Q^3 + \cdots \]

\[ QS = SQ = Q + Q^2 + Q^3 + Q^4 + \cdots \]

\[ S(I - Q) = (I - Q)S = I. \]

What’s missing? \( \lim_{n \to \infty} Q^n = 0 \).

Fundamental matrix:

\[ S = (I - Q)^{-1} \]
Mean Number of Steps until Absorption

Solving the fundamental matrix

\[ S = I + Q + Q^2 + Q^3 + \cdots \]
\[ QS = SQ = Q + Q^2 + Q^3 + Q^4 + \cdots \]
\[ S(I - Q) = (I - Q)S = I. \]

What’s missing? \( \lim_{n \to \infty} Q^n = 0. \)

Fundamental matrix:

\[ S = (I - Q)^{-1} \]

Define \( M_i = \) mean number of steps until absorption, starting in \( i \in T. \)

\[ M_i = \sum_{j \in T} S_{i,j}, \quad \text{or} \quad M = S e, \quad \text{where} \quad e = [1, \ldots, 1]^T. \]
Probability of Absorption

For \( i \in \mathcal{T}, \ell \in \mathcal{A} \). Define

\[
B_{i,\ell} = \mathbb{P}(\text{eventually being absorbed by state } \ell | X_0 = i)
= \mathbb{P}(X_1 = \ell | X_0 = i) + \mathbb{P}(X_2 = \ell, X_1 \in \mathcal{T} | X_0 = i) + \mathbb{P}(X_3 = \ell, X_2 \in \mathcal{T} | X_0 = i) + \cdots
= R_{i,\ell} + \sum_{j \in \mathcal{T}} \mathbb{P}(X_2 = \ell | X_1 = j, X_0 = i) \mathbb{P}(X_1 = j | X_0 = i)
+ \sum_{j \in \mathcal{T}} \mathbb{P}(X_3 = \ell | X_2 = j, X_0 = i) \mathbb{P}(X_2 = j | X_0 = i) + \cdots
= R_{i,\ell} + \sum_{j \in \mathcal{T}} R_{j,\ell} Q_{i,j} + \sum_{j \in \mathcal{T}} R_{j,\ell} Q_{i,j}^{(2)} + \sum_{j \in \mathcal{T}} R_{j,\ell} Q_{i,j}^{(3)} + \cdots
= (R)_{i,\ell} + (QR)_{i,\ell} + (Q^2R)_{i,\ell} + (Q^3R)_{i,\ell} + \cdots
\]
Probability of Absorption

For \( i \in \mathcal{T}, \ell \in A \). Define

\[
B_{i, \ell} = \mathbb{P}(\text{eventually being absorbed by state } \ell|X_0 = i)
= \mathbb{P}(X_1 = \ell|X_0 = i) + \mathbb{P}(X_2 = \ell, X_1 \in \mathcal{T}|X_0 = i) + \mathbb{P}(X_3 = \ell, X_2 \in \mathcal{T}|X_0 = i) + \cdots
= R_{i, \ell} + \sum_{j \in \mathcal{T}} \mathbb{P}(X_2 = \ell|X_1 = j, X_0 = i)\mathbb{P}(X_1 = j|X_0 = i)
+ \sum_{j \in \mathcal{T}} \mathbb{P}(X_3 = \ell|X_2 = j, X_0 = i)\mathbb{P}(X_2 = j|X_0 = i) + \cdots
= R_{i, \ell} + \sum_{j \in \mathcal{T}} R_{j, \ell} Q_{i, j} + \sum_{j \in \mathcal{T}} R_{j, \ell} Q_{i, j}^{(2)} + \sum_{j \in \mathcal{T}} R_{j, \ell} Q_{i, j}^{(3)} + \cdots
= (R)_{i, \ell} + (QR)_{i, \ell} + (Q^2 R)_{i, \ell} + (Q^3 R)_{i, \ell} + \cdots
\]

In matrix notation:

\[
B = R + QR + Q^2 R + Q^3 R + \cdots = (I + Q + Q^2 + \cdots) \cdot R = SR = (I - Q)^{-1} R.
\]
Ever Hitting Probabilities Revisited

Using the fundamental matrix $S$, we can compute ever-hitting probability

$$f_{i,j} = \Pr(j \text{ ever visited } | \text{ after leaving } i) < 1.$$  

For $i, j \in \mathcal{T}, i \neq j$

$$S_{i,j} = \mathbb{E}[\text{# of visits in } j | X_0 = i]$$

$$= \mathbb{E}[\text{# of visits in } j | j \text{ is ever visited, } X_0 = i] \cdot f_{i,j}$$

$$= \mathbb{E}[\text{# of visits in } j | X_0 = j] \cdot f_{i,j} = S_{j,j} \cdot f_{i,j}$$
Ever Hitting Probabilities Revisited

Using the fundamental matrix $S$, we can compute ever-hitting probability

$$f_{i,j} = \mathbb{P}(j \text{ ever visited } | \text{ after leaving } i) < 1.$$ 

- For $i, j \in \mathcal{T}, i \neq j$

  $$S_{i,j} = \mathbb{E}[\text{# of visits in } j | X_0 = i]$$
  $$= \mathbb{E}[\text{# of visits in } j | j \text{ is ever visited, } X_0 = i] \cdot f_{i,j}$$
  $$= \mathbb{E}[\text{# of visits in } j | X_0 = j] \cdot f_{i,j} = S_{j,j} \cdot f_{i,j}$$

- For $i = j \in \mathcal{T}$,

  $$S_{j,j} = \mathbb{E}[\text{# of visits in } j | X_0 = j]$$
  $$= \mathbb{E}[\text{# in } j | j \text{ ever again, } X_0 = j] f_{j,j} + \mathbb{E}[\text{# in } j | j \text{ never again, } X_0 = j](1 - f_{j,j})$$
  $$= [1 + S_{j,j}] \cdot f_{j,j} + 1 \cdot (1 - f_{j,j})$$
  $$= 1 + S_{j,j} \cdot f_{j,j}$$
Ever Hitting Probabilities Revisited

Using the fundamental matrix $S$, we can compute ever-hitting probability

$$f_{i,j} = \mathbb{P}(j \text{ ever visited} \mid \text{after leaving } i) < 1.$$ 

- For $i, j \in \mathcal{T}, i \neq j$

  $$S_{i,j} = \mathbb{E}[\# \text{ of visits in } j | X_0 = i] = \mathbb{E}[\# \text{ of visits in } j \mid j \text{ is ever visited, } X_0 = i] \cdot f_{i,j} = \mathbb{E}[\# \text{ of visits in } j \mid X_0 = j] \cdot f_{i,j} = S_{j,j} \cdot f_{i,j}$$

- For $i = j \in \mathcal{T}$,

  $$S_{j,j} = \mathbb{E}[\# \text{ of visits in } j | X_0 = j] = \mathbb{E}[\# \text{ in } j \mid j \text{ ever again, } X_0 = j] f_{j,j} + \mathbb{E}[\# \text{ in } j \mid j \text{ never again, } X_0 = j](1 - f_{j,j}) = [1 + S_{j,j}] \cdot f_{j,j} + 1 \cdot (1 - f_{j,j}) = 1 + S_{j,j} \cdot f_{j,j}$$

$$f_{i,j} = \begin{cases} \frac{S_{i,j}}{S_{j,j}}, & \text{if } i \neq j; \\ \frac{S_{j,j}-1}{S_{j,j}}, & \text{if } i = j. \end{cases}$$
An Example: Credit Ratings Migration

- Financial securities (e.g., bonds) move from one rating category to another.
- Ratings transition matrix (from rating agency’s historical data): probability of a bond moving from one rating to another during a certain period of time (e.g., 1 year).
- Below is a table giving a rating transition matrix produced from historical data by Standard and Poor’s (S&P).

<table>
<thead>
<tr>
<th>Initial ratings</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.9366</td>
<td>0.0583</td>
<td>0.0040</td>
<td>0.0009</td>
<td>0.0002</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>0.0066</td>
<td>0.9172</td>
<td>0.0694</td>
<td>0.0049</td>
<td>0.0006</td>
<td>0.0009</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
<tr>
<td>A</td>
<td>0.0007</td>
<td>0.0225</td>
<td>0.9176</td>
<td>0.0518</td>
<td>0.0049</td>
<td>0.0020</td>
<td>0.0001</td>
<td>0.0004</td>
</tr>
<tr>
<td>BBB</td>
<td>0.0003</td>
<td>0.0026</td>
<td>0.0483</td>
<td>0.8924</td>
<td>0.0444</td>
<td>0.0081</td>
<td>0.0016</td>
<td>0.0023</td>
</tr>
<tr>
<td>BB</td>
<td>0.0003</td>
<td>0.0006</td>
<td>0.0044</td>
<td>0.0666</td>
<td>0.8323</td>
<td>0.0746</td>
<td>0.0105</td>
<td>0.0107</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0.0010</td>
<td>0.0032</td>
<td>0.0046</td>
<td>0.0572</td>
<td>0.8362</td>
<td>0.0384</td>
<td>0.0594</td>
</tr>
<tr>
<td>CCC</td>
<td>0.0015</td>
<td>0</td>
<td>0.0029</td>
<td>0.0088</td>
<td>0.0191</td>
<td>0.1028</td>
<td>0.6123</td>
<td>0.2526</td>
</tr>
<tr>
<td>Default</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

*Source:* Standard & Poor’s, January 2001
DTMC

**Infinite state space**
Example: random walk
\( S = \{ ..., -2, -1, 0, 1, 2, ... \} \)

**Finite state space**
Examples: Markov mouse \( S = \{ 1, 2, ..., 9 \} \),
Gambler’s ruin \( S = \{ 0, 1, 2, ..., N-1, N \} \)

**Reducible** (multiple classes)
Example: Gambler’s ruin

**Irreducible** (one class)
Examples: Markov mouse, Pooh bear on honey trees
Note: we say the DTMC is “GOOD”! Stationary distribution exists, i.e., the balance equation \( \pi P = \pi \) and \( \sum_k \pi_k = 1 \) has a unique solution. In addition, \( \pi_k \) is the long-run proportion of time the DTMC is in \( k \).

**All classes are recurrent**
Focus on each class and decompose the system into isolated subsystems

**Some class is transient**
Put the transition probability matrix (TPM) into its canonical form, which can be used to compute quantities such as: mean number until absorption and probability of being absorbed by a recurrent class, starting with a transient state, etc.

**Aperiodic**
Examples: Pooh bear
Note: we say the DTMC is “GREAT”! Limiting steady-state distribution exists, i.e., \( P_{ij}^{(n)} = P_{ij}^n \xrightarrow{n} \pi_j \)

**Periodic**
Examples: Markov mouse with period \( d = 2 \)
Limiting steady-state distribution does NOT exist, i.e., \( P_{ij}^{(dn)} = P_{ij}^{dn} \xrightarrow{n} \pi_j d \)
Time Reversibility
Consider a DTMC \( \{X_n\} \) in stationary with TPM \( P \) and steady state \( \pi \). Define its reverse-time process

\[
\{\leftarrow X_n\} = \{X_n, X_{n-1}, X_{n-2}, \ldots\}
\]

Properties:

(i) Inverse-time process \( \{\leftarrow X_n\} \) is a DTMC, that is,

\[
P(\leftarrow X_m = j | X_{m+1} = i, X_{m+2} = i_m, X_{m+3} = i_{m+3}, \ldots) = P(\leftarrow X_m = j | X_{m+1} = i)
\]

(ii) The transition probability of \( \{\leftarrow X_n\} \) is

\[
\leftarrow P_{i,j} = \pi_j P_{j,i} \pi_i.
\]
Reverse Time DTMC

Definition (Reverse-Time DTMC)
Consider a DTMC \( \{ X_n \} \) in stationary with TPM \( P \) and steady state \( \pi \). Define its reverse-time process
\[
\{ \leftarrow X_n \} = \{ X_n, X_{n-1}, X_{n-2}, \ldots \}
\]

Properties:

(i) Inverse-time process \( \{ \leftarrow X_n \} \) is a DTMC, that is,
\[
\mathbb{P}(X_m = j | X_{m+1} = i, X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots) = \mathbb{P}(X_m = j | X_{m+1} = i)
\]

(ii) The transition probability of \( \{ \leftarrow X_n \} \) is
\[
\leftarrow P_{i,j} \equiv \mathbb{P}(X_m = j | X_{m+1} = i) = \frac{\pi_j P_{j,i}}{\pi_i}.
\]
Reverse Time DTMC

Proof of property (i):

- Given present, future \( \perp \) past (and past \( \perp \) future).
Reverse Time DTMC

Proof of property (i):
- Given present, future ⊥ past (and past ⊥ future).
- More rigorously:

\[
LHS = \frac{\mathbb{P}(X_m = j, X_{m+1} = i, X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots)}{\mathbb{P}(X_{m+1} = i, X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots)}
= \frac{\mathbb{P}(X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots | X_{m+1} = i, X_m = j) \mathbb{P}(X_{m+1} = i, X_m = j)}{\mathbb{P}(X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots | X_{m+1} = i) \mathbb{P}(X_{m+1} = i)}
= \frac{\mathbb{P}(X_{m+1} = i, X_m = j)}{\mathbb{P}(X_{m+1} = i)} = \mathbb{P}(X_m = j | X_{m+1} = i)
\]
Reverse Time DTMC

Proof of property (i):

- Given present, future \( \perp \) past (and past \( \perp \) future).
- More rigorously:

\[
LHS = \frac{\mathbb{P}(X_m = j, X_{m+1} = i, X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots)}{\mathbb{P}(X_{m+1} = i, X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots)}
\]

\[
= \frac{\mathbb{P}(X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots | X_{m+1} = i, X_m = j) \mathbb{P}(X_{m+1} = i, X_m = j)}{\mathbb{P}(X_{m+2} = i_{m+2}, X_{m+3} = i_{m+3}, \ldots | X_{m+1} = i) \mathbb{P}(X_{m+1} = i)}
\]

\[
= \frac{\mathbb{P}(X_{m+1} = i, X_m = j)}{\mathbb{P}(X_{m+1} = i)} = \mathbb{P}(X_m = j | X_{m+1} = i)
\]

Proof of property (ii):

\[
P_{i,j} = \mathbb{P}(X_m = j | X_{m+1} = i) = \frac{\mathbb{P}(X_m = j, X_{m+1} = i)}{\mathbb{P}(X_{m+1} = i)}
\]

\[
= \frac{\mathbb{P}(X_{m+1} = i | X_m = j) \mathbb{P}(X_m = j)}{\mathbb{P}(X_{m+1} = i)} = \frac{\pi_j P_{j,i}}{\pi_i}.
\]
Time Reversibility

Definition (Time Reversibility)

We say DTMC \( \{X_n\} \) is time reversible (T-R), if \( \{\bar{X}_n\} \) and \( \{X_n\} \) have the same probability law

\[ P = \bar{P}, \quad \left( P_{i,j} = \bar{P}_{i,j} = \frac{P_{j,i} \pi_j}{\pi_i} \right), \]

or equivalently,

\[ \pi_i P_{i,j} = \pi_j P_{j,i} \quad \text{for all } i, j \in S. \]
Time Reversibility

Definition (Time Reversibility)

We say DTMC \( \{X_n\} \) is time reversible (T-R), if \( \{X_n\} \) and \( \{X_n\} \) have the same probability law

\[
P = \overleftarrow{P}, \quad \left( P_{i,j} = \overleftarrow{P}_{i,j} = \frac{P_{j,i} \pi_j}{\pi_i} \right),
\]

or equivalently,

\[
\pi_i P_{i,j} = \pi_j P_{j,i} \quad \text{for all } i, j \in S.
\]

Interpretation:

A DTMC is T-R if, for all \( i, j \in S \),

long-run rate (prop. of transitions) from \( i \to j = \) long-run rate from \( j \to i \)

Example 1

In general, a DTMC may not be T-R, that is, \( \{X_n\} \) and \( \{\hat{X}_n\} \) may have different dynamics (\( P_{i,j} \neq \hat{P}_{i,j} \)).

Example: DTMC with \( S = \{1, 2, 3\} \)

\[
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{bmatrix}, \quad \pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)
\]

Checking the T-R criterion:

\[
\frac{1}{3} \cdot 1 = \pi_1 P_{1,2} \neq \pi_2 P_{2,1} = \frac{1}{3} \cdot 0
\]

Hence, the DTMC is NOT time reversible.
Example 2: Random Walk on a Bounded Set

Consider a random walk in a bounded set \( S = \{0, 1, \ldots, N\} \) with transition probabilities \( P_{i,i+1} = p_i = 1 - P_{i,i-1} \) for \( 1 \leq i \leq N - 1 \), \( P_{N,N-1} = P_{0,1} = 1 \). Let \( S_n \) denote the position at time \( n \), then \( \{S_n, n \geq 0\} \) is a DTMC.

Question:

- Is the DTMC time-reversible?
Time Reversibility

Proposition (A Sufficient Condition)

If there exists a vector $\pi^* = (\pi_1^*, \pi_2^*, \ldots)$, s.t., $\pi_k^* > 0$, $\sum_{k \in S} \pi_k^* = 1$, satisfying

$$\pi_i^* P_{i,j} = \pi_j^* P_{j,i}, \quad \text{for all } i, j \in S$$

then

1. $\pi^* = \pi$ (that is, $\pi^*$ is the stationary distribution)
2. The DTMC is T-R.
Time Reversibility

Proposition (A Sufficient Condition)

If there exists a vector \( \pi^* = (\pi_1^*, \pi_2^*, \ldots) \), s.t., \( \pi_k^* > 0 \), \( \sum_{k \in S} \pi_k^* = 1 \), satisfying

\[
\pi_i^* P_{i,j} = \pi_j^* P_{j,i}, \quad \text{for all } i, j \in S
\]  

(1)

then

1. \( \pi^* = \pi \) (that is, \( \pi^* \) is the stationary distribution)
2. The DTMC is T-R.

Proof: summing both sides of (1) over \( i \) yields

\[
\sum_i \pi_i^* P_{i,j} = \sum_i \pi_j^* P_{j,i} = \pi_j^* \sum_i P_{j,i} = \pi_j^*,
\]

which is the balance equation! Hence, \( \pi = \pi^* \).

In addition, we have

\[
\pi_i P_{i,j} = \pi_j P_{j,i},
\]

so that DTMC is T-R.
Random Walks on Weighted Graphs

**Definition (RWoWG)**

A graph has \( n \) nodes and \( m \) arcs. Each arc (undirected) \((i, j)\) has a weight \( w_{i,j} = w_{j,i} \). A particle moves from node \( i \rightarrow j \) with a probability that is proportional to the weight \( w_{i,j} \), that is,

\[
P_{i,j} = \frac{w_{i,j}}{\sum_k w_{i,k}} = \frac{\text{weight on arc } (i, j)}{\text{sum of weights on all outgoing arcs from } i},
\]

for all \( i, j \).

Let \( X_n \) be the location of the particle at time \( n \). If the RWoWG is connected, then \( \{X_n, n \geq 0\} \) is a DTMC which is irreducible and positive-recurrence.
Random Walks on Weighted Graphs

**Definition (RWoWG)**

A graph has \( n \) nodes and \( m \) arcs. Each arc (undirected) \((i, j)\) has a weight \( w_{i,j} = w_{j,i} \). A particle moves from node \( i \rightarrow j \) with a probability that is proportional to the weight \( w_{i,j} \), that is,

\[
P_{i,j} = \frac{w_{i,j}}{\sum_k w_{i,k}} = \frac{\text{weight on arc } (i, j)}{\text{sum of weights on all outgoing arcs from } i}
\]

for all \( i, j \).

Let \( X_n \) be the location of the particle at time \( n \).

If the RWoWG is connected, then \( \{X_n, n \geq 0\} \) is a DTMC which is irreducible and positive-recurrence.
Theorem (Steady state of RWoWG)

The DTMC of a connected RWoWG

1. is time-reversible,
2. has stationary distribution

\[
\pi_i = \frac{\sum_{k \in S} w_{i,k}}{\sum_{k \in S} \sum_{\ell \in S} w_{k,\ell}}, \quad i \in S.
\]
Random Walks on Weighted Graphs

Theorem (Steady state of RWoWG)

The DTMC of a connected RWoWG

1. is time-reversible,
2. has stationary distribution

\[ \pi_i = \frac{\sum_{k \in S} w_{i,k}}{\sum_{k \in S} \sum_{\ell \in S} w_{k,\ell}}, \quad i \in S. \] (2)

Intuition:

\[ \pi_i = \frac{\text{sum of weights on all incoming arcs to } i}{\sum_{k \in S} \left( \text{sum of weights on all incoming arcs to } k \right)} \]
Theorem (Steady state of RWoWG)

The DTMC of a connected RWoWG

1. is time-reversible,
2. has stationary distribution

\[ \pi_i = \frac{\sum_{k \in S} w_{i,k}}{\sum_{k \in S} \sum_{\ell \in S} w_{k,\ell}}, \quad i \in S. \]  

(2)

Intuition:

\[ \pi_i = \frac{\text{sum of weights on all incoming arcs to } i}{\sum_{k \in S} (\text{sum of weights on all incoming arcs to } k)} \]

Proof: We verify that \( \pi_i \) in (2) indeed solves Equation (1):

\[ \pi_i P_{i,j} = \frac{\sum_{k \in S} w_{i,k}}{\sum_{\ell \in S} \sum_{k \in S} w_{\ell,k}} \cdot \frac{w_{i,j}}{\sum_{k \in S} w_{i,k}} = \frac{\sum_{k \in S} w_{j,k}}{\sum_{\ell \in S} \sum_{k \in S} w_{\ell,k}} \cdot \frac{w_{j,i}}{\sum_{k \in S} w_{j,k}} = \pi_j P_{j,i}, \]

where the 2nd equality holds because \( w_{i,j} = w_{j,i} \).
Examples of RWoWG: Markov Mouse in a Closed Maze Revisited

Questions:
- Is the DTMC time reversible?
- Are there easy ways to compute $\pi$?
Examples of RWoWG: Random Walk on a Bounded Set Revisited

Consider a random walk in a bounded set \( S = \{0, 1, \ldots, N\} \) with transition probabilities \( P_{i,i+1} = p_i = 1 - P_{i,i-1} \) for \( 1 \leq i \leq N - 1 \), \( P_{N,N-1} = P_{0,1} = 1 \). Let \( S_n \) denote the position at time \( n \), then \( \{S_n, n \geq 0\} \) is a DTMC.

Questions:

- Can we transform this problem to a RWoWG?
- If yes, give an expression of \( \pi_i, i = 0, 1, \ldots, N \).
Extensions of DTMC

- Continuous-time Markov chain (will be covered later in Topic 4);
- Phase-type distributions and matrix geometrics;
- Hidden Markov chain (HMC, with both observable states and hidden states);
- Markov decision process (MDP, with “actions”);
  MDP is an important building block for reinforcement learning.
- Partially observable MDP (POMDP, with both hidden states and actions);
- Markov chain Monte Carlo (MCMC, a computer simulation algorithm).