STK2130 – Lecture 6 - part 2

A. B. Huseby

Department of Mathematics
University of Oslo, Norway

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:

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

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

Chapter 4.8 Time Reversible Markov Chains (cont.)

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:

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

or equivalently if and only if:

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

Chapter 4.8 Time Reversible Markov Chains (cont.)

Theorem (Time reversible chain)

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

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

Then the Markov chain is time reversible.

PROOF: If x_i, $i \in \mathcal{S}$ satisfy (1), then it follows that:

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

We have proved that the equations (2) 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:

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

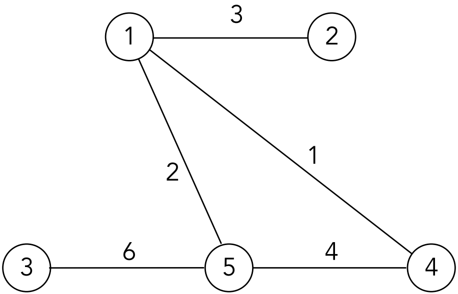
$$= \frac{(M-j+1) \cdots (M-1)M}{j(j-1) \cdots 2 \cdot 1} \pi_0$$

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

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

Example 4.38

Undirected graph with weighted edges.



Here the transition probabilities are given by:

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

Thus, e.g.,

$$P_{1,2} = \tfrac{3}{6}, \quad P_{1,4} = \tfrac{1}{6}, \quad P_{1,5} = \tfrac{2}{6}$$

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:

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

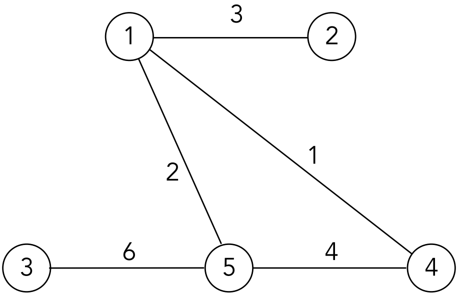
Hence,

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

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:

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

Chapter 4.8 Time Reversible Markov Chains (cont.)

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:

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

Chapter 4.8 Time Reversible Markov Chains (cont.)

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.

Doubly stochastic transition matrices

Let $\boldsymbol{P}$ be the transition probability matrix of a Markov chain $\{X_n\}$ with state space $\mathcal{S}$ where $|\mathcal{S}| = n < \infty$. We then know that:

$$\sum_{j \in \mathcal{S}} P_{ij} \cdot 1 = 1, \qquad \text{for all } i \in \mathcal{S}.$$

If $\boldsymbol{e} = (1, \ldots, 1)$ is an n-dimensional vector, we can rewrite this as:

$$\boldsymbol{P}\boldsymbol{e}^T = \boldsymbol{e}^T.$$

We say that $\boldsymbol{P}$ is doubly stochastic if we also have:

$$\sum_{i \in \mathcal{S}} P_{ij} \cdot 1 = 1, \qquad \text{for all } j \in \mathcal{S}.$$

In matrix notation this condition can be expressed as:

$$\boldsymbol{e}\boldsymbol{P} = \boldsymbol{e}.$$

Doubly stochastic transition matrices (cont.)

Proposition

Let $\boldsymbol{P}$ be the transition probability matrix of a Markov chain $\{X_n\}$ with state space $\mathcal{S}$ where $|\mathcal{S}| = n < \infty$. Then $\boldsymbol{P}$ is doubly stochastic if and only if the stationary distribution is uniform, i.e.: $\boldsymbol{\pi} = \frac{1}{n}\boldsymbol{e}$.

PROOF:

$\boldsymbol{P}$ is doubly stochastic

$$\Updownarrow$$

$$\boldsymbol{e}\boldsymbol{P} = \boldsymbol{e}$$

$$\Updownarrow$$

$$\frac{1}{n}\boldsymbol{e}\boldsymbol{P} = \frac{1}{n}\boldsymbol{e}$$

$$\Updownarrow$$

$$\boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}$$

$$\Updownarrow$$

The stationary distribution is uniform

Doubly stochastic transition matrices (cont.)

Proposition

Let $\boldsymbol{P}$ *be the transition probability matrix of a Markov chain* $\{X_n\}$ *with state space* $\mathcal{S}$ *where* $|\mathcal{S}| = n < \infty$. *Assume that* $\boldsymbol{P}$ *is doubly stochastic. Then the Markov chain is time reversible if and only if* $\boldsymbol{P}$ *is symmetric.*

PROOF: Since $\boldsymbol{P}$ is doubly stochastic $\pi_i = \frac{1}{n}$ for all $i \in \mathcal{S}$. Hence, we get:

$$Q_{ij} = \frac{\pi_j P_{ji}}{\pi_i} = \frac{\frac{1}{n} P_{ji}}{\frac{1}{n}} = P_{ji} \quad i, j \in \mathcal{S}$$

Thus, $\boldsymbol{Q} = \boldsymbol{P}$ if and only if $P_{ji} = P_{ij}$ for all $i, j \in \mathcal{S}$.

That is, $\boldsymbol{Q} = \boldsymbol{P}$ if and only if $\boldsymbol{P}$ is symmetric.

Example

Let $\boldsymbol{P}$ be the transition probability matrix of a Markov chain $\{X_n\}$ with state space $\mathcal{S} = \{0, 1, 2\}$, where:

$$\boldsymbol{P} = \begin{bmatrix} 0.2 & 0.4 & 0.4 \\ 0.4 & 0.2 & 0.4 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}$$

Since $\boldsymbol{P}$ is doubly stochastic and symmetric, the Markov chain is time reversible.

Example

Let $\boldsymbol{P}$ be the transition probability matrix of a Markov chain $\{X_n\}$ with state space $\mathcal{S} = \{0, 1, 2\}$, where:

$$\boldsymbol{P} = \begin{bmatrix} 0.0 & 0.1 & 0.9 \\ 0.9 & 0.0 & 0.1 \\ 0.1 & 0.9 & 0.0 \end{bmatrix}$$

Since $\boldsymbol{P}$ is doubly stochastic and not symmetric, the Markov chain is not time reversible.

Time Reversible Markov Chains (cont.)

Proposition (4.9)

Consider an irreducible Markov chain with transition probability matrix $\boldsymbol{P}$ *and state space* $\mathcal{S}$*. If we can find a stationary distribution* $\boldsymbol{\pi}$ *on* $\mathcal{S}$*, and a transition probability matrix* $\boldsymbol{Q}$ *such that:*

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

then $\boldsymbol{Q}$ *is the transition probability matrix of the reversed chain, and* $\boldsymbol{\pi}$ *is the stationary distribution for both the original and reversed chain.*

PROOF: The result is an immediate consequence of the previously established formula:

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

Example 4.40

Let L denote the lifetime of a type of light bulb, and assume that the distribution of L in days is:

$$P(L = i) = p_i, \quad i = 1, 2, \ldots$$

We have an infinite supply of light bulb of this kind, with lifetimes $L_1, L_2, \ldots$ being independent and with the same distribution as L. Each time a bulb fails, it is replaced by a new one the next day.

$$X_n = \text{The age of the light bulb at day } n, \quad n = 1, 2, \ldots$$

Then $\{X_n\}$ is a Markov chain with state space $\mathcal{S} = \{1, 2, \ldots\}$ and with transition probabilities:

$$P_{i,1} = \frac{P(L = i)}{P(L \geq i)} = \frac{p_i}{\sum_{j=i}^{\infty} p_j}, \quad P_{i,i+1} = 1 - P_{i,1}, \quad i = 1, 2, \ldots$$

Example 4.40 (cont.)

We then claim that the reversed chain has transition probabilities:

$$Q_{i,i-1} = 1, \quad i > 1$$
$$Q_{1,i} = P(L = i) = p_i, \quad i \geq 1$$

To show this we need to find a stationary distribution $\pi_1, \pi_2, \ldots$ such that:

$$\pi_i P_{ij} = \pi_j Q_{ji}, \quad i, j = 1, 2, \ldots$$

We start out by letting $j = 1$ and determine $\pi_1, \pi_2, \ldots$ such that:

$$\pi_i P_{i,1} = \pi_i \frac{P(L = i)}{P(L \geq i)} = \pi_1 Q_{1,i} = \pi_1 P(L = i).$$

This is equivalent to:

$$\pi_i = \pi_1 P(L \geq i)$$

Example 4.40 (cont.)

Summing over all i and using that the π_i-s add up to 1, we get:

$$
\begin{aligned}
1 = \sum_{i=1}^{\infty} \pi_i &= \pi_1 \sum_{i=1}^{\infty} P(L \geq i) \\
&= \pi_1 \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} p_j = \pi_1 \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_j I(j \geq i) \\
&= \pi_1 \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} p_j I(i \leq j) = \pi_1 \sum_{j=1}^{\infty} p_j \sum_{i=1}^{\infty} I(i \leq j) \\
&= \pi_1 \sum_{j=1}^{\infty} p_j \cdot j = \pi_1 E[L].
\end{aligned}
$$

Hence, it follows that $\pi_1 = (E[L])^{-1}$, and thus:

$$
\pi_i = \pi_1 P(L \geq i) = \frac{P(L \geq i)}{E[L]}, \quad i = 1, 2, \ldots
$$

Example 4.40 (cont.)

We note that if $j > 1$, we have $P_{ij} = Q_{ji} = 0$ for $j \neq i + 1$. Hence, it remains to verify that:

$$\pi_i P_{ij} = \pi_j Q_{ji}, \quad i = 1, 2, \ldots, \quad j = i + 1.$$

Using the expressions for $\pi_1, \pi_2, \ldots$ and that $Q_{i+1,i} = 1$ for $i > 1$, this is equivalent to:

$$\frac{P(L \geq i)}{E[L]} \left(1 - \frac{P(L = i)}{P(L \geq i)}\right) = \frac{P(L \geq i + 1)}{E[L]}$$

By simplifying this equation we get:

$$P(L \geq i) - P(L = i) = P(L \geq i + 1)$$

which is trivially true.