STAT253/317 Lecture 5 Time Reversibility (§4.8)

4.8.1 Backward Markov Chain

If $\{\ldots, X_{n-1}, X_n, X_{n+1}, \ldots\}$ is a Markov chain, the backward chain $\{\ldots, X_{n+1}, X_n, X_{n-1}, \ldots\}$ is also a Markov chain.

Proof:

$$\mathrm{P}(X_m = j \mid X_{m+1} = i, X_{m+2}, X_{m+3}, \ldots)$$

$$= \frac{\mathrm{P}(X_m = j, X_{m+1} = i, X_{m+2}, X_{m+3}, \ldots)}{\mathrm{P}(X_{m+1} = i, X_{m+2}, X_{m+3}, \ldots)}$$

$$= \frac{\mathrm{P}(X_{m+2}, X_{m+3}, \ldots \mid X_m = j, X_{m+1} = i)\mathrm{P}(X_m = j, X_{m+1} = i)}{\mathrm{P}(X_{m+2}, X_{m+3}, \ldots \mid X_{m+1} = i)\mathrm{P}(X_{m+1} = i)}$$

$$= \frac{\mathrm{P}(X_{m+2}, X_{m+3}, \ldots \mid X_{m+1} = i)\mathrm{P}(X_m = j, X_{m+1} = i)}{\mathrm{P}(X_{m+2}, X_{m+3}, \ldots \mid X_{m+1} = i)\mathrm{P}(X_{m+1} = i)} \quad \text{(Markov Prope}$$

$$= \frac{\mathrm{P}(X_m = j, X_{m+1} = i)}{\mathrm{P}(X_{m+1} = i)} = \mathrm{P}(X_m = j \mid X_{m+1} = i)$$

Transition Probabilities of the Backward Markov Chain

Consider a Markov chain $\{X_n : n = 0, 1, 2, \ldots\}$ with transition probabilities $\{P_{ij}\}$.

Let $\{\pi_j^{(m)} = P(X_m = j)\}_{j\geq 0}$ be the marginal distribution of X_m.

The transition probabilities $\{Q_{ij}^{(m)}\}$ of the backward Markov chain are

$$\begin{aligned} Q_{ij}^{(m)} &= \mathrm{P}(X_m = j \mid X_{m+1} = i) \\ &= \frac{\mathrm{P}(X_m = j, X_{m+1} = i)}{\mathrm{P}(X_{m+1} = i)} \\ &= \frac{\mathrm{P}(X_m = j)\mathrm{P}(X_{m+1} = i \mid X_m = j)}{\mathrm{P}(X_{m+1} = i)} = \frac{\pi_j^{(m)} P_{ji}}{\pi_i^{(m+1)}} \end{aligned}$$

We can see the backward Markov chain is **NOT stationary** because the transition probabilities $Q_{ij}^{(m)}$ depend on m.

How to make a backward Markov chain stationary?

To make the backward Markov chain **stationary**, the forward chain must start with its stationary distribution $\{\pi_j\}$ so that

$$P(X_m = j) = \pi_j \quad \text{for all } m$$

the transition probabilities $\{Q_{ij}\}$ of the backward Markov chain is

$$Q_{ij} = \frac{\pi_j P_{ji}}{\pi_i}$$

which does not depend on m

Time Reversible Markov Chains & Detailed Balanced Equations

A Markov chain is said to be *time reversible* iff

$$Q_{ij} = P_{ij},$$

i.e., **it behaves exactly the same** no matter **running forward or backward** when in the stationary state.

Because Q_{ij} equals $\pi_j P_{ji}/\pi_i$, a Markov chain is time reversible if and only if its stationary distribution $\{\pi_j\}$ satisfies the equations

$$\pi_i P_{ij} = \pi_j P_{ji} \quad \text{for all } i, j.$$

This set of equations is called the **detailed balanced equation**.

Balanced Equations v.s. Detailed Balanced Equations

Recall a distribution π_j for a Markov chain is said to be stationary if and only if it satisfies

$$\pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij} \quad \text{for all } j \in \mathfrak{X}.$$

This set of equations is called the **balanced equations**.

Balanced Equations v.s. Detailed Balanced Equations

Recall a distribution π_j for a Markov chain is said to be stationary if and only if it satisfies

$$\pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij} \quad \text{for all } j \in \mathfrak{X}.$$

This set of equations is called the **balanced equations**.

A solution to the **detailed balanced equations** must also be a solution to the **balanced equations**, because

$$\sum\nolimits_{i \in \mathfrak{X}} \pi_i P_{ij} = \sum\nolimits_{i \in \mathfrak{X}} \pi_j P_{ji} = \pi_j \sum\nolimits_{i \in \mathfrak{X}} P_{ji} = \pi_j \cdot 1 = \pi_j$$

Balanced Equations v.s. Detailed Balanced Equations

Recall a distribution π_j for a Markov chain is said to be stationary if and only if it satisfies

$$\pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij} \quad \text{for all } j \in \mathfrak{X}.$$

This set of equations is called the **balanced equations**.

A solution to the **detailed balanced equations** must also be a solution to the **balanced equations**, because

$$\sum\nolimits_{i \in \mathfrak{X}} \pi_i P_{ij} = \sum\nolimits_{i \in \mathfrak{X}} \pi_j P_{ji} = \pi_j \sum\nolimits_{i \in \mathfrak{X}} P_{ji} = \pi_j \cdot 1 = \pi_j$$

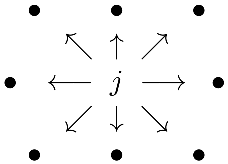
It is possible that the balanced equations have solutions but the detailed balanced equations do not.

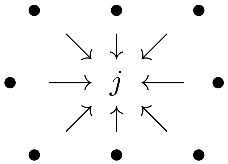
Interpretation of the Balanced Equation

$$\pi_j = \sum_{i \in \mathfrak{X}} \pi_i P_{ij} \quad \text{for all } j \in \mathfrak{X}$$

$$\Leftrightarrow \quad \pi_j(1 - P_{jj}) = \sum_{i \in \mathfrak{X}, i \neq j} \pi_i P_{ij} \quad \text{for all } j \in \mathfrak{X}$$

rate of transitions **out of** state j = rate of transitions **into** state j





Interpretation of the Detailed Balanced Equation

$$\pi_i P_{ij} = \pi_j P_{ji}$$

rate of transitions from i to j = rate of transitions from j to i

$i \longrightarrow j$ $\qquad$ $j \longrightarrow i$

Balanced Eqns v.s. Detailed Balanced Eqns.

- For balanced equations,

  $$\text{the \# of equations} = \text{\# of states} = \text{\# of unknowns}$$

- For detailed balanced equations,

  $$\text{\# of equations} = \text{\# of pairs of states} > \text{\# of unknowns}$$

- Detailed Balanced Equations are easier to solve than the Balanced Equations as the former ones involve only two unknowns in each equation
- One can start by solving the detailed balanced equations for the stationary distribution. If you can find one, it'll also be the solution for the balanced equations. That also proves the Markov chain if positive recurrent if it's irreducible (2nd limit theorem).
- However, if the detailed balanced equations have no solutions, it doesn't prove the Markov chain to be null current or transient since the balanced equations might still have a solution.

Example 4.35

Consider a random walk with states $0, 1, \ldots, M$ and transition probabilities

$$
\begin{aligned}
P_{i,i+1} &= \alpha_i = 1 - P_{i,i-1}, \quad \text{for } i = 1, \ldots, M-1, \\
P_{0,1} &= \alpha_0 = 1 - P_{0,0}, \\
P_{M,M} &= \alpha_M = 1 - P_{M,M-1}
\end{aligned}
$$

$$
0 \xrightarrow{\alpha_0} 1 \xrightarrow{\alpha_1} 2 \cdots \longrightarrow M-1 \xrightarrow{\alpha_{M-1}} \underset{\alpha_M}{\underset{\circlearrowleft}{M}}
$$

$$
\underset{1-\alpha_0}{\underset{\circlearrowleft}{0}} \xleftarrow{1-\alpha_1} 1 \xleftarrow{1-\alpha_2} 2 \cdots \longleftarrow M-1 \xleftarrow{1-\alpha_M} M
$$

Example 4.35 (Cont'd)

The stationary distribution π can be solved via the detailed balanced equation

$$\pi_i P_{i,i-1} = \pi_i(1-\alpha_i) = \pi_{i-1}P_{i-1,i} = \pi_{i-1}\alpha_{i-1}$$

So

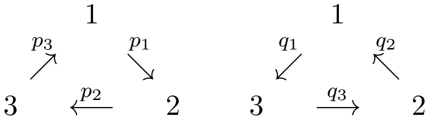
$$\pi_i = \frac{\alpha_{i-1}}{1-\alpha_i}\pi_{i-1} = \ldots = \frac{\alpha_{i-1}\alpha_{i-2}\ldots\alpha_0}{(1-\alpha_i)(1-\alpha_{i-1})\ldots(1-\alpha_1)}\pi_0$$

Since $\sum_0^M \pi_i = 1$, one can solve π_0 via

$$\pi_0\left[1+\sum_{i=1}^M \frac{\alpha_{i-1}\alpha_{i-2}\ldots\alpha_0}{(1-\alpha_i)(1-\alpha_{i-1})\ldots(1-\alpha_1)}\right] = 1$$

A Non-Time-Reversible Markov Chain

In Exercise 4.34 (a flea moving around the vertices of a triangle),



where $p_i + q_i = 1$

the transition probabilities, and the stationary distribution are respectively

$$\begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & p_1 & q_1 \\ q_2 & 0 & p_2 \\ p_3 & q_3 & 0 \end{pmatrix} \end{array}, \quad \text{and } \pi = \left(\frac{1 - p_2 q_3}{C}, \frac{1 - p_3 q_1}{C}, \frac{1 - p_1 q_2}{C}\right)$$

where $C = 3 - p_2 q_3 - p_3 q_1 - p_1 q_2$. One can easily verify that

$$\pi_1 P_{12} = \pi_1 p_1 \neq \pi_2 P_{21} = \pi_2 q_2$$

The chain is NOT time reversible.

Other Non-Time-Reversible Markov Chains

- A Markov chain with **transient states** cannot be time-reversible because then running forward and backward in time will not be equivalent.

- If there exists two states i and j such that

$$P_{ij} > 0 \quad \text{but} \quad P_{ji} = 0$$

  then the Markov chain cannot be time-reversible because then when running backward in time

$$Q_{ij} = \frac{\pi_j P_{ji}}{\pi_i} = 0 \neq P_{ij}.$$

Theorem 4.2

An ergodic Markov chain for which $P_{ij} = 0$ whenever $P_{ji} = 0$ is time reversible if and only if starting in state i, any path back to i has the same probability as the reversed path. That is, if

$$P_{ii_1} P_{i_1 i_2} \dots P_{i_k i} = P_{ii_k} P_{i_k i_{k-1}} \dots P_{i_1 i}$$

for all states $i, i_1, \dots, i_k$.

$$\begin{array}{ccccc} & i_k & \longrightarrow & i & \\ \nearrow & & & & \searrow \\ i_{k-1} & & & & i_1 \\ \nwarrow & & & & \swarrow \\ & \cdots & \longleftarrow & i_2 & \end{array} \qquad \begin{array}{ccccc} & i_k & \longleftarrow & i & \\ \swarrow & & & & \nwarrow \\ i_{k-1} & & & & i_1 \\ \searrow & & & & \nearrow \\ & \cdots & \longrightarrow & i_2 & \end{array}$$

Theorem 4.2 — Proof of Necessity

If a Markov chain is time reversible, we have

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \pi_k P_{kj} = \pi_j P_{jk}.$$

implying (if $P_{ij}P_{jk} > 0$) that

$$\frac{\pi_i}{\pi_k} = \frac{P_{ji}P_{kj}}{P_{ij}P_{jk}},$$

but $\pi_i P_{ik} = \pi_k P_{ki}$ also implies $\pi_i/\pi_k = P_{ki}/P_{ik}$. Thus

$$P_{ik}P_{kj}P_{ji} = P_{ij}P_{jk}P_{ki}.$$

This proves for the case $i \to j \to k \to i$. The general case for longer cycle can be proved similarly

Theorem 4.2 — Proof of Sufficiency

Consider the cycle $i \to i_1 \to i_2 \to \ldots \to i_k \to j \to i$.

$$P_{ii_1} P_{i_1 i_2} \ldots P_{i_k j} P_{ji} = P_{ij} P_{ji_k} P_{i_k i_{k-1}} \ldots P_{i_1 i}$$

Summing the preceding over all states $i_1, \ldots, i_k$ yields

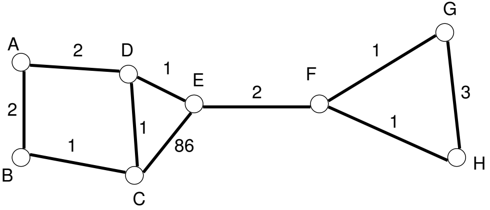
$$P_{ij}^{(k)} P_{ji} = P_{ij} P_{ji}^{(k)}$$

Letting $k \to \infty$ yields

$$\underbrace{\lim_{k\to\infty} P_{ij}^{(k)}}_{=\pi_j} P_{ji} = P_{ij} \underbrace{\lim_{k\to\infty} P_{ji}^{(k)}}_{=\pi_i}$$

in which $\lim_{k\to\infty} P_{ij}^{(k)} = \pi_j$ for all j since the Markov chain is ergodic.
This proves the theorem.

Example 4.36 Random Walk on a Weighted Graph (p.241)



A graph = a set of vertices (or nodes) + a set of arcs (or edges) connecting some pairs of vertices. We consider random walk on a connected graph such that

- each pair (i, j) of vertices are connected by at most one arc;
- all arcs are undirected: arc (i, j) = arc (j, i);
- there is a path consists of arcs connecting any pair of vertices;
- each arc (i, j) is associated with a weight $w_{ij} > 0$
  - $w_{ij} = 0$ if there is not arc connecting (i, j)
  - $w_{ij} = w_{ji}$

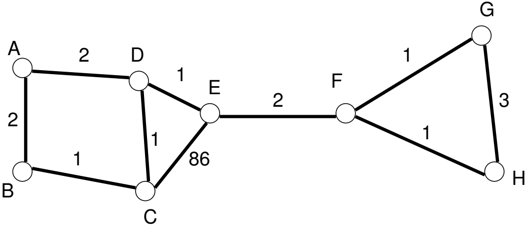
Example 4.36 Random Walk on a Weighted Graph (p.241)

A particle moving from vertices to vertices that if at any time the particle is at node i, then it will next move to node j with probability

$$P_{ij} = \frac{w_{ij}}{\sum_k w_{ik}},$$

E.g., in the graph below, there are two arcs from vertices B with weights $w_{BA} = 2$ and $W_{BC} = 1$ respectively. So,

$$P_{BA} = \frac{w_{BA}}{w_{BA} + w_{BC}} = \frac{2}{2+1} = \frac{2}{3}, \quad P_{BC} = \frac{w_{BC}}{w_{BA} + w_{BC}} = \frac{1}{2+1} = \frac{1}{3}.$$



Random walk on a graph is **irreducible** because the graph is **connected**.

Random Walk on a Weighted Graph is Time Reversible

Solving the detailed balanced equation:

$$\pi_i P_{ij} = \frac{\pi_i w_{ij}}{\sum_k w_{ik}} = \frac{\pi_j w_{ji}}{\sum_k w_{jk}} = \pi_j P_{ji} \quad \text{for all } i, j$$

or, equivalently, since $w_{ij} = w_{ji}$,

$$\frac{\pi_i}{\sum_k w_{ik}} = \frac{\pi_j}{\sum_k w_{jk}} \quad \text{for all } i, j,$$

which means $\frac{\pi_i}{\sum_k w_{ik}}$ is a constant c for all i, i.e.,

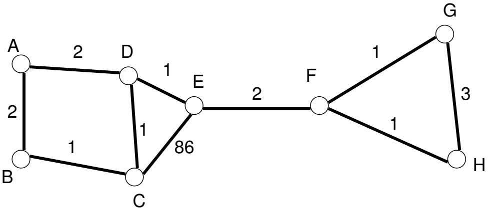
$$\pi_i = c \sum_k w_{ik}.$$

Since $1 = \sum_j \pi_j = c \sum_j \sum_k w_{jk}$, we know $c = 1/(\sum_j \sum_k w_{jk})$, and hence

$$\pi_i = \frac{\sum_k w_{ik}}{\sum_j \sum_k w_{jk}}$$

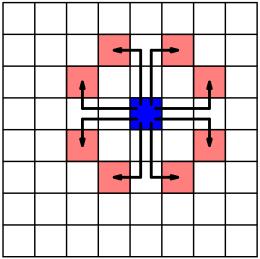
is a solution to the detailed balanced eq. The process is therefore time-reversible.

Random Walk on a Weighted Graph

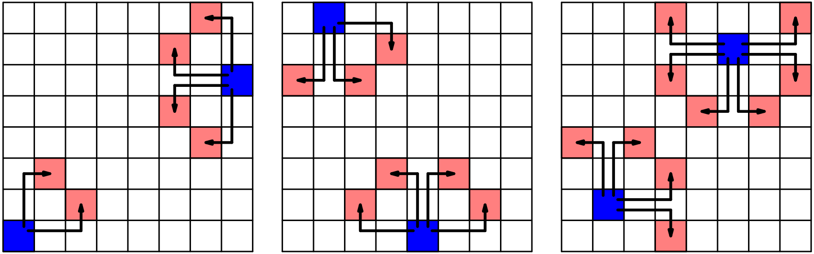


Vertices i	$\sum_k w_{ik}$	π_i
A	$2+2=4$	$\pi_A = 4/200$
B	$2+1=3$	$\pi_B = 3/200$
C	$1+1+86=88$	$\pi_C = 88/200$
D	$2+1+1=4$	$\pi_D = 4/200$
E	$1+2+86=89$	$\pi_E = 89/200$
F	$2+1+1=4$	$\pi_F = 4/200$
G	$1+3=4$	$\pi_G = 4/200$
H	$1+3=4$	$\pi_H = 4/200$
Sum	$\sum_i \sum_k w_{ik} = 200$	

Random Knight on a Chessboard

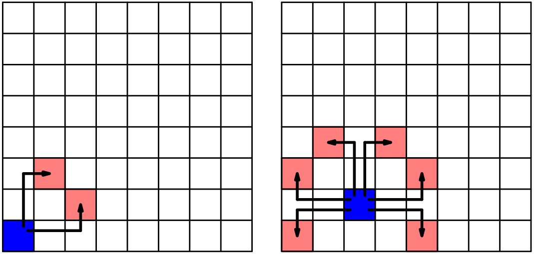


- The Knight moves in an **L** shape in any direction.
- At the blue square, the Knight can move to any of the 8 red squares.
- From a square near the boundary, the Knight has fewer possible moves as it cannot move out of the Chessboard (see the 3 graphs below.)



Random Knight on a Chessboard

- A Knight moves randomly on an empty chessboard.
- In each step, it's equally like to take any of its legal moves. E.g., at the corner, it has prob. $1/2$ each to move to either of the two red squares, from which it has prob. $1/6$ each to move to any of the 6 possible squares.



- Each move is indep. of the history of moves up to that time.
- The position of on knight after nth move is a Markov chain where states are the 64 squares on the chessboard.

Random Knight on a Chessboard

The Knight's random walk on a Chessboard is also a random walk on weighted graph where

- the vertices are the 64 squares on the chessboard;
- there is an arc between any two squares that Knight can move in 1 step;
- all the arcs have weight $w_{ij} = 1$.

The transition probability of a random walk on weighted graph from square i to square j is

$$P_{ij} = \frac{w_{ij}}{\sum_k w_{ik}} = \frac{1}{\text{\# of squares that connected with square } i \text{ with an arc}}$$
$$= \frac{1}{\text{\# of legal moves from square } i}$$

which is exactly the random walk of the knight.

Random Knight on a Chessboard

Using the property of random walks on a graph, the stationary distribution of the Knight's random walk is

$$\pi_i = \frac{\sum_k w_{ik}}{\sum_j \sum_k w_{jk}} = \frac{\text{\# of legal moves from square } i}{\sum_j (\text{\# of legal moves from square } j)}$$

The numbers of legal moves from the squares are as follows:

2	3	4	4	4	4	3	2
3	4	6	6	6	6	4	3
4	6	8	8	8	8	6	4
4	6	8	8	8	8	6	4
4	6	8	8	8	8	6	4
4	6	8	8	8	8	6	4
3	4	6	6	6	6	4	3
2	3	4	4	4	4	3	2

The sum of the number of possible moves over all squares is

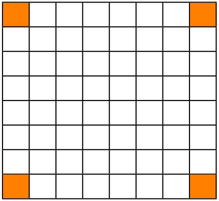
$$2 \times 4 + 3 \times 8 + 4 \times 20 + 6 \times 16 + 8 \times 16 = 336.$$

The long run proportion of time that the Knight is in a specific square is simply the counts in the table above divided by 336.

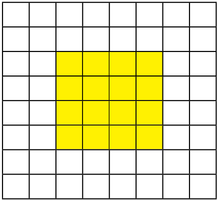
Return Time of a Random Knight

Recall that $1/\pi_i = \mathbb{E}[T_i]$ is the expected time between two visits of the Markov chain to state i.

Starting from one of the four corners, it takes $1/\pi_i = 336/2 = 168$ moves on average for a Knight to return to its initial position.



Starting from the center of the chessboard, it takes $1/\pi_i = 336/8 = 42$ moves on average for a Knight to return to its initial position



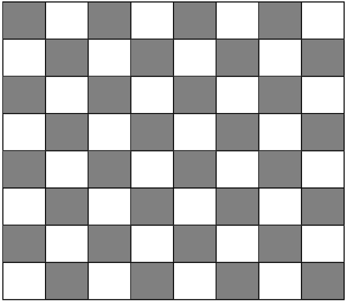
More Questions

- Is this Markov chain irreducible? That is, can the Knight visit every square from every square?
- What is the period of this Markov chain?

More Questions

- Is this Markov chain irreducible? That is, can the Knight visit every square from every square?

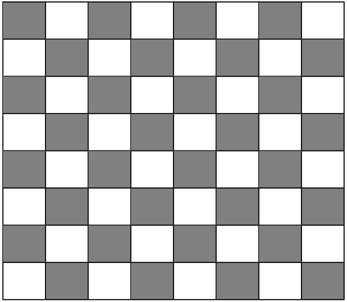
- What is the period of this Markov chain?



More Questions

- Is this Markov chain irreducible? That is, can the Knight visit every square from every square?

- What is the period of this Markov chain?



Every "L" move can only from a gray square to a white square or a white square to a gray square.