Chapter 2: Markov Chains

L. Breuer
University of Kent, UK

October 3, 2010

Theorem 2.6

Let $\mathcal{X}$ denote a homogeneous Markov chain with transition matrix P.

Theorem 2.6

Let $\mathcal{X}$ denote a homogeneous Markov chain with transition matrix P. Then the relation

$$\mathbb{P}(X_{n+m} = j | X_n = i) = P^m(i, j)$$

holds for all $m, n \in \mathbb{N}_0$ and $i, j \in E$,

Theorem 2.6

Let $\mathcal{X}$ denote a homogeneous Markov chain with transition matrix P. Then the relation

$$\mathbb{P}(X_{n+m} = j | X_n = i) = P^m(i,j)$$

holds for all $m, n \in \mathbb{N}_0$ and $i, j \in E$, with $P^m(i,j)$ denoting the (i,j)th entry of the mth power of the matrix P.

Theorem 2.6

Let $\mathcal{X}$ denote a homogeneous Markov chain with transition matrix P. Then the relation

$$\mathbb{P}(X_{n+m} = j | X_n = i) = P^m(i,j)$$

holds for all $m, n \in \mathbb{N}_0$ and $i, j \in E$, with $P^m(i,j)$ denoting the (i,j)th entry of the mth power of the matrix P. In particular, P^0 equals the identity matrix.

Theorem 2.6

Let $\mathcal{X}$ denote a homogeneous Markov chain with transition matrix P. Then the relation

$$\mathbb{P}(X_{n+m} = j | X_n = i) = P^m(i,j)$$

holds for all $m, n \in \mathbb{N}_0$ and $i, j \in E$, with $P^m(i,j)$ denoting the (i,j)th entry of the mth power of the matrix P. In particular, P^0 equals the identity matrix.

Proof:

This follows by induction on m.

Theorem 2.6

Let $\mathcal{X}$ denote a homogeneous Markov chain with transition matrix P. Then the relation

$$\mathbb{P}(X_{n+m} = j | X_n = i) = P^m(i,j)$$

holds for all $m, n \in \mathbb{N}_0$ and $i, j \in E$, with $P^m(i,j)$ denoting the (i,j)th entry of the mth power of the matrix P. In particular, P^0 equals the identity matrix.

Proof:

This follows by induction on m. For $m = 1$ the statement holds by definition of P.

Theorem 2.6

Let $\mathcal{X}$ denote a homogeneous Markov chain with transition matrix P. Then the relation

$$\mathbb{P}(X_{n+m} = j | X_n = i) = P^m(i,j)$$

holds for all $m, n \in \mathbb{N}_0$ and $i, j \in E$, with $P^m(i,j)$ denoting the (i,j)th entry of the mth power of the matrix P. In particular, P^0 equals the identity matrix.
Proof:
This follows by induction on m. For $m = 1$ the statement holds by definition of P. For $m > 1$ we can write

Proof of Theorem 2.6

$$\mathbb{P}(X_{n+m} = j | X_n = i) = \frac{\mathbb{P}\left(X_{n+m} = j, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

Proof of Theorem 2.6

$$\mathbb{P}(X_{n+m} = j | X_n = i) = \frac{\mathbb{P}\left(X_{n+m} = j, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

$$= \sum_{k \in E} \frac{\mathbb{P}\left(X_{n+m} = j, X_{n+m-1} = k, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

Proof of Theorem 2.6

$$\mathbb{P}(X_{n+m} = j | X_n = i) = \frac{\mathbb{P}(X_{n+m} = j, X_n = i)}{\mathbb{P}(X_n = i)}$$

$$= \sum_{k \in E} \frac{\mathbb{P}(X_{n+m} = j, X_{n+m-1} = k, X_n = i)}{\mathbb{P}(X_n = i)}$$

$$= \sum_{k \in E} \frac{\mathbb{P}(X_{n+m} = j, X_{n+m-1} = k, X_n = i)}{\mathbb{P}(X_{n+m-1} = k, X_n = i)} \frac{\mathbb{P}(X_{n+m-1} = k, X_n = i)}{\mathbb{P}(X_n = i)}$$

Proof of Theorem 2.6

$$\mathbb{P}(X_{n+m} = j | X_n = i) = \frac{\mathbb{P}\left(X_{n+m} = j, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

$$= \sum_{k \in E} \frac{\mathbb{P}\left(X_{n+m} = j, X_{n+m-1} = k, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

$$= \sum_{k \in E} \frac{\mathbb{P}\left(X_{n+m} = j, X_{n+m-1} = k, X_n = i\right)}{\mathbb{P}\left(X_{n+m-1} = k, X_n = i\right)} \frac{\mathbb{P}\left(X_{n+m-1} = k, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

$$= \sum_{k \in E} \mathbb{P}\left(X_{n+m} = j | X_{n+m-1} = k, X_n = i\right) \cdot P^{m-1}(i, k)$$

Proof of Theorem 2.6

$$\mathbb{P}(X_{n+m} = j | X_n = i) = \frac{\mathbb{P}(X_{n+m} = j, X_n = i)}{\mathbb{P}(X_n = i)}$$

$$= \sum_{k \in E} \frac{\mathbb{P}(X_{n+m} = j, X_{n+m-1} = k, X_n = i)}{\mathbb{P}(X_n = i)}$$

$$= \sum_{k \in E} \frac{\mathbb{P}(X_{n+m} = j, X_{n+m-1} = k, X_n = i)}{\mathbb{P}(X_{n+m-1} = k, X_n = i)} \frac{\mathbb{P}(X_{n+m-1} = k, X_n = i)}{\mathbb{P}(X_n = i)}$$

$$= \sum_{k \in E} \mathbb{P}(X_{n+m} = j | X_{n+m-1} = k, X_n = i) \cdot P^{m-1}(i,k)$$

$$= \sum_{k \in E} p_{kj} \cdot P^{m-1}(i,k) = P^m(i,j)$$

Chapter 14: Chapman–Kolmogorov equations

Thus the probabilities for transitions in m steps are given by the mth power of the transition matrix P.

Chapman–Kolmogorov equations

Thus the probabilities for transitions in m steps are given by the mth power of the transition matrix P. The rule $P^{m+n} = P^m P^n$ for the multiplication of matrices

Chapman–Kolmogorov equations

Thus the probabilities for transitions in m steps are given by the mth power of the transition matrix P. The rule $P^{m+n} = P^m P^n$ for the multiplication of matrices and theorem 2.6 lead to the decompositions

$$\mathbb{P}(X_{m+n} = j | X_0 = i) = \sum_{k \in E} \mathbb{P}(X_m = k | X_0 = i) \cdot \mathbb{P}(X_n = j | X_0 = k)$$

Chapman–Kolmogorov equations

Thus the probabilities for transitions in m steps are given by the mth power of the transition matrix P. The rule $P^{m+n} = P^m P^n$ for the multiplication of matrices and theorem 2.6 lead to the decompositions

$$\mathbb{P}(X_{m+n} = j | X_0 = i) = \sum_{k \in E} \mathbb{P}(X_m = k | X_0 = i) \cdot \mathbb{P}(X_n = j | X_0 = k)$$

which are known as the **Chapman–Kolmogorov equations**.

Stopping times

Let τ denote a random variable with values in $\mathbb{N}_0 \cup \{\infty\}$,

Stopping times

Let τ denote a random variable with values in $\mathbb{N}_0 \cup \{\infty\}$, such that the condition

$$\mathbb{P}(\tau \leq n|\mathcal{X}) = \mathbb{P}(\tau \leq n|X_0, \dots, X_n)$$

holds for all $n \in \mathbb{N}_0$.

Stopping times

Let τ denote a random variable with values in $\mathbb{N}_0 \cup \{\infty\}$, such that the condition

$$\mathbb{P}(\tau \leq n | \mathcal{X}) = \mathbb{P}(\tau \leq n | X_0, \ldots, X_n)$$

holds for all $n \in \mathbb{N}_0$. Such a random variable is called a (discrete) **stopping time** for $\mathcal{X}$.

Theorem 2.7: Strong Markov property

Let $\mathcal{X}$ denote a Markov chain and τ a stopping time for $\mathcal{X}$ with $\mathbb{P}(\tau < \infty) = 1$.

Theorem 2.7: Strong Markov property

Let $\mathcal{X}$ denote a Markov chain and τ a stopping time for $\mathcal{X}$ with $\mathbb{P}(\tau < \infty) = 1$. Then the relation

$$\mathbb{P}(X_{\tau+m} = j | X_0 = i_0, \ldots, X_\tau = i_\tau) = \mathbb{P}(X_m = j | X_0 = i_\tau)$$

holds for all $m \in \mathbb{N}$ and $i_0, \ldots, i_\tau, j \in E$.

Proof of theorem 2.7

$$\mathbb{P}(X_{\tau+m} = j | X_0 = i_0, \dots, X_\tau = i_\tau)$$

Proof of theorem 2.7

$$\mathbb{P}(X_{\tau+m} = j | X_0 = i_0, \dots, X_\tau = i_\tau)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n, X_{\tau+m} = j | X_0 = i_0, \dots, X_n = i_n)$$

Proof of theorem 2.7

$$\mathbb{P}(X_{\tau+m} = j | X_0 = i_0, \dots, X_\tau = i_\tau)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n, X_{\tau+m} = j | X_0 = i_0, \dots, X_n = i_n)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(X_{n+m} = j | \tau = n, X_0 = i_0, \dots, X_n = i_n)$$

$$\times \mathbb{P}(\tau = n | X_0 = i_0, \dots, X_n = i_n)$$

Proof of theorem 2.7

$$\mathbb{P}(X_{\tau+m} = j | X_0 = i_0, \dots, X_\tau = i_\tau)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n, X_{\tau+m} = j | X_0 = i_0, \dots, X_n = i_n)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(X_{n+m} = j | \tau = n, X_0 = i_0, \dots, X_n = i_n)$$

$$\times \mathbb{P}(\tau = n | X_0 = i_0, \dots, X_n = i_n)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(X_{n+m} = j | \tau = n, X_n = i_n) \cdot \mathbb{P}(\tau = n | \mathcal{X})$$

Proof of theorem 2.7

$$\mathbb{P}(X_{\tau+m} = j | X_0 = i_0, \dots, X_\tau = i_\tau)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n, X_{\tau+m} = j | X_0 = i_0, \dots, X_n = i_n)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(X_{n+m} = j | \tau = n, X_0 = i_0, \dots, X_n = i_n)$$

$$\times \mathbb{P}(\tau = n | X_0 = i_0, \dots, X_n = i_n)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(X_{n+m} = j | \tau = n, X_n = i_n) \cdot \mathbb{P}(\tau = n | \mathcal{X})$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(\tau = n | \mathcal{X}) \cdot \mathbb{P}(X_m = j | X_0 = i_\tau)$$

Classification of states

Let $\mathcal{X}$ denote a Markov chain with state space E and transition matrix P.

Classification of states

Let $\mathcal{X}$ denote a Markov chain with state space E and transition matrix P. We call a state $j \in E$ **accessible** from a state $i \in E$

Classification of states

Let $\mathcal{X}$ denote a Markov chain with state space E and transition matrix P. We call a state $j \in E$ **accessible** from a state $i \in E$ if there is a number $m \in \mathbb{N}_0$ with $P(X_m = j | X_0 = i) > 0$.

Classification of states

Let $\mathcal{X}$ denote a Markov chain with state space E and transition matrix P. We call a state $j \in E$ **accessible** from a state $i \in E$ if there is a number $m \in \mathbb{N}_0$ with $P(X_m = j | X_0 = i) > 0$. This relation shall be denoted by $i \to j$.

Let $\mathcal{X}$ denote a Markov chain with state space E and transition matrix P. We call a state $j \in E$ **accessible** from a state $i \in E$ if there is a number $m \in \mathbb{N}_0$ with $P(X_m = j | X_0 = i) > 0$. This relation shall be denoted by $i \to j$. If for two states $i, j \in E$ the relations $i \to j$ and $j \to i$ hold,

Classification of states

Let $\mathcal{X}$ denote a Markov chain with state space E and transition matrix P. We call a state $j \in E$ **accessible** from a state $i \in E$ if there is a number $m \in \mathbb{N}_0$ with $P(X_m = j | X_0 = i) > 0$. This relation shall be denoted by $i \to j$. If for two states $i, j \in E$ the relations $i \to j$ and $j \to i$ hold, then i and j are said to **communicate**,

Classification of states

Let $\mathcal{X}$ denote a Markov chain with state space E and transition matrix P. We call a state $j \in E$ **accessible** from a state $i \in E$ if there is a number $m \in \mathbb{N}_0$ with $P(X_m = j | X_0 = i) > 0$. This relation shall be denoted by $i \to j$. If for two states $i, j \in E$ the relations $i \to j$ and $j \to i$ hold, then i and j are said to **communicate**, in notation $i \leftrightarrow j$.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.
Proof:
Because of $P^0 = I$, communication is reflexive.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition. Thus it remains to show transitivity.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition. Thus it remains to show transitivity. For this, assume $i \leftrightarrow j$ and $j \leftrightarrow k$ for three states $i, j, k \in E$.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition. Thus it remains to show transitivity. For this, assume $i \leftrightarrow j$ and $j \leftrightarrow k$ for three states $i, j, k \in E$. This means that there are numbers $m, n \in \mathbb{N}_0$ with $P^m(i, j) > 0$ and $P^n(j, k) > 0$.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition. Thus it remains to show transitivity. For this, assume $i \leftrightarrow j$ and $j \leftrightarrow k$ for three states $i, j, k \in E$. This means that there are numbers $m, n \in \mathbb{N}_0$ with $P^m(i, j) > 0$ and $P^n(j, k) > 0$. Hence, by the Chapman–Kolmogorov equation, we obtain

$$\mathbb{P}(X_{m+n} = k | X_0 = i) = \sum_{h \in E} \mathbb{P}(X_m = h | X_0 = i) \cdot \mathbb{P}(X_n = k | X_0 = h)$$

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition. Thus it remains to show transitivity. For this, assume $i \leftrightarrow j$ and $j \leftrightarrow k$ for three states $i, j, k \in E$. This means that there are numbers $m, n \in \mathbb{N}_0$ with $P^m(i,j) > 0$ and $P^n(j,k) > 0$. Hence, by the Chapman–Kolmogorov equation, we obtain

$$\mathbb{P}(X_{m+n} = k | X_0 = i) = \sum_{h \in E} \mathbb{P}(X_m = h | X_0 = i) \cdot \mathbb{P}(X_n = k | X_0 = h)$$

$$\geq \mathbb{P}(X_m = j | X_0 = i) \cdot \mathbb{P}(X_n = k | X_0 = j) > 0$$

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition. Thus it remains to show transitivity. For this, assume $i \leftrightarrow j$ and $j \leftrightarrow k$ for three states $i, j, k \in E$. This means that there are numbers $m, n \in \mathbb{N}_0$ with $P^m(i,j) > 0$ and $P^n(j,k) > 0$. Hence, by the Chapman–Kolmogorov equation, we obtain

$$\mathbb{P}(X_{m+n} = k | X_0 = i) = \sum_{h \in E} \mathbb{P}(X_m = h | X_0 = i) \cdot \mathbb{P}(X_n = k | X_0 = h)$$

$$\geq \mathbb{P}(X_m = j | X_0 = i) \cdot \mathbb{P}(X_n = k | X_0 = j) > 0$$

which proves $i \to k$.

Theorem 2.8

The relation $\leftrightarrow$ of communication between states is an equivalence relation.

Proof:

Because of $P^0 = I$, communication is reflexive. Symmetry holds by definition. Thus it remains to show transitivity. For this, assume $i \leftrightarrow j$ and $j \leftrightarrow k$ for three states $i, j, k \in E$. This means that there are numbers $m, n \in \mathbb{N}_0$ with $P^m(i, j) > 0$ and $P^n(j, k) > 0$. Hence, by the Chapman–Kolmogorov equation, we obtain

$$\mathbb{P}(X_{m+n} = k | X_0 = i) = \sum_{h \in E} \mathbb{P}(X_m = h | X_0 = i) \cdot \mathbb{P}(X_n = k | X_0 = h)$$

$$\geq \mathbb{P}(X_m = j | X_0 = i) \cdot \mathbb{P}(X_n = k | X_0 = j) > 0$$

which proves $i \to k$. The remaining proof of $k \to i$ is completely analogous.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**. A communication class $C \subset E$ that does not allow access to states outside itself, i.e.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**. A communication class $C \subset E$ that does not allow access to states outside itself, i.e. for which the implication

$$i \to j, \quad i \in C \qquad \Rightarrow \qquad j \in C$$

holds,

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**. A communication class $C \subset E$ that does not allow access to states outside itself, i.e. for which the implication

$$i \to j, \quad i \in C \qquad \Rightarrow \qquad j \in C$$

holds, is called **closed**.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**. A communication class $C \subset E$ that does not allow access to states outside itself, i.e. for which the implication

$$i \to j, \quad i \in C \qquad \Rightarrow \qquad j \in C$$

holds, is called **closed**. If a closed equivalence class consists only of one state, then this state shall be called **absorbing**.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**. A communication class $C \subset E$ that does not allow access to states outside itself, i.e. for which the implication

$$i \to j, \quad i \in C \qquad \Rightarrow \qquad j \in C$$

holds, is called **closed**. If a closed equivalence class consists only of one state, then this state shall be called **absorbing**. If a Markov chain has only one communication class, i.e.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**. A communication class $C \subset E$ that does not allow access to states outside itself, i.e. for which the implication

$$i \to j, \quad i \in C \qquad \Rightarrow \qquad j \in C$$

holds, is called **closed**. If a closed equivalence class consists only of one state, then this state shall be called **absorbing**. If a Markov chain has only one communication class, i.e. if all states are communicating, then it is called **irreducible**.

Communication classes

Because of this result and the countability, we can divide the state space E of a Markov chain into a partition of countably many equivalence classes with respect to the communication of states. Any such equivalence class shall be called **communication class**. A communication class $C \subset E$ that does not allow access to states outside itself, i.e. for which the implication

$$i \to j, \quad i \in C \qquad \Rightarrow \qquad j \in C$$

holds, is called **closed**. If a closed equivalence class consists only of one state, then this state shall be called **absorbing**. If a Markov chain has only one communication class, i.e. if all states are communicating, then it is called **irreducible**. Otherwise it is called **reducible**.

Examples

Example 2.9: Let $\mathcal{X}$ denote a discrete random walk (see example 2.2) with the specification $\pi_1 = p$ and $\pi_{-1} = 1 - p$ for some parameter $0 < p < 1$.

Examples

Example 2.9: Let $\mathcal{X}$ denote a discrete random walk (see example 2.2) with the specification $\pi_1 = p$ and $\pi_{-1} = 1 - p$ for some parameter $0 < p < 1$. Then $\mathcal{X}$ is irreducible.

Examples

Example 2.9: Let $\mathcal{X}$ denote a discrete random walk (see example 2.2) with the specification $\pi_1 = p$ and $\pi_{-1} = 1 - p$ for some parameter $0 < p < 1$. Then $\mathcal{X}$ is irreducible.

Example 2.10: The Bernoulli process (see example 2.3) with non–trivial parameter $0 < p < 1$ is reducible.

Examples

Example 2.9: Let $\mathcal{X}$ denote a discrete random walk (see example 2.2) with the specification $\pi_1 = p$ and $\pi_{-1} = 1 - p$ for some parameter $0 < p < 1$. Then $\mathcal{X}$ is irreducible.

Example 2.10: The Bernoulli process (see example 2.3) with non–trivial parameter $0 < p < 1$ is reducible. Every state $i \in \mathbb{N}_0$ forms an own communication class.

Examples

Example 2.9: Let $\mathcal{X}$ denote a discrete random walk (see example 2.2) with the specification $\pi_1 = p$ and $\pi_{-1} = 1 - p$ for some parameter $0 < p < 1$. Then $\mathcal{X}$ is irreducible.

Example 2.10: The Bernoulli process (see example 2.3) with non–trivial parameter $0 < p < 1$ is reducible. Every state $i \in \mathbb{N}_0$ forms an own communication class. None of these is closed,

Examples

Example 2.9: Let $\mathcal{X}$ denote a discrete random walk (see example 2.2) with the specification $\pi_1 = p$ and $\pi_{-1} = 1 - p$ for some parameter $0 < p < 1$. Then $\mathcal{X}$ is irreducible.

Example 2.10: The Bernoulli process (see example 2.3) with non–trivial parameter $0 < p < 1$ is reducible. Every state $i \in \mathbb{N}_0$ forms an own communication class. None of these is closed, thus there are no absorbing states.

First visit times

Define τ_j as the stopping time of the **first visit** to the state $j \in E$, i.e.

First visit times

Define τ_j as the stopping time of the **first visit** to the state $j \in E$, i.e.

$$\tau_j := \min\{n \in \mathbb{N} : X_n = j\}$$

First visit times

Define τ_j as the stopping time of the **first visit** to the state $j \in E$, i.e.

$$\tau_j := \min\{n \in \mathbb{N} : X_n = j\}$$

Denote the distribution of τ_j by

$$F_k(i,j) := \mathbb{P}(\tau_j = k | X_0 = i)$$

for all $i, j \in E$ and $k \in \mathbb{N}$.

Lemma 2.12

For all states $i, j \in E$

$$F_k(i,j) = \begin{cases} p_{ij}, & k = 1 \\ \sum_{h \neq j} p_{ih} F_{k-1}(h,j), & k \geq 2 \end{cases}$$

Lemma 2.12

For all states $i, j \in E$

$$F_k(i,j) = \begin{cases} p_{ij}, & k = 1 \\ \sum_{h \neq j} p_{ih} F_{k-1}(h,j), & k \geq 2 \end{cases}$$

Proof:
For $k = 1$, the definition yields

$$F_1(i,j) = \mathbb{P}(\tau_j = 1 | X_0 = i) = \mathbb{P}(X_1 = j | X_0 = i) = p_{ij}$$

for all $i, j \in E$.

Proof of lemma 2.12

For $k \geq 2$, conditioning upon X_1 yields

$$F_k(i,j) = \mathbb{P}(X_1 \neq j, \dots, X_{k-1} \neq j, X_k = j | X_0 = i)$$

Proof of lemma 2.12

For $k \geq 2$, conditioning upon X_1 yields

$$F_k(i,j) = \mathbb{P}(X_1 \neq j, \ldots, X_{k-1} \neq j, X_k = j | X_0 = i)$$

$$= \sum_{h \neq j} \mathbb{P}(X_1 = h | X_0 = i)$$

$$\times \mathbb{P}(X_2 \neq j, \ldots, X_{k-1} \neq j, X_k = j | X_0 = i, X_1 = h)$$

Proof of lemma 2.12

For $k \geq 2$, conditioning upon X_1 yields

$$F_k(i,j) = \mathbb{P}(X_1 \neq j, \ldots, X_{k-1} \neq j, X_k = j | X_0 = i)$$

$$= \sum_{h \neq j} \mathbb{P}(X_1 = h | X_0 = i)$$

$$\times \mathbb{P}(X_2 \neq j, \ldots, X_{k-1} \neq j, X_k = j | X_0 = i, X_1 = h)$$

$$= \sum_{h \neq j} p_{ih} \cdot \mathbb{P}(X_1 \neq j, \ldots, X_{k-2} \neq j, X_{k-1} = j | X_0 = h)$$

Proof of lemma 2.12

For $k \geq 2$, conditioning upon X_1 yields

$$F_k(i,j) = \mathbb{P}(X_1 \neq j, \dots, X_{k-1} \neq j, X_k = j | X_0 = i)$$

$$= \sum_{h \neq j} \mathbb{P}(X_1 = h | X_0 = i)$$
$$\times \mathbb{P}(X_2 \neq j, \dots, X_{k-1} \neq j, X_k = j | X_0 = i, X_1 = h)$$

$$= \sum_{h \neq j} p_{ih} \cdot \mathbb{P}(X_1 \neq j, \dots, X_{k-2} \neq j, X_{k-1} = j | X_0 = h)$$

$$= \sum_{h \neq j} p_{ih} F_{k-1}(h,j)$$

Probability of ever visiting a state

Now define

$$f_{ij} := \mathbb{P}(\tau_j < \infty | X_0 = i) = \sum_{k=1}^{\infty} F_k(i,j)$$

for all $i, j \in E$,

Probability of ever visiting a state

Now define

$$f_{ij} := \mathbb{P}(\tau_j < \infty | X_0 = i) = \sum_{k=1}^{\infty} F_k(i,j)$$

for all $i, j \in E$, which represents the probability of ever visiting state j after beginning in state i.

Probability of ever visiting a state

Now define

$$f_{ij} := \mathbb{P}(\tau_j < \infty | X_0 = i) = \sum_{k=1}^{\infty} F_k(i,j)$$

for all $i, j \in E$, which represents the probability of ever visiting state j after beginning in state i. Summing up over all $k \in \mathbb{N}$ in the formula of Lemma 2.12 leads to

$$f_{ij} = p_{ij} + \sum_{h \neq j} p_{ih} f_{hj}$$

for all $i, j \in E$.

Probability of ever visiting a state

Now define

$$f_{ij} := \mathbb{P}(\tau_j < \infty | X_0 = i) = \sum_{k=1}^{\infty} F_k(i,j)$$

for all $i, j \in E$, which represents the probability of ever visiting state j after beginning in state i. Summing up over all $k \in \mathbb{N}$ in the formula of Lemma 2.12 leads to

$$f_{ij} = p_{ij} + \sum_{h \neq j} p_{ih} f_{hj}$$

for all $i, j \in E$. The proof is left as an exercise.

Total number of visits

Define

$$N_j := \sum_{n=0}^{\infty} \mathbb{I}_{\{X_n = j\}}$$

Total number of visits

Define

$$N_j := \sum_{n=0}^{\infty} \mathbb{I}_{\{X_n = j\}}$$

where $\mathbb{I}$ denotes the indicator function, i.e.

Total number of visits

Define

$$N_j := \sum_{n=0}^{\infty} \mathbb{I}_{\{X_n = j\}}$$

where $\mathbb{I}$ denotes the indicator function, i.e.

$$\mathbb{I}_A := \begin{cases} 1, & \text{if A is true} \\ 0, & \text{if A is false} \end{cases}$$

Total number of visits

Define

$$N_j := \sum_{n=0}^{\infty} \mathbb{I}_{\{X_n=j\}}$$

where $\mathbb{I}$ denotes the indicator function, i.e.

$$\mathbb{I}_A := \begin{cases} 1, & \text{if A is true} \\ 0, & \text{if A is false} \end{cases}$$

This is the random variable of the **total number of visits** to the state $j \in E$.

Theorem 2.13

Let $\mathcal{X}$ denote a Markov chain with state space E. Then

Theorem 2.13

Let $\mathcal{X}$ denote a Markov chain with state space E. Then

$$\mathbb{P}(N_j = m | X_0 = j) = f_{jj}^{m-1}(1 - f_{jj})$$

for $m \in \mathbb{N}$,

Theorem 2.13

Let $\mathcal{X}$ denote a Markov chain with state space E. Then

$$\mathbb{P}(N_j = m | X_0 = j) = f_{jj}^{m-1}(1 - f_{jj})$$

for $m \in \mathbb{N}$, and for $i \neq j$

Theorem 2.13

Let $\mathcal{X}$ denote a Markov chain with state space E. Then

$$\mathbb{P}(N_j = m | X_0 = j) = f_{jj}^{m-1}(1 - f_{jj})$$

for $m \in \mathbb{N}$, and for $i \neq j$

$$\mathbb{P}(N_j = m | X_0 = i) = \begin{cases} 1 - f_{ij}, & m = 0 \\ f_{ij} f_{jj}^{m-1}(1 - f_{jj}), & m \geq 1 \end{cases}$$

Proof of theorem 2.13

Define $\tau_j^{(1)} := \tau_j$

Proof of theorem 2.13

Define $\tau_j^{(1)} := \tau_j$ and $\tau_j^{(k+1)} := \min\{n > \tau_j^{(k)} : X_n = j\}$ for all $k \in \mathbb{N}$,

Proof of theorem 2.13

Define $\tau_j^{(1)} := \tau_j$ and $\tau_j^{(k+1)} := \min\{n > \tau_j^{(k)} : X_n = j\}$ for all $k \in \mathbb{N}$, with the convention that $\min \emptyset = \infty$.

Proof of theorem 2.13

Define $\tau_j^{(1)} := \tau_j$ and $\tau_j^{(k+1)} := \min\{n > \tau_j^{(k)} : X_n = j\}$ for all $k \in \mathbb{N}$, with the convention that $\min \emptyset = \infty$. Note that $\tau_j^{(k)} = \infty$ implies $\tau_j^{(l)} = \infty$ for all $l > k$.

Proof of theorem 2.13

Define $\tau_j^{(1)} := \tau_j$ and $\tau_j^{(k+1)} := \min\{n > \tau_j^{(k)} : X_n = j\}$ for all $k \in \mathbb{N}$, with the convention that $\min \emptyset = \infty$. Note that $\tau_j^{(k)} = \infty$ implies $\tau_j^{(l)} = \infty$ for all $l > k$.

Then the sequence $(\tau_j^{(k)} : k \in \mathbb{N})$ is a sequence of stopping times.

Proof of theorem 2.13

Define $\tau_j^{(1)} := \tau_j$ and $\tau_j^{(k+1)} := \min\{n > \tau_j^{(k)} : X_n = j\}$ for all $k \in \mathbb{N}$, with the convention that $\min \emptyset = \infty$. Note that $\tau_j^{(k)} = \infty$ implies $\tau_j^{(l)} = \infty$ for all $l > k$.

Then the sequence $(\tau_j^{(k)} : k \in \mathbb{N})$ is a sequence of stopping times. Further,

$$\{N_j = m\} = \begin{cases} \bigcap_{k=1}^{m-1} \{\tau_j^{(k)} < \infty\} \cap \{\tau_j^{(m)} = \infty\} & \text{on } \{X_0 = j\} \\ \bigcap_{k=1}^{m} \{\tau_j^{(k)} < \infty\} \cap \{\tau_j^{(m+1)} = \infty\} & \text{on } \{X_0 \neq j\} \end{cases}$$

Proof of theorem 2.13

Define $\tau_j^{(1)} := \tau_j$ and $\tau_j^{(k+1)} := \min\{n > \tau_j^{(k)} : X_n = j\}$ for all $k \in \mathbb{N}$, with the convention that $\min \emptyset = \infty$. Note that $\tau_j^{(k)} = \infty$ implies $\tau_j^{(l)} = \infty$ for all $l > k$.

Then the sequence $(\tau_j^{(k)} : k \in \mathbb{N})$ is a sequence of stopping times. Further,

$$\{N_j = m\} = \begin{cases} \bigcap_{k=1}^{m-1}\{\tau_j^{(k)} < \infty\} \cap \{\tau_j^{(m)} = \infty\} & \text{on } \{X_0 = j\} \\ \bigcap_{k=1}^{m}\{\tau_j^{(k)} < \infty\} \cap \{\tau_j^{(m+1)} = \infty\} & \text{on } \{X_0 \neq j\} \end{cases}$$

$$= \begin{cases} \bigcap_{k=1}^{m-1}\{\tau_j^{(k)} - \tau_j^{(k-1)} < \infty\} \cap \{\tau_j^{(m)} - \tau_j^{(m-1)} = \infty\}, & X_0 = j \\ \bigcap_{k=1}^{m}\{\tau_j^{(k)} - \tau_j^{(k-1)} < \infty\} \cap \{\tau_j^{(m+1)} - \tau_j^{(m)} = \infty\}, & X_0 \neq j \end{cases}$$

with $\tau_j^{(0)} := 0$.

Proof of theorem 2.13 (contd.)

The strong Markov property yields for $X_0 = i$

$$\mathbb{P}(\tau_j^{(k)} - \tau_j^{(k-1)} < \infty) = \begin{cases} f_{ij}, & k = 1 \\ f_{jj}, & k > 1 \end{cases}$$

Proof of theorem 2.13 (contd.)

The strong Markov property yields for $X_0 = i$

$$\mathbb{P}(\tau_j^{(k)} - \tau_j^{(k-1)} < \infty) = \begin{cases} f_{ij}, & k = 1 \\ f_{jj}, & k > 1 \end{cases}$$

and independence of the events $\{\tau_j^{(k)} - \tau_j^{(k-1)} <= \infty\}$, $k = 1, \ldots, m+1$.

Corollary 2.14

Summing over all $m \in \mathbb{N}$ in the above theorem leads to

Corollary 2.14

Summing over all $m \in \mathbb{N}$ in the above theorem leads to

$$\mathbb{P}(N_j < \infty | X_0 = j) = \begin{cases} 1, & f_{jj} < 1 \\ 0, & f_{jj} = 1 \end{cases}$$

Corollary 2.14

Summing over all $m \in \mathbb{N}$ in the above theorem leads to

$$\mathbb{P}(N_j < \infty | X_0 = j) = \begin{cases} 1, & f_{jj} < 1 \\ 0, & f_{jj} = 1 \end{cases}$$

i.e. depending on f_{jj} there are almost surely infinitely many visits to a state $j \in E$.

Recurrence / Transience

This result gives rise to the following definitions:

Recurrence / Transience

This result gives rise to the following definitions: A state $j \in E$ is called **recurrent** if $f_{jj} = 1$

Recurrence / Transience

This result gives rise to the following definitions: A state $j \in E$ is called **recurrent** if $f_{jj} = 1$ and **transient** otherwise.

Recurrence / Transience

This result gives rise to the following definitions: A state $j \in E$ is called **recurrent** if $f_{jj} = 1$ and **transient** otherwise. Further define the **potential matrix** $R = (r_{ij})_{i,j \in E}$ of the Markov chain

Recurrence / Transience

This result gives rise to the following definitions: A state $j \in E$ is called **recurrent** if $f_{jj} = 1$ and **transient** otherwise. Further define the **potential matrix** $R = (r_{ij})_{i,j \in E}$ of the Markov chain by its entries

$$r_{ij} := \mathbb{E}(N_j | X_0 = i) = \sum_{n=0}^{\infty} P^n(i,j)$$

for all $i, j \in E$.

Corollary 2.15

For all $i, j \in E$ the relations

$$r_{jj} = (1 - f_{jj})^{-1} \qquad \text{and} \qquad r_{ij} = f_{ij} r_{jj}$$

hold,

Corollary 2.15

For all $i, j \in E$ the relations

$$r_{jj} = (1 - f_{jj})^{-1} \qquad \text{and} \qquad r_{ij} = f_{ij} r_{jj}$$

hold, with the conventions $0^{-1} := \infty$ and $0 \cdot \infty := 0$ included.

Corollary 2.15

For all $i, j \in E$ the relations

$$r_{jj} = (1 - f_{jj})^{-1} \qquad \text{and} \qquad r_{ij} = f_{ij} r_{jj}$$

hold, with the conventions $0^{-1} := \infty$ and $0 \cdot \infty := 0$ included. In particular, the expected number r_{jj} of visits to the state $j \in E$ is finite if j is transient

Corollary 2.15

For all $i, j \in E$ the relations

$$r_{jj} = (1 - f_{jj})^{-1} \qquad \text{and} \qquad r_{ij} = f_{ij} r_{jj}$$

hold, with the conventions $0^{-1} := \infty$ and $0 \cdot \infty := 0$ included. In particular, the expected number r_{jj} of visits to the state $j \in E$ is finite if j is transient and infinite if j is recurrent.

Theorem 2.16

Recurrence and transience of states are class properties with respect to the relation $\leftrightarrow$.

Theorem 2.16

Recurrence and transience of states are class properties with respect to the relation $\leftrightarrow$. Furthermore, a recurrent communication class is always closed.

Theorem 2.16

Recurrence and transience of states are class properties with respect to the relation $\leftrightarrow$. Furthermore, a recurrent communication class is always closed.

Proof:
Assume that $i \in E$ is transient and $i \leftrightarrow j$.

Theorem 2.16

Recurrence and transience of states are class properties with respect to the relation $\leftrightarrow$. Furthermore, a recurrent communication class is always closed.

Proof:
Assume that $i \in E$ is transient and $i \leftrightarrow j$. Then there are numbers $m, n \in \mathbb{N}$ with $0 < P^m(i,j) \leq 1$ and $0 < P^n(j,i) \leq 1$.

Theorem 2.16

Recurrence and transience of states are class properties with respect to the relation $\leftrightarrow$. Furthermore, a recurrent communication class is always closed.

Proof:
Assume that $i \in E$ is transient and $i \leftrightarrow j$. Then there are numbers $m, n \in \mathbb{N}$ with $0 < P^m(i,j) \leq 1$ and $0 < P^n(j,i) \leq 1$. The inequalities

$$\sum_{k=0}^{\infty} P^k(i,i) \geq \sum_{h=0}^{\infty} P^{m+h+n}(i,i) \geq P^m(i,j)P^n(j,i) \sum_{k=0}^{\infty} P^k(j,j)$$

now imply $r_{jj} < \infty$.

Theorem 2.16

Recurrence and transience of states are class properties with respect to the relation $\leftrightarrow$. Furthermore, a recurrent communication class is always closed.

Proof:
Assume that $i \in E$ is transient and $i \leftrightarrow j$. Then there are numbers $m, n \in \mathbb{N}$ with $0 < P^m(i,j) \leq 1$ and $0 < P^n(j,i) \leq 1$. The inequalities

$$\sum_{k=0}^{\infty} P^k(i,i) \geq \sum_{h=0}^{\infty} P^{m+h+n}(i,i) \geq P^m(i,j)P^n(j,i) \sum_{k=0}^{\infty} P^k(j,j)$$

now imply $r_{jj} < \infty$. According to corollary 2.15 this means that j is transient, too.

Proof of theorem 2.16 (contd.)

If j is recurrent,

Proof of theorem 2.16 (contd.)

If j is recurrent, then the same inequalities lead to

$$r_{ii} \geq P^m(i,j)P^n(j,i)r_{jj} = \infty$$

Proof of theorem 2.16 (contd.)

If j is recurrent, then the same inequalities lead to

$$r_{ii} \geq P^m(i,j)P^n(j,i)r_{jj} = \infty$$

which signifies that i is recurrent, too.

Proof of theorem 2.16 (contd.)

If j is recurrent, then the same inequalities lead to

$$r_{ii} \geq P^m(i,j)P^n(j,i)r_{jj} = \infty$$

which signifies that i is recurrent, too. Since the above arguments are symmetric in i and j, the proof of the first statement is complete.

Proof of theorem 2.16 (contd.)

If j is recurrent, then the same inequalities lead to

$$r_{ii} \geq P^m(i,j)P^n(j,i)r_{jj} = \infty$$

which signifies that i is recurrent, too. Since the above arguments are symmetric in i and j, the proof of the first statement is complete.

For the second statement

Proof of theorem 2.16 (contd.)

If j is recurrent, then the same inequalities lead to

$$r_{ii} \geq P^m(i,j)P^n(j,i)r_{jj} = \infty$$

which signifies that i is recurrent, too. Since the above arguments are symmetric in i and j, the proof of the first statement is complete.

For the second statement assume that $i \in E$ belongs to a communication class $C \subset E$ and $p_{ij} > 0$ for some state $j \in E \setminus C$.

Proof of theorem 2.16 (contd.)

If j is recurrent, then the same inequalities lead to

$$r_{ii} \geq P^m(i,j)P^n(j,i)r_{jj} = \infty$$

which signifies that i is recurrent, too. Since the above arguments are symmetric in i and j, the proof of the first statement is complete.

For the second statement assume that $i \in E$ belongs to a communication class $C \subset E$ and $p_{ij} > 0$ for some state $j \in E \setminus C$. Then

$$f_{ii} = p_{ii} + \sum_{h \neq i} p_{ih} f_{hi} \leq 1 - p_{ij} < 1$$

since $f_{ji} = 0$ (otherwise $i \leftrightarrow j$).

Proof of theorem 2.16 (contd.)

If j is recurrent, then the same inequalities lead to

$$r_{ii} \geq P^m(i,j)P^n(j,i)r_{jj} = \infty$$

which signifies that i is recurrent, too. Since the above arguments are symmetric in i and j, the proof of the first statement is complete.

For the second statement assume that $i \in E$ belongs to a communication class $C \subset E$ and $p_{ij} > 0$ for some state $j \in E \setminus C$. Then

$$f_{ii} = p_{ii} + \sum_{h \neq i} p_{ih} f_{hi} \leq 1 - p_{ij} < 1$$

since $f_{ji} = 0$ (otherwise $i \leftrightarrow j$). Thus i is transient, which proves the second statement.

Theorem 2.17

If the state $j \in E$ is transient, then $\lim_{n\to\infty} P^n(i,j) = 0$,

Theorem 2.17

If the state $j \in E$ is transient, then $\lim_{n\to\infty} P^n(i,j) = 0$, regardless of the initial state $i \in E$.

Theorem 2.17

If the state $j \in E$ is transient, then $\lim_{n\to\infty} P^n(i,j) = 0$, regardless of the initial state $i \in E$.

Proof:
If the state j is transient, then

$$r_{jj} = (1 - f_{jj})^{-1} < \infty$$

by corollary 2.15.

Theorem 2.17

If the state $j \in E$ is transient, then $\lim_{n\to\infty} P^n(i,j) = 0$, regardless of the initial state $i \in E$.

Proof:
If the state j is transient, then

$$r_{jj} = (1 - f_{jj})^{-1} < \infty$$

by corollary 2.15. Further, by the same corollary,

$$r_{ij} = f_{ij} r_{jj} < \infty$$

Theorem 2.17

If the state $j \in E$ is transient, then $\lim_{n\to\infty} P^n(i,j) = 0$, regardless of the initial state $i \in E$.

Proof:
If the state j is transient, then

$$r_{jj} = (1 - f_{jj})^{-1} < \infty$$

by corollary 2.15. Further, by the same corollary,

$$r_{ij} = f_{ij} r_{jj} < \infty$$

Since $r_{ij} = \sum_{n=0}^{\infty} P^n(i,j)$, the statement follows.