STK2130 – Week 4

A. B. Huseby

Department of Mathematics
University of Oslo, Norway

First-passage probabilities

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

We want to calculate the following probability:

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

In order to analyze this we introduce:

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

where we let $N = \infty$ is $X_n \notin \mathcal{A}$ for all n.

NOTE: N represents the first time the Markov chain enters $\mathcal{A}$.

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

NOTE: When $\{X_n\}$ enters $\mathcal{A}$, $\{W_n\}$ is absorbed in state A.

First-passage probabilities (cont.)

The transition probabilities of $\{W_n\}$, denoted $Q_{i,j}$, are given by:

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

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

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

We now have:

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

First-passage probabilities (cont.)

We now consider the following probability:

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

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

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:

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

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:

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

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:

$$\pi_1 = 0.75\pi_1 + 0.35\pi_2$$
$$\pi_2 = 0.25\pi_1 + 0.65\pi_2$$

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:

$$\pi_1 = 0.35/(1 - 0.40) = 0.5833$$
$$\pi_2 = 1 - \pi_1 = 0.4167$$

Chapter 4.3. Classification of States

Let $\{X_n\}$ be a Markov chain with state space $\mathcal{S}$ and transition probability matrix $\boldsymbol{P}$.

State j is said to be accessible from state i, denoted as $i \to j$, if $P_{ij}^n > 0$ for some $n \geq 0$.

Note that we have:

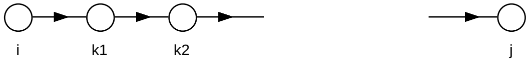
$$\max_n P_{ij}^n \leq P(\bigcup_{n=1}^{\infty}\{X_n = j\}|X_0 = i)$$
$$\leq \sum_{n=0}^{\infty} P\{X_n = j|X_0 = i\} = \sum_{n=0}^{\infty} P_{ij}^n.$$

Hence, $i \to j$ if and only if:

$$P(\bigcup_{n=1}^{\infty}\{X_n = j\}|X_0 = i) > 0.$$

A state diagram for a Markov chain is a directed graph where the nodes represent the states and the edges represent possible one-step transitions. More precisely, the state diagram contains an edge from node i to node j if and only if $P_{ij} > 0$.

If $i, j \in \mathcal{S}$, then $i \to j$ if and only if the state diagram contains at least one directed path from i to j.



If such a path exists, we have:

$$P_{ij}^{n} \geq P_{i,k_1} \cdot P_{k_1,k_2} \cdots P_{k_{n-2},k_{n-1}} \cdot P_{k_{n-1},j} > 0.$$

Communicating states

States i and j communicate, denoted as $i \leftrightarrow j$, if $i \to j$ and $j \to i$.

The relation $\leftrightarrow$ is an equivalence relation. That is $\leftrightarrow$ satisfies the following properties:

- Reflexivity: $i \leftrightarrow i$.
- Symmetry: $i \leftrightarrow j$ if and only if $j \leftrightarrow i$.
- Transitivity: $i \leftrightarrow j$ and $j \leftrightarrow k$ implies $i \leftrightarrow k$.

Reflexivity follows since we always have $P_{ii}^0 = 1 > 0$. Symmetry follows directly from the definition.

Communicating states (cont.)

To prove transitivity we assume that $i \leftrightarrow j$ and $j \leftrightarrow k$.

Hence, in particular $i \to j$ and $j \to k$, implying that there exists $m, n \geq 0$ such that $P_{ij}^m > 0$ and $P_{jk}^n > 0$.

By the Chapman-Kolmogorov equations, we have:

$$P_{ik}^{m+n} = \sum_{r \in \mathcal{S}} P_{ir}^m P_{rk}^n \geq P_{ij}^m \cdot P_{jk}^n > 0.$$

Hence, by definition $i \to k$.

By a similar argument we can show that $k \to i$ as well.

Hence, we conclude that $i \leftrightarrow k$.

Communicating states (cont.)

Two states that communicate are said to be in the same (equivalence) class.

Two classes of states are either identical or disjoint.

PROOF: Assume that $\mathcal{A}, \mathcal{B} \subseteq \mathcal{S}$ represent two equivalence classes, and assume that $\mathcal{A} \cap \mathcal{B} \neq \emptyset$. That is, there exists a state i such that $i \in \mathcal{A} \cap \mathcal{B}$.

Then choose $j \in \mathcal{A}$ and $k \in \mathcal{B}$ arbitrarily.

Now, $i, j \in \mathcal{A}$ implies that $i \leftrightarrow j$ and $i, k \in \mathcal{B}$ implies that $i \leftrightarrow k$.

Hence, by transitivity we also have $j \leftrightarrow k$. That is, j and k belong to the same equivalence class.

Since this holds for any $j \in \mathcal{A}$ and $k \in \mathcal{B}$, this implies that $\mathcal{A} = \mathcal{B}$ ■

The equivalence classes partition the state space $\mathcal{S}$ into a number of disjoint sets. A Markov chain is called irreducible if the number of equivalence classes is one.

Example 4.15

Consider a Markov chain with state space $\mathcal{S} = \{0, 1, 2\}$ and transition probability matrix:

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

We then observe:

$$\text{Since } P_{01} = \tfrac{1}{2} > 0, \text{ it follows that } 0 \rightarrow 1$$
$$\text{Since } P_{10} = \tfrac{1}{2} > 0, \text{ it follows that } 1 \rightarrow 0$$
$$\text{Since } P_{12} = \tfrac{1}{4} > 0, \text{ it follows that } 1 \rightarrow 2$$
$$\text{Since } P_{21} = \tfrac{1}{3} > 0, \text{ it follows that } 2 \rightarrow 1$$

Hence, $0 \leftrightarrow 1$ and $1 \leftrightarrow 2$, and by transitivity $0 \leftrightarrow 2$ as well. Thus, the Markov chain is irreducible.

Example 4.15 (cont.)

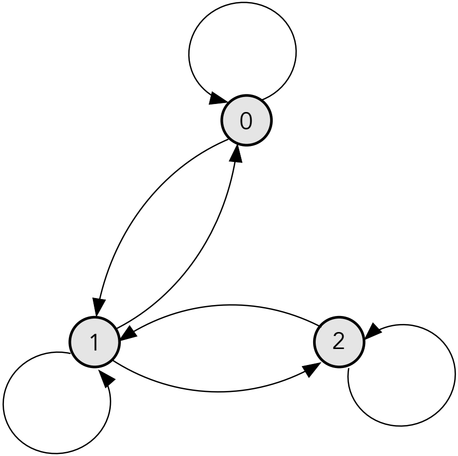


Figure: State diagram of an irreducible Markov chain with one class $\{0, 1, 2\}$

Example 4.16

A Markov chain with state space $\mathcal{S} = \{0, 1, 2, 3\}$ and matrix:

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{aligned} P_{01} = P_{10} &= \tfrac{1}{2}, \text{ implying that } 0 \leftrightarrow 1 \\ P_{ij} &= 0, \text{ implying that } i \not\to j, \quad i = 0, 1, \quad j = 2, 3 \\ P_{2i} &= \tfrac{1}{4}, \text{ implying that } 2 \to i, \quad i = 0, 1, 2, 3 \\ P_{3i} &= 0, \text{ implying that } 3 \not\to i, \quad i = 0, 1, 2 \end{aligned}$$

The Markov chain has classes $\{0, 1\}$, $\{2\}$ and $\{3\}$, and is not irreducible.

Example 4.16 (cont.)

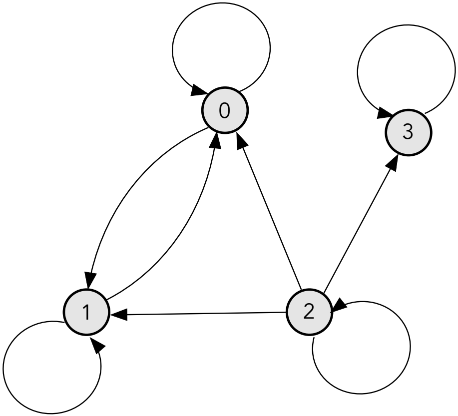


Figure: State diagram of a Markov chain with three classes $\{0, 1\}$, $\{2\}$ and $\{3\}$.

Recurrent and transient states

We consider the probabilities:

$$f_i = P(\bigcup_{r=1}^{\infty}\{X_r = i\}|X_0 = i\}, \quad i \in \mathcal{S}.$$

- State i is recurrent if $f_i = 1$.
- State i is transient if $f_i < 1$.

Assume that $X_0 = i$, and let N_i denote the number of times $X_n = i$.

- If i is recurrent, then $P(N_i = \infty) = 1$.
- If i is transient, then $P(N_i = n) = f_i^{n-1}(1 - f_i)$, $n = 1, 2, \ldots$.

If i is transient and $X_0 = i$, then N_i has a geometric distribution with $E[N_i] = 1/(1 - f_i)$.

Proposition 4.1

Let $I_i^{(n)} = I(X_n = i)$, $n = 0, 1, \ldots$. We can then write:

$$N_i = \sum_{n=0}^{\infty} I_i^{(n)}$$

Hence, we have:

$$\begin{aligned} E[N_i | X_0 = i] &= \sum_{n=0}^{\infty} E[I_i^{(n)} | X_0 = i] \\ &= \sum_{n=0}^{\infty} P[X_n = i | X_0 = i] = \sum_{n=0}^{\infty} P_{ii}^n \end{aligned}$$

- State i is recurrent, if $\sum_{n=1}^{\infty} P_{ii}^n = \infty$.
- State i is transient, if $\sum_{n=1}^{\infty} P_{ii}^n < \infty$.

Corollary 4.2

If state i is recurrent, and $i \leftrightarrow j$, then state j is recurrent as well. Thus, recurrence is a class property.

PROOF: Since $i \leftrightarrow j$, there exists k and m such that $P_{ij}^k > 0$ and $P_{ji}^m > 0$.

Hence, for any $n = 1, 2, \ldots$ we have:

$$P_{jj}^{m+n+k} \geq P_{ji}^m \cdot P_{ii}^n \cdot P_{ij}^k.$$

Summing over all n, and using that i is recurrent, $P_{ij}^k > 0$ and $P_{ji}^m > 0$ we get:

$$\sum_{n=1}^{\infty} P_{jj}^{m+n+k} \geq P_{ji}^m \cdot P_{ij}^k \cdot \sum_{n=1}^{\infty} P_{ii}^n = \infty$$

Hence, we conclude that j is recurrent as well ■

Corollary 4.2 (cont.)

• Corollary 4.2 also implies that transience is a class property. For if state i is transient and $i \leftrightarrow j$, then state j must also be transient. For if j were recurrent then, by Corollary 4.2, i would also be recurrent and hence could not be transient.

• Corollary 4.2 along with the fact that not all states in a finite Markov chain can be transient leads to the conclusion that all states of a finite irreducible Markov chain are recurrent.

Example 4.17

Consider a Markov chain with state space $\mathcal{S} = \{0, 1, 2, 3\}$ and transition probability matrix:

$$\boldsymbol{P} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

It is easy to verify that $i \leftrightarrow j$ for all $i, j \in \mathcal{S}$. Hence, the Markov chain is irreducible and thus all states must be recurrent ■

Example 4.17 (cont.)

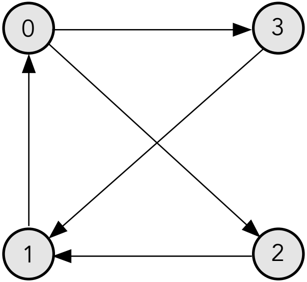


Figure: State diagram of an irreducible Markov chain with one class $\{0, 1, 2, 3\}$

Example 4.18

Consider a Markov chain with state space $\mathcal{S} = \{0, 1, 2, 3, 4\}$ and transition probability matrix:

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

This chain has classes $\{0, 1\}$, $\{2, 3\}$ and $\{4\}$.

The first two classes are recurrent and the third transient ■

Example 4.18 (cont.)

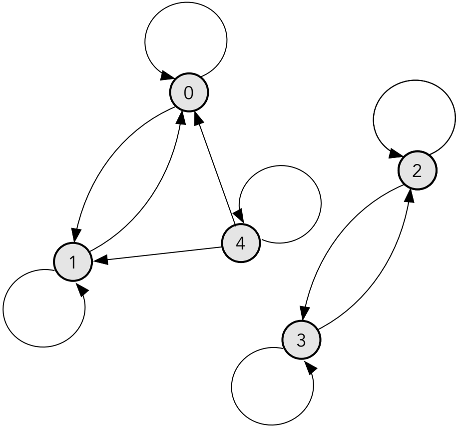


Figure: State diagram of a Markov chain with classes $\{0, 1\}$, $\{2, 3\}$ and $\{4\}$

Example 4.19 - Random walk

Consider a Markov chain with state space $\mathcal{S} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ and where $0 < p < 1$ and:

$$P_{i,i+1} = p, \quad P_{i,i-1} = (1-p), \quad i \in \mathcal{S}.$$

It is obvious that $i \leftrightarrow j$ for all $i, j \in \mathcal{S}$. Hence, according to Corollary 4.2 all states are either recurrent or transient.

In order to check for recurrence, it is sufficient to check if $\sum_{n=1}^{\infty} P_{00}^n = \infty$.

We then observe that X_n is odd if n is odd, and X_n is even if n is even. Hence, since 0 is even, we have:

$$P_{00}^{2n-1} = 0, \quad n = 1, 2, \dots$$

$$P_{00}^{2n} = \binom{2n}{n} p^n (1-p)^n = \frac{(2n)!}{n!n!}[p(1-p)]^n, \quad n = 1, 2, \dots$$

Example 4.19 - Random walk (cont.)

We then use Stirling's formula for $n!$:

$$n! \approx n^{n+1/2} e^{-n} \sqrt{2\pi}$$

From this we get:

$$\frac{(2n)!}{n!n!} \approx \frac{(2n)^{2n+1/2} e^{-2n} \sqrt{2\pi}}{(n^{n+1/2} e^{-n} \sqrt{2\pi})^2} = \frac{(2n)^{2n+1/2} e^{-2n} \sqrt{2\pi}}{n^{2n+1} e^{-2n} (2\pi)} = \frac{2^{2n}}{\sqrt{n\pi}} = \frac{4^n}{\sqrt{n\pi}}$$

Hence:

$$P_{00}^{2n} = \frac{(2n)!}{n!n!} [p(1-p)]^n \approx \frac{(4p(1-p))^n}{\sqrt{n\pi}}$$

Example 4.19 - Random walk (cont.)

This implies that:

$$\sum_{n=1}^{\infty} P_{00}^{2n} \approx \sum_{n=1}^{\infty} \frac{(4p(1-p))^n}{\sqrt{n\pi}}$$

This series is divergent if and only if $p = \frac{1}{2}$.

Hence, the states are recurrent if and only if $p = \frac{1}{2}$.