Chapter 2: Markov Chains

L. Breuer
University of Kent, UK

September 28, 2010

Definition-1

Definition-1

Let X_n with $n \in \mathbb{N}_0$ denote random variables on a discrete space E.

Definition-1

Let X_n with $n \in \mathbb{N}_0$ denote random variables on a discrete space E. The sequence $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is called a **stochastic chain**.

Definition-1

Let X_n with $n \in \mathbb{N}_0$ denote random variables on a discrete space E. The sequence $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is called a **stochastic chain**. If $\mathbb{P}$ is a probability measure for $\mathcal{X}$ such that

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n) \quad (1)$$

for all $i_0, \ldots, i_n, j \in E$ and $n \in \mathbb{N}_0$,

Definition-1

Let X_n with $n \in \mathbb{N}_0$ denote random variables on a discrete space E. The sequence $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is called a **stochastic chain**. If $\mathbb{P}$ is a probability measure for $\mathcal{X}$ such that

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n) \quad (1)$$

for all $i_0, \ldots, i_n, j \in E$ and $n \in \mathbb{N}_0$, then the sequence $\mathcal{X}$ shall be called a **Markov chain** on E.

Definition-1

Let X_n with $n \in \mathbb{N}_0$ denote random variables on a discrete space E. The sequence $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is called a **stochastic chain**. If $\mathbb{P}$ is a probability measure for $\mathcal{X}$ such that

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = j | X_n = i_n) \quad (1)$$

for all $i_0, \ldots, i_n, j \in E$ and $n \in \mathbb{N}_0$, then the sequence $\mathcal{X}$ shall be called a **Markov chain** on E. The probability measure $\mathbb{P}$ is called the distribution of $\mathcal{X}$,

Definition-1

Let X_n with $n \in \mathbb{N}_0$ denote random variables on a discrete space E. The sequence $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is called a **stochastic chain**. If $\mathbb{P}$ is a probability measure for $\mathcal{X}$ such that

$$\mathbb{P}\left(X_{n+1} = j | X_0 = i_0, \ldots, X_n = i_n\right) = \mathbb{P}\left(X_{n+1} = j | X_n = i_n\right) \quad (1)$$

for all $i_0, \ldots, i_n, j \in E$ and $n \in \mathbb{N}_0$, then the sequence $\mathcal{X}$ shall be called a **Markov chain** on E. The probability measure $\mathbb{P}$ is called the distribution of $\mathcal{X}$, and E is called the **state space** of $\mathcal{X}$.

Definition-2

If the conditional probabilities $\mathbb{P}\left(X_{n+1} = j | X_n = i_n\right)$ are independent of the time index $n \in \mathbb{N}_0$,

Definition-2

If the conditional probabilities $\mathbb{P}(X_{n+1} = j | X_n = i_n)$ are independent of the time index $n \in \mathbb{N}_0$, then we call the Markov chain $\mathcal{X}$ **homogeneous**

Definition-2

If the conditional probabilities $\mathbb{P}(X_{n+1} = j | X_n = i_n)$ are independent of the time index $n \in \mathbb{N}_0$, then we call the Markov chain $\mathcal{X}$ **homogeneous** and denote

$$p_{ij} := \mathbb{P}(X_{n+1} = j | X_n = i)$$

for all $i, j \in E$.

Definition-2

If the conditional probabilities $\mathbb{P}(X_{n+1} = j | X_n = i_n)$ are independent of the time index $n \in \mathbb{N}_0$, then we call the Markov chain $\mathcal{X}$ **homogeneous** and denote

$$p_{ij} := \mathbb{P}(X_{n+1} = j | X_n = i)$$

for all $i, j \in E$. The probability p_{ij} is called **transition probability** from state i to state j.

Definition-2

If the conditional probabilities $\mathbb{P}(X_{n+1} = j | X_n = i_n)$ are independent of the time index $n \in \mathbb{N}_0$, then we call the Markov chain $\mathcal{X}$ **homogeneous** and denote

$$p_{ij} := \mathbb{P}(X_{n+1} = j | X_n = i)$$

for all $i, j \in E$. The probability p_{ij} is called **transition probability** from state i to state j. The matrix $P := (p_{ij})_{i,j \in E}$ shall be called **transition matrix** of the chain $\mathcal{X}$.

Definition-2

If the conditional probabilities $\mathbb{P}(X_{n+1} = j | X_n = i_n)$ are independent of the time index $n \in \mathbb{N}_0$, then we call the Markov chain $\mathcal{X}$ **homogeneous** and denote

$$p_{ij} := \mathbb{P}(X_{n+1} = j | X_n = i)$$

for all $i, j \in E$. The probability p_{ij} is called **transition probability** from state i to state j. The matrix $P := (p_{ij})_{i,j \in E}$ shall be called **transition matrix** of the chain $\mathcal{X}$. Condition (1) is referred to as the **Markov property**.

Example 2.1

Example 2.1

If $(X_n : n \in \mathbb{N}_0)$ are random variables on a discrete space E, which are stochastically independent and identically distributed (shortly: iid),

Example 2.1

If $(X_n : n \in \mathbb{N}_0)$ are random variables on a discrete space E, which are stochastically independent and identically distributed (shortly: iid), then the chain $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is a homogeneous Markov chain.

Example 2.2: Discrete Random Walk

Example 2.2: Discrete Random Walk

Set $E := \mathbb{Z}$

Example 2.2: Discrete Random Walk

Set $E := \mathbb{Z}$ and let $(S_n : n \in \mathbb{N})$ be a sequence of iid random variables with values in $\mathbb{Z}$ and distribution π.

Example 2.2: Discrete Random Walk

Set $E := \mathbb{Z}$ and let $(S_n : n \in \mathbb{N})$ be a sequence of iid random variables with values in $\mathbb{Z}$ and distribution π. Define $X_0 := 0$ and $X_n := \sum_{k=1}^{n} S_k$ for all $n \in \mathbb{N}$.

Example 2.2: Discrete Random Walk

Set $E := \mathbb{Z}$ and let $(S_n : n \in \mathbb{N})$ be a sequence of iid random variables with values in $\mathbb{Z}$ and distribution π. Define $X_0 := 0$ and $X_n := \sum_{k=1}^{n} S_k$ for all $n \in \mathbb{N}$. Then the chain $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is a homogeneous Markov chain

Example 2.2: Discrete Random Walk

Set $E := \mathbb{Z}$ and let $(S_n : n \in \mathbb{N})$ be a sequence of iid random variables with values in $\mathbb{Z}$ and distribution π. Define $X_0 := 0$ and $X_n := \sum_{k=1}^{n} S_k$ for all $n \in \mathbb{N}$. Then the chain $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is a homogeneous Markov chain with transition probabilities $p_{ij} = \pi_{j-i}$.

Example 2.2: Discrete Random Walk

Set $E := \mathbb{Z}$ and let $(S_n : n \in \mathbb{N})$ be a sequence of iid random variables with values in $\mathbb{Z}$ and distribution π. Define $X_0 := 0$ and $X_n := \sum_{k=1}^{n} S_k$ for all $n \in \mathbb{N}$. Then the chain $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$ is a homogeneous Markov chain with transition probabilities $p_{ij} = \pi_{j-i}$. This chain is called **discrete random walk**.

Example 2.3: Bernoulli process

Example 2.3: Bernoulli process

Set $E := \mathbb{N}_0$ and choose any parameter $0 < p < 1$.

Example 2.3: Bernoulli process

Set $E := \mathbb{N}_0$ and choose any parameter $0 < p < 1$. The definitions $X_0 := 0$ as well as

$$p_{ij} := \begin{cases} p, & j = i + 1 \\ 1 - p, & j = i \end{cases}$$

for $i \in \mathbb{N}_0$

Example 2.3: Bernoulli process

Set $E := \mathbb{N}_0$ and choose any parameter $0 < p < 1$. The definitions $X_0 := 0$ as well as

$$p_{ij} := \begin{cases} p, & j = i + 1 \\ 1 - p, & j = i \end{cases}$$

for $i \in \mathbb{N}_0$ determine a homogeneous Markov chain $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$.

Example 2.3: Bernoulli process

Set $E := \mathbb{N}_0$ and choose any parameter $0 < p < 1$. The definitions $X_0 := 0$ as well as

$$p_{ij} := \begin{cases} p, & j = i + 1 \\ 1 - p, & j = i \end{cases}$$

for $i \in \mathbb{N}_0$ determine a homogeneous Markov chain $\mathcal{X} = (X_n : n \in \mathbb{N}_0)$. It is called **Bernoulli process** with parameter p.

Stochastic Matrix

Let P denote the transition matrix of a Markov chain on E.

Stochastic Matrix

Let P denote the transition matrix of a Markov chain on E. Then as an immediate consequence of its definition we obtain $p_{ij} \in [0, 1]$ for all $i, j \in E$

Stochastic Matrix

Let P denote the transition matrix of a Markov chain on E. Then as an immediate consequence of its definition we obtain $p_{ij} \in [0,1]$ for all $i,j \in E$ and $\sum_{j \in E} p_{ij} = 1$ for all $i \in E$.

Stochastic Matrix

Let P denote the transition matrix of a Markov chain on E. Then as an immediate consequence of its definition we obtain $p_{ij} \in [0,1]$ for all $i, j \in E$ and $\sum_{j \in E} p_{ij} = 1$ for all $i \in E$. A matrix P with these properties is called a **stochastic matrix** on E.

Theorem 2.4

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

Theorem 2.4

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

$$\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i\right) = p_{i,j_1} \cdot \ldots \cdot p_{j_{m-1},j_m}$$

holds for all $n \in \mathbb{N}_0$, $m \in \mathbb{N}$, and $i, j_1, \ldots, j_m \in E$.

Theorem 2.4

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

$$\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i\right) = p_{i,j_1} \cdot \ldots \cdot p_{j_{m-1},j_m}$$

holds for all $n \in \mathbb{N}_0$, $m \in \mathbb{N}$, and $i, j_1, \ldots, j_m \in E$.
Proof:
This follows by induction on m.

Theorem 2.4

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

$$\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i) = p_{i,j_1} \cdot \ldots \cdot p_{j_{m-1},j_m}$$

holds for all $n \in \mathbb{N}_0$, $m \in \mathbb{N}$, and $i, j_1, \ldots, j_m \in E$.
Proof:
This follows by induction on m. For $m = 1$ the statement holds by definition of P.

Theorem 2.4

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

$$\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i\right) = p_{i,j_1} \cdot \ldots \cdot p_{j_{m-1},j_m}$$

holds for all $n \in \mathbb{N}_0$, $m \in \mathbb{N}$, and $i, j_1, \ldots, j_m \in E$.
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 2.4

$$\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i)$$
$$= \frac{\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m, X_n = i)}{\mathbb{P}(X_n = i)}$$

Proof of theorem 2.4

$$\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i)$$

$$= \frac{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m} = j_m, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

$$= \frac{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m} = j_m, X_n = i\right)}{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1}, X_n = i\right)} \times \frac{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1}, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

Proof of theorem 2.4

$$\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i)$$

$$= \frac{\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m, X_n = i)}{\mathbb{P}(X_n = i)}$$

$$= \frac{\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m, X_n = i)}{\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1}, X_n = i)}$$

$$\times \frac{\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1}, X_n = i)}{\mathbb{P}(X_n = i)}$$

$$= \mathbb{P}(X_{n+m} = j_m | X_n = i, X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1})$$

$$\times p_{i,j_1} \cdot \ldots \cdot p_{j_{m-2},j_{m-1}}$$

Proof of theorem 2.4

$$\mathbb{P}(X_{n+1} = j_1, \ldots, X_{n+m} = j_m | X_n = i)$$

$$= \frac{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m} = j_m, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

$$= \frac{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m} = j_m, X_n = i\right)}{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1}, X_n = i\right)} \times \frac{\mathbb{P}\left(X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1}, X_n = i\right)}{\mathbb{P}\left(X_n = i\right)}$$

$$= \mathbb{P}\left(X_{n+m} = j_m | X_n = i, X_{n+1} = j_1, \ldots, X_{n+m-1} = j_{m-1}\right) \times p_{i,j_1} \cdot \ldots \cdot p_{j_{m-2},j_{m-1}}$$

$$= p_{j_{m-1},j_m} \cdot p_{i,j_1} \cdot \ldots \cdot p_{j_{m-2},j_{m-1}}$$

because of the induction hypothesis and the Markov property.