STK2130 – Lecture 2

A. B. Huseby

Department of Mathematics
University of Oslo, Norway

Section 4.1 – Introduction (from Lecture 1)

Markov Chains

A discrete time, discrete space stochastic process $\{X_0, X_1, X_2, \dots\}$ is called a time homogenuous Markov chain if the transition probabilities satisfy the following:

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) = P_{ij}$$

for all states $i_0, i_1, \dots, i_{n-1}, i, j$ and all $n \geq 0$.

NOTE: The Markov assumption implies that given X_n, the next state, X_{n+1}, is independent of all previous states $X_0, X_1, \dots, X_{n-1}$. The process is time homogenuous since P_{ij} is independent of n.

One-step transition probability matrix:

$$\boldsymbol{P} = [P_{ij}]_{i,j \in \mathcal{S}}$$

where $P_{ij} \geq 0$ for all $i, j \in \mathcal{S}$, and $\sum_{j \in \mathcal{S}} P_{ij} = 1$ for all $i \in \mathcal{S}$.

Section 4.2 – Chapman-Kolmogorov Equations (from Lecture 1)

Chapman-Kolmogorov Equations

***n*-step transition probabilities:**

$$P_{ij}^{n} = P\{X_{n+k} = j | X_k = i\}, \quad n \geq 1$$

***n*-step transition probability matrix:**

$$\boldsymbol{P}^{(n)} = \left[P_{ij}^{n}\right]_{i,j \in \mathcal{S}}$$

Chapman-Kolmogorov Equations: By conditioning on X_n we get:

$$P_{ij}^{n+m} = \sum_{k \in \mathcal{S}} P_{ik}^{n} \cdot P_{kj}^{m}$$

By using the definition of matrix products it follows that:

$$\boldsymbol{P}^{(n+m)} = \boldsymbol{P}^{(n)} \cdot \boldsymbol{P}^{(m)}$$

$$\boldsymbol{P}^{(n)} = \boldsymbol{P}^{(n-1)} \cdot \boldsymbol{P}^{(1)}$$

$$\boldsymbol{P}^{(n)} = \boldsymbol{P}^{n}$$

First-passage probabilities

Consider a Markov chain $\{X_n\}$ with state space $\mathcal{S}$ and transition probability matrix $\boldsymbol{P}$, and let $\mathcal{A}$ be a non-empty proper subset of $\mathcal{S}$.

We want to calculate the following probability:

$$\beta = P(X_k \in \mathcal{A} \text{ for some } 1 \leq k \leq m \,|\, X_0 = i)$$

where we assume that $i \notin \mathcal{A}$.

In order to analyze this we introduce:

$$N = \min\{n > 0 : X_n \in \mathcal{A}\},$$

Thus, N represents the first time the Markov chain enters $\mathcal{A}$.

First-passage probabilities (cont.)

In order to solve this problem we introduce a new Markov chain $\{W_n\}$ derived from $\{X_n\}$ as follows:

$$W_n = \begin{cases} X_n & \text{if } n < N \\ A & \text{if } n \geq N \end{cases}$$

To avoid confusion, we let the state A be some arbitrary state in $\mathcal{A}$. Then the state space of $\{W_n\}$ becomes $(\mathcal{S} \setminus \mathcal{A}) \cup \{A\}$.

We observe that when $\{X_n\}$ enters $\mathcal{A}$, i.e., at time N, $\{W_n\}$ is absorbed in state A.

First-passage probabilities (cont.)

The transition probability matrix of $\{W_n\}$, denoted $\boldsymbol{Q}$, is given by:

$$Q_{i,j} = P_{i,j}, \qquad i \notin \mathcal{A}, j \notin \mathcal{A},$$

$$Q_{i,A} = \sum_{j \in \mathcal{A}} P_{i,j}, \quad i \notin \mathcal{A},$$

$$Q_{A,A} = 1.$$

We now have:

$$\begin{aligned} \beta &= P(X_k \in \mathcal{A} \text{ for some } 1 \leq k \leq m | X_0 = i) \\ &= P(W_m = A | X_0 = i) \\ &= P(W_m = A | W_0 = i) = Q^m_{i,A}. \end{aligned}$$

Example 4.12

In a sequence of independent flips of a fair coin, we introduce a Markov chain $\{X_n\}$ defined as:

$$X_n = \text{The length of the current sequence of consecutive heads}$$

The state space of this Markov chain is $\mathcal{S} = \{0, 1, 2, \ldots\}$, and we note that $X_0 = 0$.

Moreover, the transition probability matrix of $\{X_n\}$ is given by:

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

Example 4.12 (cont.)

We then let $\mathcal{A} = \{3, 4, 5, \ldots\}$, and introduce:

$$N = \min\{n : X_n \in \mathcal{A}\},$$

Thus, N denotes the number of flips until there is a run of 3 consecutive heads.

We want to calculate $P(N \leq 8)$ and $P(N = 8)$.

To solve this problem we introduce a Markov chain $\{W_n\}$ with states $\mathcal{T} = \{0, 1, 2, 3\}$ defined relative to the sequence of coin flips as follows:

$$W_n = \begin{cases} X_n & \text{if } n < N \\ 3 & \text{if } n \geq N \end{cases}$$

Example 4.12 (cont.)

We note that state 3 is an absorbing state for $\{W_n\}$, and that the transition probability matrix of $\{W_n\}$ is given by:

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

Example 4.12 (cont.)

We then calculate the 2, 4, 7 and 8 steps transition probability matrices:

$$\boldsymbol{Q}^{(2)} = \begin{bmatrix} 0.5000 & 0.2500 & 0.2500 & 0.0000 \\ 0.5000 & 0.2500 & 0.0000 & 0.2500 \\ 0.2500 & 0.2500 & 0.0000 & 0.5000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

$$\boldsymbol{Q}^{(4)} = \begin{bmatrix} 0.4375 & 0.2500 & 0.1250 & 0.1875 \\ 0.3750 & 0.1875 & 0.1250 & 0.3125 \\ 0.2500 & 0.1250 & 0.0625 & 0.5625 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

Example 4.12 (cont.)

$$\boldsymbol{Q}^{(7)} = \begin{bmatrix} 0.3438 & 0.1875 & 0.1016 & 0.3672 \\ 0.2891 & 0.1563 & 0.0859 & 0.4688 \\ 0.1875 & 0.1016 & 0.0547 & 0.6563 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

$$\boldsymbol{Q}^{(8)} = \begin{bmatrix} 0.3164 & 0.1719 & 0.0938 & 0.4180 \\ 0.2656 & 0.1445 & 0.0781 & 0.5117 \\ 0.1719 & 0.0938 & 0.0508 & 0.6836 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

Example 4.12 (cont.)

Using these matrices we get that:

$$P(N \leq 8) = Q_{0,3}^{8} = 0.4180$$

$$\begin{aligned} P(N = 8) &= P(N \leq 8) - P(N \leq 7) \\ &= Q_{0,3}^{8} - Q_{0,3}^{7} \\ &= 0.4180 - 0.3672 = 0.0508 \end{aligned}$$

Alternatively:

$$P(N = 8) = Q_{0,2}^{7} \cdot 0.5 = 0.1016 \cdot 0.5 = 0.0508$$

Example 4.13

Let $\{X_n\}$ be a Markov chain with state set $\mathcal{S} = \{0, 1, 2, 3\}$ and transition probability matrix $\boldsymbol{P}$.

We then introduce:

$$N = \min\{n \geq 4 : X_{n-3} = 1, X_{n-2} = 2, X_{n-1} = 1, X_n = 2\}$$

Thus, N is the number of steps we need to observe the sequence $(1, 2, 1, 2)$.

PROBLEM: For a given value of $k \geq 4$ calculate $P(N \leq k)$ given that $X_0 = 0$.

Example 4.13 (cont.)

To solve this problem, we start out by introducing the following process $\{Z_n\}$:

$$Z_n = \text{The minimal number of steps before observing the first } (1,2,1,2)$$

for any given time step n of the original process. The state space of $\{Z_n\}$ is $\mathcal{T} = \{0,1,2,3,4\}$ and we observe that:

$$\begin{aligned} P(Z_{n+1} = 0 | Z_n = 0) &= 1 \\ P(Z_{n+1} = 0 | Z_n = 1) &= P_{1,2} \\ P(Z_{n+1} = 1 | Z_n = 2) &= P_{2,1} \\ P(Z_{n+1} = 2 | Z_n = 3) &= P_{1,2} \end{aligned}$$

Determining $P(Z_{n+1} = 3 | Z_n = 4)$ is more difficult, but we may express this by conditioning on X_n as follows:

$$P(Z_{n+1} = 3 | Z_n = 4) = \sum_{i \in \{0,2,3\}} P(Z_{n+1} = 3 | Z_n = 4, X_n = i) \cdot P(X_n = i)$$

Example 4.13 (cont.)

NOTE: The term $P(Z_{n+1} = 3 | Z_n = 4, X_n = 1) \cdot P(X_n = 1)$ is skipped since $X_n = 1$ implies that $Z_n < 4$.

The problem with the transition probability $P(Z_{n+1} = 3 | Z_n = 4)$ is that it depends on n. In order to avoid this problem, we need a process which contains more information about the state of the original process.

Thus, we introduce a new Markov chain $\{Y_n\}$ defined as follows:

$$Y_n = Z_n + X_n \cdot I(Z_n = 4)$$

The state space of $\{Y_n\}$ is $\mathcal{T} = \{0, 1, 2, 3, 4, 6, 7\}$.

NOTE: $\mathcal{T}$ does not contain the state 5 since the combination where $Z_n = 4$ and $X_n = 1$ is not possible.

NOTE: If $Y_n \geq 4$, then $X_n = Y_n - 4$. Thus, in this case the state of the original process can be derived from Y_n. As a result the transition probabilities of $\{Y_n\}$ does not depend on n.

Example 4.13 (cont.)

The transition probability matrix of $\{Y_n\}$ with state space $\mathcal{T} = \{0, 1, 2, 3, 4, 6, 7\}$ is:

$$\boldsymbol{Q} = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_{1,2} & 0 & 0 & P_{1,1} & P_{1,0} & 0 & P_{1,3} \\ 0 & P_{2,1} & 0 & 0 & P_{2,0} & P_{2,2} & P_{2,3} \\ 0 & 0 & P_{1,2} & P_{1,1} & P_{1,0} & 0 & P_{1,3} \\ \hdashline 0 & 0 & 0 & P_{0,1} & P_{0,0} & P_{0,2} & P_{0,3} \\ 0 & 0 & 0 & P_{2,1} & P_{2,0} & P_{2,2} & P_{2,3} \\ 0 & 0 & 0 & P_{3,1} & P_{3,0} & P_{3,2} & P_{3,3} \end{array}\right]$$

NOTE: The $Y_n = 0$, the process is absorbed in this state.

If $Y_n = 1$ or $Y_n = 3$, then $X_n = 1$. If $Y_n = 2$, then $X_n = 2$.

The entries below the horizontal dashed line corresponds to cases where $Y_n \geq 4$, implying that $X_n = Y_n - 4$.

Example 4.13 (cont.)

Since we have assumed that $X_0 = 0$, it follows that $Y_0 = 4$.

Moreover, recalling that:

$$N = \min\{n \geq 4 : X_{n-3} = 1, X_{n-2} = 2, X_{n-1} = 1, X_n = 2\}$$

and using the fact that state 0 is an absorbing state for $\{Y_n\}$, it follows that:

$$N \leq k \quad \Leftrightarrow \quad Y_k = 0$$

Hence, we get that:

$$P(N \leq k) = P(Y_k = 0 | Y_0 = 4) = Q_{4,0}^k$$

First-passage probabilities (cont.)

Consider again a Markov chain $\{X_n\}$ with state space $\mathcal{S}$ and transition probability matrix $\boldsymbol{P}$, and let $\mathcal{A}$ be a non-empty proper subset of $\mathcal{S}$.

We now consider the following probability:

$$\alpha = P(X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq m-1, X_m = j | X_0 = i)$$

As before we let $N = \min\{n > 0 : X_n \in \mathcal{A}\}$ and:

$$W_n = \begin{cases} X_n & \text{if } n < N \\ A & \text{if } n \geq N \end{cases}$$

As before, we let the state A be some arbitrary state in $\mathcal{A}$, and hence, the state space of $\{W_n\}$ becomes $(\mathcal{S} \setminus \mathcal{A}) \cup \{A\}$.

First-passage probabilities (cont.)

CASE 1. $i \notin \mathcal{A}$, $j \notin \mathcal{A}$

In this case the event $\{X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq m-1, X_m = j\}$ is equivalent to the event $\{W_m = j\}$.

Hence, it follows that:

$$\alpha = P(W_m = j | W_0 = i) = Q^m_{i,j}.$$

First-passage probabilities (cont.)

CASE 2. $i \notin \mathcal{A}$ and $j \in \mathcal{A}$

In this case we have:

$$\alpha = \sum_{r \notin \mathcal{A}} P(X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq m-2, X_{m-1} = r, X_m = j | X_0 = i)$$

$$= \sum_{r \notin \mathcal{A}} P(X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq m-2, X_{m-1} = r | X_0 = i)$$
$$\cdot P(X_m = j | X_0 = i, X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq m-2, X_{m-1} = r)$$

$$= \sum_{r \notin \mathcal{A}} P(X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq m-2, X_{m-1} = r | X_0 = i) \cdot P(X_m = j | X_{m-1} = r)$$

$$= \sum_{r \notin \mathcal{A}} Q_{i,r}^{m-1} \cdot P_{r,j}$$

First-passage probabilities (cont.)

CASE 3. $i \in \mathcal{A}$ and $j \notin \mathcal{A}$

In this case we have:

$$\alpha = \sum_{r \notin \mathcal{A}} P(X_1 = r, X_k \notin \mathcal{A} \text{ for all } 2 \leq k \leq m-1, X_m = j | X_0 = i)$$

$$= \sum_{r \notin \mathcal{A}} P(X_1 = r | X_0 = i)$$
$$\cdot P(X_k \notin \mathcal{A} \text{ for all } 2 \leq k \leq m-1, X_m = j | X_0 = i, X_1 = r)$$

$$= \sum_{r \notin \mathcal{A}} P(X_1 = r | X_0 = i)$$
$$\cdot P(X_k \notin \mathcal{A} \text{ for all } 2 \leq k \leq m-1, X_m = j | X_1 = r)$$

$$= \sum_{r \notin \mathcal{A}} P_{i,r} \cdot Q_{r,j}^{m-1}$$

First-passage probabilities (cont.)

CASE 4. $i \in \mathcal{A}$ and $j \in \mathcal{A}$

By combining the previous arguments we get:

$$\alpha = \sum_{r,s \notin \mathcal{A}} P_{i,r} \cdot Q_{r,s}^{m-2} \cdot P_{s,j}$$

Example 4.14

Let $\{X_n\}$ be a Markov chain with state space $\mathcal{S} = \{1, 2, 3, 4, 5\}$, and let $\mathcal{A} = \{3, 4, 5\}$.

We now consider the following probability:

$$\alpha = P(X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq 3, X_4 = 2 | X_0 = 1)$$

We then let $N = \min\{n > 0 : X_n \in \mathcal{A}\}$ and:

$$W_n = \begin{cases} X_n & \text{if } n < N \\ 3 & \text{if } n \geq N \end{cases}$$

The state space of $\{W_n\}$ becomes $\mathcal{T} = \{1, 2, 3\}$.

Example 4.14 (cont.)

The transition probability matrix of $\{W_n\}$, denoted $\boldsymbol{Q}$, is given by:

$$Q_{i,j} = P_{i,j}, \qquad i, j = 1, 2,$$

$$Q_{i,3} = 1 - P_{i,1} - P_{i,2}, \quad i = 1, 2,$$

$$Q_{3,3} = 1.$$

We then assume that:

$$P_{1,1} = 0.3, \qquad P_{1,2} = 0.3$$

$$P_{2,1} = 0.1, \qquad P_{2,2} = 0.2$$

Example 4.14 (cont.)

Hence, we get that:

$$\boldsymbol{Q} = \begin{bmatrix} 0.3 & 0.3 & 0.4 \\ 0.1 & 0.2 & 0.7 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

From this we calculate that:

$$\boldsymbol{Q}^{(4)} = \begin{bmatrix} 0.0219 & 0.0285 & 0.9496 \\ 0.0095 & 0.0124 & 0.9781 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}$$

Hence, we get that:

$$\alpha = P(X_k \notin \mathcal{A} \text{ for all } 1 \leq k \leq 3, X_4 = 2 | X_0 = 1)$$

$$= P(W_4 = 2 | W_0 = 1) = Q^4_{1,2} = 0.0285$$

Unconditional probabilities

We introduce the probabilities:

$$\pi_i^{(n)} = P\{X_n = i\}, \quad i \in \mathcal{S}, \quad n = 0, 1, 2, \ldots$$

We then have:

$$\begin{aligned}
\pi_j^{(n+m)} &= \sum_{i \in \mathcal{S}} P\{X_{n+m} = j \cap X_n = i\} \\
&= \sum_{i \in \mathcal{S}} P\{X_n = i\} \cdot P\{X_{n+m} = j | X_n = i\} \\
&= \sum_{i \in \mathcal{S}} \pi_i^{(n)} \cdot P_{ij}^m.
\end{aligned}$$

In particular:

$$\pi_j^{(1)} = \sum_{i \in \mathcal{S}} \pi_i^{(0)} \cdot P_{ij}, \qquad \pi_j^{(n+1)} = \sum_{i \in \mathcal{S}} \pi_i^{(n)} \cdot P_{ij}$$

Unconditional probabilities (cont.)

Assume that:

$$\lim_{n\to\infty} \pi_i^{(n)} = \pi_i, \quad i \in \mathcal{S}.$$

Then we obviously also have:

$$\lim_{n\to\infty} \pi_j^{(n+1)} = \pi_j, \quad j \in \mathcal{S}.$$

In particular, if $\mathcal{S} = \{1, \ldots, k\}$ and $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_k)$, then:

$$\pi_j = \lim_{n\to\infty} \pi_j^{(n+1)} = \lim_{n\to\infty} \sum_{i=1}^{k} \pi_i^{(n)} \cdot P_{ij} = \sum_{i=1}^{k} \lim_{n\to\infty} \pi_i^{(n)} \cdot P_{ij} = \sum_{i=1}^{k} \pi_i \cdot P_{ij}$$

Thus, $\boldsymbol{\pi}$ must satisfy:

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

Example 4.8 revisited

$$P\{\text{Rain tomorrow}|\text{Rain today}\} = 0.75$$
$$P\{\text{Rain tomorrow}|\text{No rain today}\} = 0.35$$

$$\boldsymbol{P} = \begin{bmatrix} 0.75 & 0.25 \\ 0.35 & 0.65 \end{bmatrix}$$

$$\boldsymbol{P}^{(2)} = \begin{bmatrix} 0.65 & 0.35 \\ 0.49 & 0.51 \end{bmatrix}$$

$$\boldsymbol{P}^{(4)} = \begin{bmatrix} 0.5940 & 0.4060 \\ 0.5684 & 0.4316 \end{bmatrix}$$

$$\boldsymbol{P}^{(12)} = \begin{bmatrix} 0.5833 & 0.4167 \\ 0.5833 & 0.4167 \end{bmatrix}$$

Example 4.8 revisited (cont.)

We now let $\boldsymbol{\pi} = (\pi_1, \pi_2)$, and consider the equation:

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

which in this case becomes:

$$\begin{aligned} \pi_1 &= 0.75\pi_1 + 0.35\pi_2 \\ \pi_2 &= 0.25\pi_1 + 0.65\pi_2 \end{aligned}$$

By inserting $\pi_2 = 1 - \pi_1$ into the first equation, we get:

$$\pi_1 = 0.75\pi_1 + 0.35(1 - \pi_1) = 0.40\pi_1 + 0.35$$

From this it follows that:

$$\begin{aligned} \pi_1 &= 0.35/(1 - 0.40) = 0.5833 \\ \pi_2 &= 1 - \pi_1 = 0.4167 \end{aligned}$$