STK2130 – Week 8

A. B. Huseby

Department of Mathematics
University of Oslo, Norway

Periodic Markov chains

We recall that a Markov chain $\{X_n\}$ is said to be periodic if it can only return to a state in a multiple of $d > 1$ steps.

EXAMPLE: Assume that $\{X_n\}$ has state space $\mathcal{S} = \{0, 1\}$, and transition matrix:

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

Assuming that $X_0 = 0$, it follows that:

$$X_n = \begin{cases} 0, & \text{if } n \text{ is even} \\ 1, & \text{if } n \text{ is odd} \end{cases}$$

Thus, this chain can return to a state (0 or 1) in a multiple of 2 steps.

QUESTION: Does periodicity only occur when the chain is deterministic?

Periodic Markov chains (cont.)

EXAMPLE 1: Assume that $\{X_n\}$ has state space $\mathcal{S} = \{0, 1, 2\}$, and transition matrix:

$$\boldsymbol{P} = \begin{bmatrix} 0.0 & 1.0 & 0.0 \\ 0.5 & 0.0 & 0.5 \\ 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Assuming that $X_0 = 1$, the chain will return to this state for $n = 2, 4, 6, \ldots$. Thus, the chain is periodic but not deterministic.

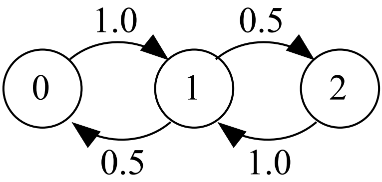


Figure: A non-deterministic periodic Markov chain

Periodic Markov chains (cont.)

EXAMPLE 2. One-dimensional random walk. If $X_0 = 0$, then X_n is even if n is even, and odd if n is odd. The chain can only return to state 0 in an even number of steps. Thus, this chain is periodic but not deterministic.

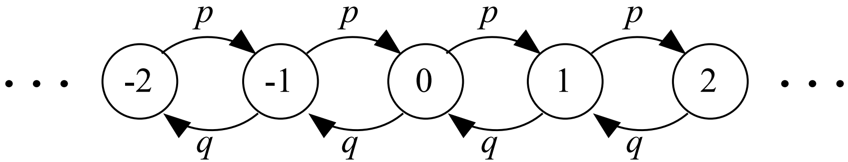


Figure: A one-dimensional random walk

Periodic Markov chains (cont.)

EXAMPLE 3: Assume that $\{X_n\}$ has state space $\mathcal{S} = \{0, 1, 2, 3, 4\}$, and transition matrix:

$$\boldsymbol{P} = \begin{bmatrix} 0.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 0.5 & 0.0 & 0.0 & 0.0 & 0.5 \\ 0.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 & 0.0 \end{bmatrix}$$

Assuming that $X_0 = 2$, the chain will return to this state for $n = 3, 6, 9, \ldots$.
Thus, the chain is periodic but not deterministic.

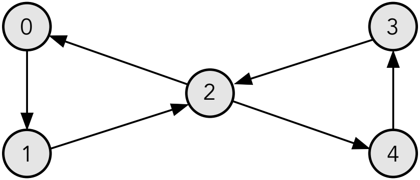


Figure: A non-deterministic periodic Markov chain

Chapter 4.7 Branching Processes

Population with X_0 individuals, each able to produce offspring of the same kind during its lifetime.

$$P_j = P(\text{An individual produces } j \text{ new offspring}), \quad j = 0, 1, 2, \ldots$$

ASSUMPTIONS: $P_0 > 0$ and $P_j < 1$ for $j = 0, 1, 2, \ldots$

$$X_n = \text{Population size in the } n\text{th generation.}, \quad n = 0, 1, 2, \ldots$$

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ P_0 & P_1 & P_2 & P_3 & \cdots \\ P_0^2 & 2P_0 \cdot P_1 & \cdots & \cdots & \cdots \\ P_0^3 & 3P_0^2 \cdot P_1 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

Chapter 4.7 Branching Processes (cont.)

NOTE:

- Since $P_{00} = 1$, then 0 is a recurrent state.
- Since $P_0 > 0$, it follows that $P_{j0} = P_0^j > 0$.

  Hence, state j is transient for all $j > 0$.
- Any finite set of transient states $\{1, 2, \ldots, n\}$ will be visited only a finite number of times.

  Hence, since $P_0 > 0$, the population size converges to 0 or ∞ with probability 1.

Chapter 4.7 Branching Processes (cont.)

Assume that $X_0 = 1$, and let μ and σ^2 denote respectively the mean and the variance of the number of offspring of an individual. Then:

$$\mu = \sum_{j=0}^{\infty} jP_j,$$

$$\sigma^2 = \sum_{j=0}^{\infty} (j-\mu)^2 P_j.$$

We also let Z_r be the number of offspring from individual r in the $(n-1)$st generation. Hence:

$$X_n = \sum_{r=1}^{X_{n-1}} Z_r$$

$$
\begin{aligned}
E[X_n] &= E[E[X_n \mid X_{n-1}]] \\
&= E[E[\sum_{r=1}^{X_{n-1}} Z_r \mid X_{n-1}]] \\
&= E[X_{n-1}\mu] = \mu E[X_{n-1}]
\end{aligned}
$$

Since we have assumed that $X_0 = 1$, it follows by induction that:

$$E[X_n] = \mu^n.$$

To find $\mathrm{Var}[X_n]$ we use that:

$$
\begin{aligned}
\mathrm{Var}[X_n] &= E[\mathrm{Var}(X_n \mid X_{n-1})] + \mathrm{Var}[E(X_n \mid X_{n-1})] \\
&= E[\mathrm{Var}(\sum_{r=1}^{X_{n-1}} Z_r \mid X_{n-1})] + \mathrm{Var}[E(\sum_{r=1}^{X_{n-1}} Z_r \mid X_{n-1})]
\end{aligned}
$$

Chapter 4.7 Branching Processes (cont.)

$$= E[\mathrm{Var}(\sum_{r=1}^{X_{n-1}} Z_r \mid X_{n-1})] + \mathrm{Var}[E(\sum_{r=1}^{X_{n-1}} Z_r \mid X_{n-1})]$$

$$= E[X_{n-1}\sigma^2] + \mathrm{Var}[X_{n-1}\mu] = \sigma^2\mu^{n-1} + \mu^2\,\mathrm{Var}[X_{n-1}]$$

$$= \sigma^2\mu^{n-1} + \mu^2(\sigma^2\mu^{n-2} + \mu^2\,\mathrm{Var}[X_{n-2}])$$

$$= \sigma^2(\mu^{n-1} + \mu^n) + \mu^4\,\mathrm{Var}[X_{n-2}]$$

$$= \cdots$$

$$= \sigma^2(\mu^{n-1} + \mu^n + \cdots \mu^{2n-2}) + \mu^{2n}\,\mathrm{Var}[X_0]$$

$$= \sigma^2(\mu^{n-1} + \mu^n + \cdots \mu^{2n-2})$$

Hence, we get:

$$\mathrm{Var}[X_n] = \begin{cases} \sigma^2\mu^{n-1}(\frac{1-\mu^n}{1-\mu}), & \text{if } \mu \neq 1 \\ n\sigma^2, & \text{if } \mu = 1 \end{cases}$$

Chapter 4.7 Branching Processes (cont.)

We then consider the probability that the population eventually dies out:

$$\pi_0 = \lim_{n \to \infty} P(X_n = 0 \mid X_0 = 1)$$

We first note that:

$$\begin{aligned}
\mu^n = E[X_n] = E[X_n \mid X_0 = 1] &= \sum_{j=1}^{\infty} j \cdot P(X_n = j \mid X_0 = 1) \\
&\geq \sum_{j=1}^{\infty} 1 \cdot P(X_n = j \mid X_0 = 1) \\
&= P(X_n \geq 1 | X_0 = 1) = 1 - P(X_n = 0 \mid X_0 = 1)
\end{aligned}$$

Hence, it follows that if $\mu < 1$, then $\pi_0 = 1$, since:

$$1 \geq \pi_0 = \lim_{n \to \infty} P(X_n = 0 \mid X_0 = 1) \geq 1 - \lim_{n \to \infty} \mu^n = 1 - 0 = 1.$$

Chapter 4.7 Branching Processes (cont.)

In general we have:

$$\begin{aligned}\pi_0 &= P(\text{The population dies out}) \\ &= \sum_{j=0}^{\infty} P(\text{The population dies out} \mid X_1 = j) P_j \\ &= \sum_{j=0}^{\infty} \pi_0^j P_j\end{aligned} \tag{1}$$

It can be shown that π_0 is the smallest positive number that satisfies (1).

NOTE: Since $\sum_{j=0}^{\infty} P_j = 1$, we see that $\pi_0 = 1$ is one solution to (1).

We now introduce the following functions:

$$\phi(z) = \sum_{j=0}^{\infty} z^j P_j, \quad \ell(z) = z,$$

and note that a solution z to the equation (1) is found by solving $\phi(z) = \ell(z)$.

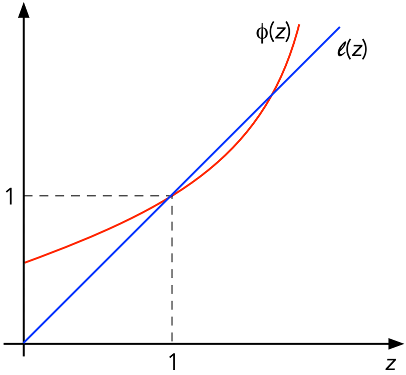
We observe that:

$$\phi(0) = \sum_{j=0}^{\infty} 0^j P_j = P_0 > 0, \qquad \phi(1) = \sum_{j=0}^{\infty} 1^j P_j = 1,$$

$$\phi'(z) = \sum_{j=1}^{\infty} j \cdot z^{j-1} P_j, \qquad \phi'(1) = \sum_{j=1}^{\infty} jP_j = \mu,$$

$$\phi''(z) = \sum_{j=2}^{\infty} j(j-1) \cdot z^{j-1} P_j, \qquad \phi''(z) > 0 \text{ for all } z > 0.$$

CASE 1. $\phi'(1) = \mu < 1$

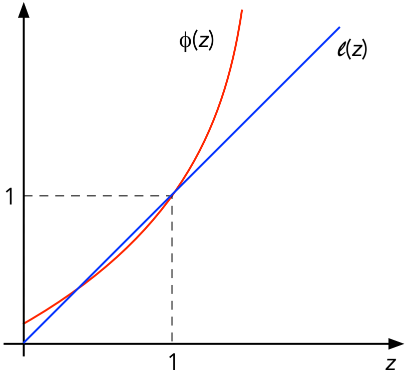


In this case $\phi(z) = \ell(z)$ for $z = 1$ and some $z > 1$.

The smallest positive number that satisfies (1) is $\pi_0 = 1$.

Chapter 4.7 Branching Processes (cont.)

CASE 1. $\phi'(1) = \mu > 1$

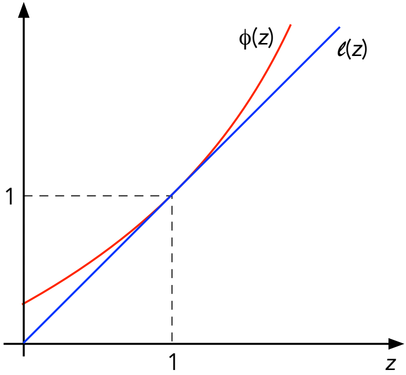


In this case $\phi(z) = \ell(z)$ for $z = 1$ and some $0 < z < 1$.

The smallest positive number that satisfies (1) is $\pi_0 < 1$.

Chapter 4.7 Branching Processes (cont.)

CASE 1. $\phi'(1) = \mu = 1$



In this case $\phi(z) = \ell(z)$ for $z = 1$ only.

The only positive number that satisfies (1) is $\pi_0 = 1$.

Chapter 4.7 Branching Processes (cont.)

CONCLUSION:

- If $\mu \leq 1$, then $\pi_0 = P(\text{The population dies out}) = 1$.
- If $\mu > 1$, then $\pi_0 = P(\text{The population dies out}) < 1$.

Example 4.34

Assume that $P_0 = \frac{1}{2}$, $P_1 = \frac{1}{4}$ and $P_2 = \frac{1}{4}$. Find π_0.

SOLUTION:

$$\begin{aligned} \mu &= 0 \cdot P_0 + 1 \cdot P_1 + 2 \cdot P_2 \\ &= 0 \cdot \tfrac{1}{2} + 1 \cdot \tfrac{1}{4} + 2 \cdot \tfrac{1}{4} = \tfrac{3}{4} < 1. \end{aligned}$$

Hence, we must have $\pi_0 = 1$.

Example 4.35

Assume that $P_0 = \frac{1}{4}$, $P_1 = \frac{1}{4}$ and $P_2 = \frac{1}{2}$. Find π_0.

SOLUTION:

$$\begin{aligned}\mu &= 0 \cdot P_0 + 1 \cdot P_1 + 2 \cdot P_2 \\ &= 0 \cdot \tfrac{1}{4} + 1 \cdot \tfrac{1}{4} + 2 \cdot \tfrac{1}{2} = \tfrac{5}{4} > 1.\end{aligned}$$

In order to find π_0 we solve (1), which in this case becomes:

$$\pi_0 = \pi_0^0 P_0 + \pi_0^1 P_1 + \pi_0^2 P_2 = \tfrac{1}{4} + \tfrac{1}{4}\pi_0 + \tfrac{1}{2}\pi_0^2.$$

or equivalently:

$$2\pi_0^2 - 3\pi_0 + 1 = 2(\pi_0 - 1)(\pi_0 - \tfrac{1}{2}) = 0$$

Hence, the smallest positive number that satisfies (1) is $\pi_0 = \frac{1}{2}$.

Chapter 4.8 Time Reversible Markov Chains

Consider an ergodic Markov chain with transition probabilities P_{ij} and stationary probabilities π_i, $i, j \in \mathcal{S}$.

Then let n be so large that we have reached a stationary state, i.e. $P_{ij}^n \approx \pi_j$.

We then consider the backwards chain $X_n, X_{n-1}, X_{n-2}, \ldots$

The backwards chain is also a Markov chain with transitions probabilities Q_{ij}, $i, j \in \mathcal{S}$ given by:

$$Q_{ij} = P(X_m = j \mid X_{m+1} = i) = \frac{P(X_m = j \cap X_{m+1} = i)}{P(X_{m+1} = i)}$$

$$= \frac{P(X_m = j) P(X_{m+1} = i \mid X_m = j)}{P(X_{m+1} = i)} = \frac{\pi_j P_{ji}}{\pi_i}.$$

We say that $\{X_n\}$ is time reversible if $Q_{ij} = P_{ij}$ for all $i, j \in \mathcal{S}$.

Hence, $\{X_n\}$ is time reversible if and only if:

$$\frac{\pi_j P_{ji}}{\pi_i} = P_{ij}, \qquad \text{for all } i, j \in \mathcal{S}.$$

or equivalently if and only if:

$$\pi_i P_{ij} = \pi_j P_{ji}, \qquad \text{for all } i, j \in \mathcal{S}.$$

Chapter 4.8 Time Reversible Markov Chains (cont.)

Assume that we can find non-negative numbers x_i, $i \in \mathcal{S}$ such that:

$$x_i P_{ij} = x_j P_{ji}, \quad \text{for all } i, j \in \mathcal{S}, \quad \text{and } \sum_{i \in \mathcal{S}} x_i = 1. \tag{2}$$

Then the Markov chain is time reversible.

PROOF: If x_i, $i \in \mathcal{S}$ satisfy (2), then it follows that:

$$\sum_{i \in \mathcal{S}} x_i P_{ij} = x_j \sum_{i \in \mathcal{S}} P_{ji} = x_j, \quad \text{for all } j \in \mathcal{S} \quad \text{and } \sum_{i \in \mathcal{S}} x_i = 1. \tag{3}$$

We have proved that the equations (3) have the unique solution:

$$x_i = \pi_i, \quad \text{for all } i \in \mathcal{S},$$

which completes the proof.

Example 4.37

Consider a Markov chain $\{X_n\}$ with state space $\mathcal{S} = \{0, 1, \ldots, M\}$ and transition probabilities:

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

In matrix form we have

$$\boldsymbol{P} = \begin{bmatrix}
1-\alpha_0 & \alpha_0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
1-\alpha_1 & 0 & \alpha_1 & 0 & \ldots & 0 & 0 & 0 \\
0 & 1-\alpha_2 & 0 & \alpha_2 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 0 & \alpha_{M-2} & 0 \\
0 & 0 & 0 & 0 & \ldots & 1-\alpha_{M-1} & 0 & \alpha_{M-1} \\
0 & 0 & 0 & 0 & \ldots & 0 & 1-\alpha_M & \alpha_M
\end{bmatrix}$$

Example 4.37 (cont.)

In this case the long run rate of transitions from i to $i+1$ must be equal to the long run rate of transitions from $i+1$ to i. From this it can be shown that:

$$\pi_i P_{i,i+1} = \pi_{i+1} P_{i+1,i}, \quad i = 0, 1, \ldots, (M-1).$$

That is, the Markov chain is time reversible.

In order to find the stationary probabilities we solve the following equations:

$$\begin{aligned}
\pi_0 \alpha_0 &= \pi_1 (1 - \alpha_1), \\
\pi_1 \alpha_1 &= \pi_2 (1 - \alpha_2), \\
&\vdots \\
\pi_{M-1} \alpha_{M-1} &= \pi_M (1 - \alpha_M)
\end{aligned}$$

Example 4.37 (cont.)

Hence, we get:

$$\pi_1 = \frac{\alpha_0}{1-\alpha_1}\pi_0,$$

$$\pi_2 = \frac{\alpha_1}{1-\alpha_2}\pi_1 = \frac{\alpha_1\alpha_0}{(1-\alpha_2)(1-\alpha_1)}\pi_0,$$

$$\vdots$$

$$\pi_M = \frac{\alpha_{M-1}}{1-\alpha_M}\pi_{M-1} = \frac{\alpha_{M-1}\cdots\alpha_1\alpha_0}{(1-\alpha_M)\cdots(1-\alpha_2)(1-\alpha_1)}\pi_0.$$

Example 4.37 (cont.)

We then use that $\sum_{j=0}^{M} \pi_j = 1$ and get:

$$\pi_0 \left[1 + \sum_{j=1}^{M} \frac{\alpha_{j-1} \cdots \alpha_1 \alpha_0}{(1-\alpha_j) \cdots (1-\alpha_2)(1-\alpha_1)} \right] = 1$$

From this it follows that:

$$\pi_0 = \left[1 + \sum_{j=1}^{M} \frac{\alpha_{j-1} \cdots \alpha_1 \alpha_0}{(1-\alpha_j) \cdots (1-\alpha_2)(1-\alpha_1)} \right]^{-1}$$

and that:

$$\pi_j = \frac{\alpha_{j-1} \cdots \alpha_1 \alpha_0}{(1-\alpha_j) \cdots (1-\alpha_2)(1-\alpha_1)} \pi_0, \quad j = 1, \ldots, M.$$

Example 4.37 (cont.)

Assume in particular that $\alpha_j = \alpha$, $j = 0, 1, \ldots, M$ and let $\beta = \alpha/(1-\alpha)$.

We then get:

$$\pi_0 = \left[1 + \sum_{j=1}^{M} \frac{\alpha^j}{(1-\alpha)^j}\right]^{-1}$$

$$= \left[\frac{1-\beta^{M+1}}{1-\beta}\right]^{-1} = \frac{1-\beta}{1-\beta^{M+1}},$$

and:

$$\pi_j = \frac{\beta^j(1-\beta)}{1-\beta^{M+1}}, \quad j = 1, \ldots, M.$$

Example 4.37 (cont.)

SPECIAL CASE: Two urns with a total of M items (molecules). At each step one item is sampled from the total population and moved from this urn to the other.

$$X_n = \text{The number of items in urn 1 at the } n\text{th step.}$$

In this case we get:

$$\alpha_j = \frac{M - j}{M}, \qquad (1 - \alpha_j) = \frac{j}{M}, \quad j = 0, 1, \ldots, M.$$

NOTE: $\alpha_0 = 1$ and $\alpha_M = 0$.

Example 4.37 (cont.)

Hence, we get:

$$\pi_0 = \left[1 + \sum_{j=1}^{M} \frac{\alpha_{j-1} \cdots \alpha_1 \alpha_0}{(1-\alpha_j)\cdots(1-\alpha_2)(1-\alpha_1)}\right]^{-1}$$

$$= \left[1 + \sum_{j=1}^{M} \frac{(M-j+1)\cdots(M-1)M}{j(j-1)\cdots 2 \cdot 1}\right]^{-1}$$

$$= \left[\sum_{j=0}^{M} \binom{M}{j}\right]^{-1} = \left[\sum_{j=0}^{M} \binom{M}{j} \cdot 1^j \cdot 1^{M-j}\right]^{-1}$$

$$= \left[(1+1)^M\right]^{-1} = \left(\frac{1}{2}\right)^M$$

Example 4.37 (cont.)

Furthermore, we get:

$$\pi_j = \frac{\alpha_{j-1} \cdots \alpha_1 \alpha_0}{(1-\alpha_j) \cdots (1-\alpha_2)(1-\alpha_1)} \pi_0$$

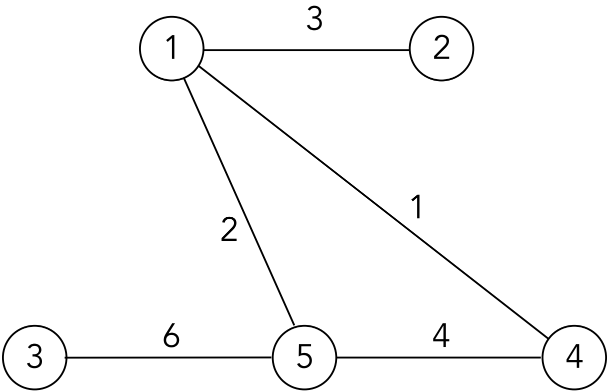
$$= \frac{(M-j+1) \cdots (M-1)M}{j(j-1) \cdots 2 \cdot 1} \pi_0$$

$$= \binom{M}{j} \left(\frac{1}{2}\right)^M, \quad j = 0, 1, 2, \ldots, M.$$

NOTE: This implies that $X_n \sim Bin(M, \frac{1}{2})$ when n is large.

Example 4.38

Undirected graph with weighted edges.



Example 4.38 (cont.)

The nodes represent states of a Markov chain with state space $\mathcal{S}$.

Thus, we define:

$$X_n = \text{The node where the process is at step } n, \quad n = 0, 1, 2, \ldots$$

We then introduce weights:

$$w_{ij} = \text{The weight associated with the edge between node } i \text{ and } j, \quad i, j \in \mathcal{S}.$$

and let:

$$P_{ij} = \frac{w_{ij}}{\sum_{k \in \mathcal{S}} w_{ik}}, \quad i, j \in \mathcal{S}.$$

Example 4.38 (cont.)

The time reversibility equations:

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad i, j \in \mathcal{S}$$

then become:

$$\pi_i \frac{w_{ij}}{\sum_{k \in \mathcal{S}} w_{ik}} = \pi_j \frac{w_{ji}}{\sum_{k \in \mathcal{S}} w_{jk}}, \quad i, j \in \mathcal{S}$$

Since $w_{ij} = w_{ji}$, the equations simplify to:

$$\frac{\pi_i}{\sum_{k \in \mathcal{S}} w_{ik}} = \frac{\pi_j}{\sum_{k \in \mathcal{S}} w_{jk}}, \quad i, j \in \mathcal{S}$$

which equivalent to:

$$\frac{\pi_i}{\sum_{k \in \mathcal{S}} w_{ik}} = c, \quad i \in \mathcal{S}$$

Example 4.38 (cont.)

Alternatively, these equations can be written as:

$$\pi_i = c \sum_{k \in \mathcal{S}} w_{ik}, \quad i \in \mathcal{S}$$

Summing over all i we get:

$$\sum_{i \in \mathcal{S}} \pi_i = c \sum_{i \in \mathcal{S}} \sum_{k \in \mathcal{S}} w_{ik} = 1.$$

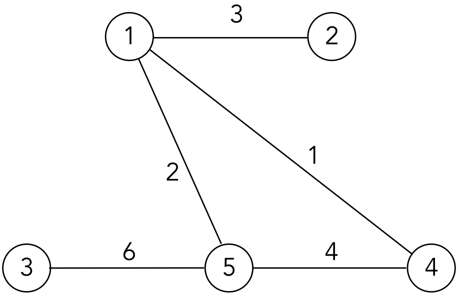
Hence,

$$c = \left[\sum_{i \in \mathcal{S}} \sum_{k \in \mathcal{S}} w_{ik} \right]^{-1}$$

Thus, we get the stationary probabilities:

$$\pi_i = \frac{\sum_{k \in \mathcal{S}} w_{ik}}{\sum_{i \in \mathcal{S}} \sum_{k \in \mathcal{S}} w_{ik}}, \quad i \in \mathcal{S}$$

Example 4.38 (cont.)



In this graph we get:

$$\pi_1 = \frac{6}{32}, \quad \pi_2 = \frac{3}{32}, \quad \pi_3 = \frac{6}{32}, \quad \pi_4 = \frac{5}{32}, \quad \pi_5 = \frac{12}{32}.$$

Chapter 4.8 Time Reversible Markov Chains (cont.)

We recall that the time reversibility equations implies that:

$$x_i P_{ij} = x_j P_{ji}$$
$$x_k P_{kj} = x_j P_{jk}$$

Assuming that $P_{ij} P_{jk} > 0$ these equations imply that:

$$x_i = x_j \frac{P_{ji}}{P_{ij}}$$
$$x_j = x_k \frac{P_{kj}}{P_{jk}}$$

Hence,

$$\frac{x_i}{x_k} = \frac{P_{kj} P_{ji}}{P_{ij} P_{jk}}$$

At the same time the time reversibility equations implies that:

$$\frac{x_i}{x_k} = \frac{P_{ki}}{P_{ik}}$$

Thus, for a valid solution to the time reversibility equations we must have that:

$$\frac{P_{kj}P_{ji}}{P_{ij}P_{jk}} = \frac{P_{ki}}{P_{ik}}$$

or equivalently:

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

Chapter 4.8 Time Reversible Markov Chains (cont.)

Theorem

A stationary 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:

$$P_{i,i_1} P_{i_1,i_2} \cdots P_{i_k,i} = P_{i,i_k} P_{i_k,i_{k-1}} \cdots P_{i_1,i}$$

for all states $i, i_1, \ldots, i_k$, $k = 1, 2, \ldots$.

PROOF: That this condition is necessary essentially follows from the argument above. We thus focus on proving sufficiency.

We fix i and j and write the condition in the theorem as:

$$P_{i,i_1} P_{i_1,i_2} \cdots P_{i_k,j} P_{j,i} = P_{i,j} P_{j,i_k} P_{i_k,i_{k-1}} \cdots P_{i_1,i}$$

By summing over all paths of length $k+1$ we get that:

$$P_{ij}^{k+1} P_{ji} = P_{ij} P_{ji}^{k+1}$$

We then sum over k from 1 to m and divide by m:

$$\frac{P_{ji}\sum_{k=1}^{m} P_{ij}^{k+1}}{m} = \frac{P_{ij}\sum_{k=1}^{m} P_{ji}^{k+1}}{m}$$

By letting $m \to \infty$ this implies that:

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

Hence, we conclude that the chain is time reversible.