MS-C2111 Stochastic Processes

Lecture 1:
Markov chains

Jukka Kohonen
Aalto University

Contents

Contents

Markov chain

Random sequence $(X_0, X_1, X_2, \dots)$ with the property:

The future state X_{t+1} is conditionally independent of the past given the current state X_t.

Andrei Markov
1856–1922

Andrei Markov
1978–

Markov chain

Random sequence $(X_0, X_1, X_2, \ldots)$ with the property:

The future state X_{t+1} *is conditionally independent of the past given the current state* X_t.

Mathematically:

Andrei Markov
1856–1922

Andrei Markov
1978–

Markov chain

Random sequence $(X_0, X_1, X_2, \ldots)$ with the property:

The future state X_{t+1} is conditionally independent of the past given the current state X_t.

Mathematically:

$$\mathbb{P}(X_{t+1} = y \mid X_t = x, H_{t-}) \;=\; \mathbb{P}(X_{t+1} = y \mid X_t = x)$$

for all events $H_{t-} = \{X_0 = x_0, \ldots, X_{t-1} = x_{t-1}\}$
and for all states x, y

Andrei Markov
1856–1922

Andrei Markov
1978–

Markov chain

Random sequence $(X_0, X_1, X_2, \ldots)$ with the property:

The future state X_{t+1} is conditionally independent of the past given the current state X_t.

Mathematically:

$$\mathbb{P}(X_{t+1} = y \,|\, X_t = x, H_{t-}) \;=\; \mathbb{P}(X_{t+1} = y \,|\, X_t = x)$$

for all events $H_{t-} = \{X_0 = x_0, \ldots, X_{t-1} = x_{t-1}\}$
and for all states x, y

State space: S = Set of possible states of the chain
(Here assumed finite)

Andrei Markov
1856–1922

Andrei Markov
1978–

Markov chain

Random sequence $(X_0, X_1, X_2, \dots)$ with the property:

The future state X_{t+1} is conditionally independent of the past given the current state X_t.

Mathematically:

$$\mathbb{P}(X_{t+1} = y \,|\, X_t = x, H_{t-}) \; = \; \mathbb{P}(X_{t+1} = y \,|\, X_t = x)$$

for all events $H_{t-} = \{X_0 = x_0, \dots, X_{t-1} = x_{t-1}\}$
and for all states x, y

State space: S = Set of possible states of the chain
(Here assumed finite)

Transition matrix: $P(x, y) = \mathbb{P}(X_{t+1} = y \,|\, X_t = x)$ is the probability to move from state x to state y
(Here assumed constant over time)

Andrei Markov
1856–1922

Andrei Markov
1978–

Example: Weather model

Develop a predictive weather model when we estimate that:

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with $1 =$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with $1 =$ cloudy,

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ...but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with 1 = cloudy, 2 =

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with 1 = cloudy, 2 = sunny

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ...but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with $1 =$ cloudy, $2 =$ sunny

Transition matrix $P =$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with 1 = cloudy, 2 = sunny

$$\text{Transition matrix} \quad P \; = \; \begin{bmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{bmatrix} \; =$$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \dots$

State space $S = \{1, 2\}$ with 1 = cloudy, 2 = sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{bmatrix} = \begin{bmatrix} & \\ & \end{bmatrix}$$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with 1 = cloudy, 2 = sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{bmatrix} = \begin{bmatrix} 0.8 & \\ & \end{bmatrix}$$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with $1 =$ cloudy, $2 =$ sunny

$$\text{Transition matrix} \quad P \;=\; \begin{bmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{bmatrix} \;=\; \begin{bmatrix} 0.8 & 0.2 \\ & \end{bmatrix}$$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with $1 =$ cloudy, $2 =$ sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & \end{bmatrix}$$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with $1 =$ cloudy, $2 =$ sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Example: Weather model

Develop a predictive weather model when we estimate that:

- A cloudy day is followed by a sunny day with probability 0.2
- A sunny day is followed by a cloudy day with probability 0.5
- What happens tomorrow **depends** on today's weather ...
- ... but **does not depend** on anything earlier. (Limited memory)

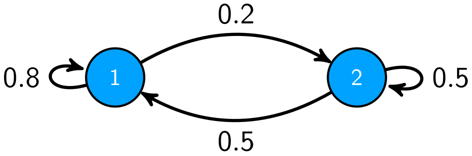
Markov chain:

X_t = Weather on day $t = 0, 1, 2, \ldots$

State space $S = \{1, 2\}$ with $1 =$ cloudy, $2 =$ sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} P(1,1) & P(1,2) \\ P(2,1) & P(2,2) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Transition diagram



Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

By conditioning on X_1 = Tuesday's weather:

$$\mathbb{P}(X_2 = 1 \mid X_0 = 1)$$
$$=$$

Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

By conditioning on X_1 = Tuesday's weather:

$$\begin{aligned} &\mathbb{P}(X_2 = 1 \,|\, X_0 = 1) \\ &\qquad = \mathbb{P}(X_1 = 1 \,|\, X_0 = 1)\mathbb{P}(X_2 = 1 \,|\, X_1 = 1, X_0 = 1) \\ &\qquad\quad + \mathbb{P}(X_1 = 2 \,|\, X_0 = 1)\mathbb{P}(X_2 = 1 \,|\, X_1 = 2, X_0 = 1) \end{aligned}$$

Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

By conditioning on X_1 = Tuesday's weather:

$$\begin{aligned}
&\mathbb{P}(X_2 = 1 \mid X_0 = 1) \\
&\quad = \mathbb{P}(X_1 = 1 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 1, X_0 = 1) \\
&\quad\quad + \mathbb{P}(X_1 = 2 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 2, X_0 = 1) \\
&\quad = P(1,1)P(1,1) + P(1,2)P(2,1)
\end{aligned}$$

Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

By conditioning on X_1 = Tuesday's weather:

$$\begin{aligned}
&\mathbb{P}(X_2 = 1 \mid X_0 = 1) \\
&\qquad = \mathbb{P}(X_1 = 1 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 1, X_0 = 1) \\
&\qquad\quad + \mathbb{P}(X_1 = 2 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 2, X_0 = 1) \\
&\qquad = P(1,1)P(1,1) + P(1,2)P(2,1) \\
&\qquad = 0.8 \times 0.8 + 0.2 \times 0.5
\end{aligned}$$

Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

By conditioning on X_1 = Tuesday's weather:

$$\begin{aligned}
&\mathbb{P}(X_2 = 1 \mid X_0 = 1) \\
&\quad = \mathbb{P}(X_1 = 1 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 1, X_0 = 1) \\
&\qquad + \mathbb{P}(X_1 = 2 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 2, X_0 = 1) \\
&\quad = P(1,1)P(1,1) + P(1,2)P(2,1) \\
&\quad = 0.8 \times 0.8 + 0.2 \times 0.5 \\
&\quad = 0.74
\end{aligned}$$

Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

By conditioning on X_1 = Tuesday's weather:

$$\begin{aligned}
&\mathbb{P}(X_2 = 1 \,|\, X_0 = 1) \\
&\quad = \mathbb{P}(X_1 = 1 \,|\, X_0 = 1)\mathbb{P}(X_2 = 1 \,|\, X_1 = 1, X_0 = 1) \\
&\qquad + \mathbb{P}(X_1 = 2 \,|\, X_0 = 1)\mathbb{P}(X_2 = 1 \,|\, X_1 = 2, X_0 = 1) \\
&\quad = P(1,1)P(1,1) + P(1,2)P(2,1) \\
&\quad = 0.8 \times 0.8 + 0.2 \times 0.5 \\
&\quad = 0.74
\end{aligned}$$

Conclusion: Wed is cloudy with probability 0.74

Predicting using the weather model

If it is cloudy on Mon (day 0), then what is the probability that it is cloudy also on Wed?

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

By conditioning on X_1 = Tuesday's weather:

$$\begin{aligned}
&\mathbb{P}(X_2 = 1 \mid X_0 = 1) \\
&\quad = \mathbb{P}(X_1 = 1 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 1, X_0 = 1) \\
&\qquad + \mathbb{P}(X_1 = 2 \mid X_0 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 2, X_0 = 1) \\
&\quad = P(1,1)P(1,1) + P(1,2)P(2,1) \\
&\quad = 0.8 \times 0.8 + 0.2 \times 0.5 \\
&\quad = 0.74
\end{aligned}$$

Conclusion: Wed is cloudy with probability 0.74

Can you predict Saturday's weather?

Example: Nonbinary weather model

State space $S = \{1, 2, 3\}$ with 1=rainy, 2=cloudy, 3=sunny

Example: Nonbinary weather model

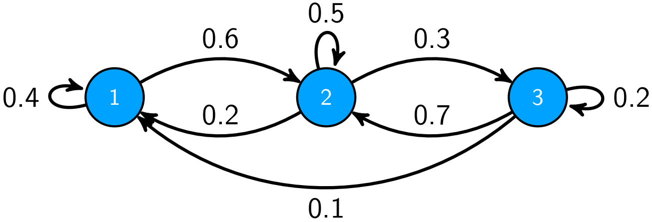
State space $S = \{1, 2, 3\}$ with 1=rainy, 2=cloudy, 3=sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{bmatrix}$$

Example: Nonbinary weather model

State space $S = \{1, 2, 3\}$ with 1=rainy, 2=cloudy, 3=sunny

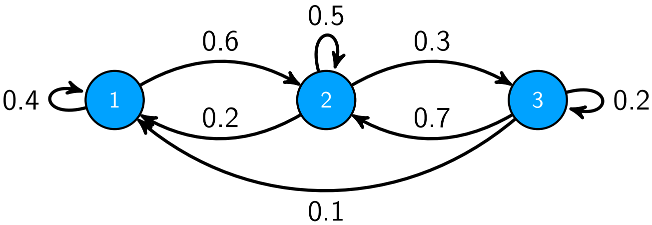
$$\text{Transition matrix} \quad P = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{bmatrix}$$



Example: Nonbinary weather model

State space $S = \{1, 2, 3\}$ with 1=rainy, 2=cloudy, 3=sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{bmatrix}$$

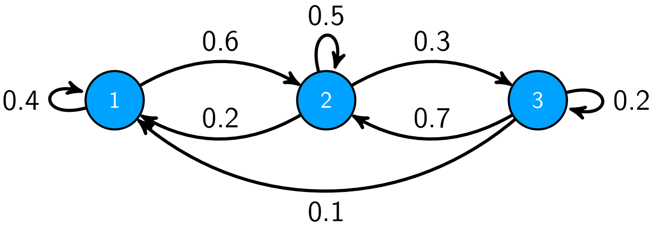


Computing weather predictions manually becomes harder.

Example: Nonbinary weather model

State space $S = \{1, 2, 3\}$ with 1=rainy, 2=cloudy, 3=sunny

$$\text{Transition matrix} \quad P = \begin{bmatrix} 0.4 & 0.6 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.7 & 0.2 \end{bmatrix}$$



Computing weather predictions manually becomes harder.
Real weather models may have thousands of states...

Example: Google's PageRank

State space $S = \{\text{Set of all web pages}\}$ with size $n = |S|$

Example: Google's PageRank

State space $S = \{\text{Set of all web pages}\}$ with size $n = |S|$

Transition matrix $P \in \mathbb{R}^{n \times n}$ with entries

$$P(x, y) \;=\; (1 - c)\,\frac{G(x, y)}{\sum_{y' \in S} G(x, y')} + c\,\frac{1}{n}$$

Example: Google's PageRank

State space $S = \{\text{Set of all web pages}\}$ with size $n = |S|$

Transition matrix $P \in \mathbb{R}^{n \times n}$ with entries

$$P(x,y) \; = \; (1-c)\,\frac{G(x,y)}{\sum_{y' \in S} G(x,y')} + c\,\frac{1}{n}$$

Adjacency matrix $G \in \{0,1\}^{n \times n}$ with entries

$$G(x,y) \; =$$

Example: Google's PageRank

State space S = {Set of all web pages} with size $n = |S|$

Transition matrix $P \in \mathbb{R}^{n \times n}$ with entries

$$P(x, y) \; = \; (1 - c)\, \frac{G(x, y)}{\sum_{y' \in S} G(x, y')} + c\, \frac{1}{n}$$

Adjacency matrix $G \in \{0, 1\}^{n \times n}$ with entries

$$G(x, y) \; = \; \begin{cases} 1 & \text{if } x \to y \\ 0 & \text{else} \end{cases}$$

Example: Google's PageRank

State space $S = \{\text{Set of all web pages}\}$ with size $n = |S|$

Transition matrix $P \in \mathbb{R}^{n \times n}$ with entries

$$P(x,y) \; = \; (1-c)\,\frac{G(x,y)}{\sum_{y' \in S} G(x,y')} + c\,\frac{1}{n}$$

Adjacency matrix $G \in \{0,1\}^{n \times n}$ with entries

$$G(x,y) \; = \; \begin{cases} 1 & \text{if } x \to y \\ 0 & \text{else} \end{cases}$$

Damping factor $c \in [0,1]$ (classically $c = 0.85$)

Example: Google's PageRank

State space $S = \{\text{Set of all web pages}\}$ with size $n = |S|$

Transition matrix $P \in \mathbb{R}^{n \times n}$ with entries

$$P(x,y) \;=\; (1-c)\,\frac{G(x,y)}{\sum_{y' \in S} G(x,y')} + c\,\frac{1}{n}$$

Adjacency matrix $G \in \{0,1\}^{n \times n}$ with entries

$$G(x,y) \;=\; \begin{cases} 1 & \text{if } x \to y \\ 0 & \text{else} \end{cases}$$

Damping factor $c \in [0,1]$ (classically $c = 0.85$)

Interpretation:
X_t = Location of a surfer at time t who browses the web by randomly selecting hyperlinks

Example: Google's PageRank

State space S = {Set of all web pages} with size $n = |S|$

Transition matrix $P \in \mathbb{R}^{n \times n}$ with entries

$$P(x,y) \;=\; (1-c)\,\frac{G(x,y)}{\sum_{y' \in S} G(x,y')} + c\,\frac{1}{n}$$

Adjacency matrix $G \in \{0,1\}^{n \times n}$ with entries

$$G(x,y) \;=\; \begin{cases} 1 & \text{if } x \to y \\ 0 & \text{else} \end{cases}$$

Damping factor $c \in [0,1]$ (classically $c = 0.85$)

Interpretation:
X_t = Location of a surfer at time t who browses the web by randomly selecting hyperlinks
c = Probability of the surfer deciding to teleport to a random page

Contents

Transient distribution

Distribution of the chain at time t is a vector μ_t with entries

$$\mu_t(x) = \mathbb{P}(X_t = x), \quad x \in S$$

telling the probability of finding the chain in state x at time t

Transient distribution

Distribution of the chain at time t is a vector μ_t with entries

$$\mu_t(x) \ = \ \mathbb{P}(X_t = x), \quad x \in S$$

telling the probability of finding the chain in state x at time t

Properties

Transient distribution

Distribution of the chain at time t is a vector μ_t with entries

$$\mu_t(x) \;=\; \mathbb{P}(X_t = x), \quad x \in S$$

telling the probability of finding the chain in state x at time t

Properties

$$\mu_t(x) \in [0,1] \quad \text{for all } x \in S$$

Transient distribution

Distribution of the chain at time t is a vector μ_t with entries

$$\mu_t(x) \;=\; \mathbb{P}(X_t = x), \quad x \in S$$

telling the probability of finding the chain in state x at time t

Properties

$$\mu_t(x) \in [0,1] \quad \text{for all } x \in S$$

$$\sum_{x \in S} \mu_t(x) \;=$$

Transient distribution

Distribution of the chain at time t is a vector μ_t with entries

$$\mu_t(x) \; = \; \mathbb{P}(X_t = x), \quad x \in S$$

telling the probability of finding the chain in state x at time t

Properties

$$\mu_t(x) \in [0,1] \quad \text{for all } x \in S$$

$$\sum_{x \in S} \mu_t(x) \; = \; \sum_{x \in S} \mathbb{P}(X_t = x) \; =$$

Transient distribution

Distribution of the chain at time t is a vector μ_t with entries

$$\mu_t(x) \ = \ \mathbb{P}(X_t = x), \quad x \in S$$

telling the probability of finding the chain in state x at time t

Properties

$$\mu_t(x) \in [0,1] \quad \text{for all } x \in S$$

$$\sum_{x \in S} \mu_t(x) \ = \ \sum_{x \in S} \mathbb{P}(X_t = x) \ = \ 1$$

Transient distribution

Distribution of the chain at time t is a vector μ_t with entries

$$\mu_t(x) \;=\; \mathbb{P}(X_t = x), \quad x \in S$$

telling the probability of finding the chain in state x at time t

Properties

$$\mu_t(x) \in [0,1] \quad \text{for all } x \in S$$

$$\sum_{x \in S} \mu_t(x) \;=\; \sum_{x \in S} \mathbb{P}(X_t = x) \;=\; 1$$

μ_0 is called the initial distribution of the chain

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

$$\mu_{t+1}(y) \;=$$

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

$$\mu_{t+1}(y) \;=\; \mathbb{P}(X_{t+1} = y)$$

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

$$\begin{aligned}\mu_{t+1}(y) &= \mathbb{P}(X_{t+1} = y) \\ &= \sum_{x \in S} \mathbb{P}(X_t = x)\, \mathbb{P}(X_{t+1} = y \,|\, X_t = x)\end{aligned}$$

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

$$\begin{aligned}
\mu_{t+1}(y) &= \mathbb{P}(X_{t+1} = y) \\
&= \sum_{x \in S} \mathbb{P}(X_t = x)\, \mathbb{P}(X_{t+1} = y \,|\, X_t = x) \\
&= \sum_{x \in S} \mu_t(x) P(x, y)
\end{aligned}$$

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

$$\begin{aligned}
\mu_{t+1}(y) &= \mathbb{P}(X_{t+1} = y) \\
&= \sum_{x \in S} \mathbb{P}(X_t = x)\, \mathbb{P}(X_{t+1} = y \mid X_t = x) \\
&= \sum_{x \in S} \mu_t(x) P(x, y)
\end{aligned}$$

Hence

$$\mu_{t+1}(y) = \sum_{x \in S} \mu_t(x) P(x, y)$$

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

$$\begin{aligned}\mu_{t+1}(y) &= \mathbb{P}(X_{t+1} = y) \\ &= \sum_{x \in S} \mathbb{P}(X_t = x)\,\mathbb{P}(X_{t+1} = y \,|\, X_t = x) \\ &= \sum_{x \in S} \mu_t(x) P(x, y)\end{aligned}$$

Hence

$$\mu_{t+1}(y) = \sum_{x \in S} \mu_t(x) P(x, y)$$

When μ_t and μ_{t+1} are interpreted as row vectors, we can write this is in matrix form as

$$\mu_{t+1} =$$

One-step time evolution

The probability of finding the chain in state y at time $t+1$ equals

$$\begin{aligned}
\mu_{t+1}(y) &= \mathbb{P}(X_{t+1} = y) \\
&= \sum_{x \in S} \mathbb{P}(X_t = x)\, \mathbb{P}(X_{t+1} = y \mid X_t = x) \\
&= \sum_{x \in S} \mu_t(x) P(x, y)
\end{aligned}$$

Hence

$$\mu_{t+1}(y) = \sum_{x \in S} \mu_t(x) P(x, y)$$

When μ_t and μ_{t+1} are interpreted as row vectors, we can write this is in matrix form as

$$\mu_{t+1} = \mu_t P$$

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \ = \ \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.
(ii) **Induction**: Assume claim OK for time $t \geq 0$.

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.

(ii) **Induction**: Assume claim OK for time $t \geq 0$. Then

$$\mu_{t+1} =$$

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \; = \; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.

(ii) **Induction**: Assume claim OK for time $t \geq 0$. Then

$$\mu_{t+1} = \mu_t P$$

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.
(ii) **Induction**: Assume claim OK for time $t \geq 0$. Then

$$\mu_{t+1} = \mu_t P = (\mu_0 P^t) P$$

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.
(ii) **Induction**: Assume claim OK for time $t \geq 0$. Then

$$\mu_{t+1} = \mu_t P = (\mu_0 P^t)P = \mu_0(P^t P)$$

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \ldots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.
(ii) **Induction**: Assume claim OK for time $t \geq 0$. Then

$$\mu_{t+1} = \mu_t P = (\mu_0 P^t) P = \mu_0 (P^t P) = \mu_0 P^{t+1}.$$

Key result

Theorem

The distribution of the chain at time instant $t = 0, 1, 2, \dots$ can be computed by

$$\mu_t \;=\; \mu_0 P^t,$$

where P^t is the t-th power of the transition matrix P.

Proof.

(i) Claim OK for $t = 0$ because $P^0 = I$ is the identity matrix.
(ii) **Induction**: Assume claim OK for time $t \geq 0$. Then

$$\mu_{t+1} = \mu_t P = (\mu_0 P^t) P = \mu_0 (P^t P) = \mu_0 P^{t+1}.$$

Hence claim OK also for time $t + 1$. $\square$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) =$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) =$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy})$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$
$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) =$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) =$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) = \mathbb{P}(\text{Mon} = \text{sunny})$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) = \mathbb{P}(\text{Mon} = \text{sunny}) = 0$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) = \mathbb{P}(\text{Mon} = \text{sunny}) = 0$

Initial distribution as row vector equals $\mu_0 = [1, 0]$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) = \mathbb{P}(\text{Mon} = \text{sunny}) = 0$

Initial distribution as row vector equals $\mu_0 = [1, 0]$

Wed weather distribution $\mu_2 = \mu_0 P^2$ equals

$$[\mu_2(1),\, \mu_2(2)] \;=\; [1,\, 0] \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^2$$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) = \mathbb{P}(\text{Mon} = \text{sunny}) = 0$

Initial distribution as row vector equals $\mu_0 = [1, 0]$

Wed weather distribution $\mu_2 = \mu_0 P^2$ equals

$$[\mu_2(1),\, \mu_2(2)] = [1,\, 0] \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^2 = [0.740,\, 0.260]$$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) = \mathbb{P}(\text{Mon} = \text{sunny}) = 0$

Initial distribution as row vector equals $\mu_0 = [1, 0]$

Wed weather distribution $\mu_2 = \mu_0 P^2$ equals

$$[\mu_2(1),\, \mu_2(2)] \;=\; [1,\, 0] \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^2 \;=\; [0.740,\, 0.260]$$

Sat weather distribution $\mu_5 = \mu_0 P^5$ equals

$$[\mu_5(1),\, \mu_5(2)] \;=\; [1,\, 0] \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^5$$

Example: Weather prediction

If it is cloudy on Mon, then what is the probability that it is cloudy also on Wed? What about Sat? ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Initial distribution $\mu_0(x) = \mathbb{P}(X_0 = x)$

$\mu_0(1) = \mathbb{P}(X_0 = 1) = \mathbb{P}(\text{Mon} = \text{cloudy}) = 1$

$\mu_0(2) = \mathbb{P}(X_0 = 2) = \mathbb{P}(\text{Mon} = \text{sunny}) = 0$

Initial distribution as row vector equals $\mu_0 = [1, 0]$

Wed weather distribution $\mu_2 = \mu_0 P^2$ equals

$$[\mu_2(1),\ \mu_2(2)] \ = \ [1,\ 0] \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^2 \ = \ [0.740,\ 0.260]$$

Sat weather distribution $\mu_5 = \mu_0 P^5$ equals

$$[\mu_5(1),\ \mu_5(2)] \ = \ [1,\ 0] \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^5 \ = \ [0.715,\ 0.285]$$

Transition matrix: Properties

$$P(x, y) \; = \; \mathbb{P}(X_{t+1} = y \,|\, X_t = x)$$

Transition matrix: Properties

$$P(x, y) \;=\; \mathbb{P}(X_{t+1} = y \mid X_t = x)$$

- All entries are probabilities, so $0 \leq P(x, y) \leq 1$
- Every row sum equals

Transition matrix: Properties

$$P(x, y) \;=\; \mathbb{P}(X_{t+1} = y \,|\, X_t = x)$$

- All entries are probabilities, so $0 \leq P(x, y) \leq 1$
- Every row sum equals

$$\sum_{y \in S} P(x, y) \;=\; \sum_{y \in S} \mathbb{P}(X_{t+1} = y \,|\, X_t = x) \;=\; 1.$$

Many-step transition probabilities

Theorem

The probability that a Markov chain moves from state x to state y during t time steps can be computed as

$$\mathbb{P}(X_t = y \,|\, X_0 = x) \;=\; P^t(x, y),$$

where $P^t(x, y)$ is the entry of the t-th power of the transition matrix corresponding to row x and column y.

Many-step transition probabilities

Theorem

The probability that a Markov chain moves from state x to state y during t time steps can be computed as

$$\mathbb{P}(X_t = y \mid X_0 = x) \; = \; P^t(x, y),$$

where $P^t(x, y)$ is the entry of the t-th power of the transition matrix corresponding to row x and column y.

Proof.

Similar induction proof works. [Lecture notes, Thm 1.7] $\square$

Contents

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 and transition matrix P, the path probabilities can be computed by

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \; =$$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 *and transition matrix* P*, the path probabilities can be computed by*

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 and transition matrix P, the path probabilities can be computed by

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=$$

$\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 and transition matrix P, the path probabilities can be computed by

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \,|\, X_0 = x_0)$$

$\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 and transition matrix P, the path probabilities can be computed by

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \,|\, X_0 = x_0) \;=\; \mu_0(x_0)P(x_0, x_1)$$

$\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 and transition matrix P, the path probabilities can be computed by

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \,|\, X_0 = x_0) \;=\; \mu_0(x_0)P(x_0, x_1)$$

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2)$$

$$=$$

$\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 *and transition matrix* P*, the path probabilities can be computed by*

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \,|\, X_0 = x_0) \;=\; \mu_0(x_0)P(x_0, x_1)$$

$$\begin{aligned} &\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\ =\; &\mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \,|\, X_0 = x_0, X_1 = x_1) \end{aligned}$$

$\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 *and transition matrix* P*, the path probabilities can be computed by*

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \dots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \,|\, X_0 = x_0) \;=\; \mu_0(x_0)P(x_0, x_1)$$

$$\begin{aligned} &\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\ =\; &\mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \,|\, X_0 = x_0, X_1 = x_1) \\ =\; &\mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \,|\, X_1 = x_1) \end{aligned}$$

$\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 and transition matrix P, the path probabilities can be computed by

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \,|\, X_0 = x_0) \;=\; \mu_0(x_0)P(x_0, x_1)$$

$$\begin{aligned}
&\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\
&= \mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \,|\, X_0 = x_0, X_1 = x_1) \\
&= \mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \,|\, X_1 = x_1) \\
&= \mu_0(x_0)P(x_0, x_1)P(x_1, x_2).
\end{aligned}$$

$\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 *and transition matrix* P*, the path probabilities can be computed by*

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \mid X_0 = x_0) \;=\; \mu_0(x_0)P(x_0, x_1)$$

$$\begin{aligned} &\mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\ =\;& \mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \mid X_0 = x_0, X_1 = x_1) \\ =\;& \mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \mid X_1 = x_1) \\ =\;& \mu_0(x_0)P(x_0, x_1)P(x_1, x_2). \end{aligned}$$

For $t \geq 3$: $\square$

Path probabilities

Theorem

For any Markov chain with initial distribution μ_0 and transition matrix P, the path probabilities can be computed by

$$\mathbb{P}(X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t) \;=\; \mu_0(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

Proof.

$$\mathbb{P}(X_0 = x_0, X_1 = x_1) \;=\; \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1 \,|\, X_0 = x_0) \;=\; \mu_0(x_0)P(x_0, x_1)$$

$$\begin{aligned}
& \mathbb{P}(X_0 = x_0, X_1 = x_1, X_2 = x_2) \\
=\; & \mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \,|\, X_0 = x_0, X_1 = x_1) \\
= & \mathbb{P}(X_0 = x_0, X_1 = x_1)\,\mathbb{P}(X_2 = x_2 \,|\, X_1 = x_1) \\
= & \mu_0(x_0)P(x_0, x_1)P(x_1, x_2).
\end{aligned}$$

For $t \geq 3$: Induction... $\square$

Occupancy of states

The frequency of state y among the first t states is

$$N_t(y) \;=\; \sum_{s=0}^{t-1} 1(X_s = y),$$

Occupancy matrix M_t has entries

$$M_t(x,y) \;=\; \mathbb{E}(N_t(y) \,|\, X_0 = x).$$

The entry of the occupancy matrix M_t for row x and column y tells the expected number of times that a chain starting at x visits y during the first t time instants.

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) =$$

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) \;=\; \mathbb{E}_x \sum_{s=0}^{t-1} 1(X_s = y)$$

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) = \mathbb{E}_x \sum_{s=0}^{t-1} 1(X_s = y) = \sum_{s=0}^{t-1} \mathbb{E}_x 1(X_s = y)$$

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) \; = \; \mathbb{E}_x \sum_{s=0}^{t-1} 1(X_s = y) \; = \; \sum_{s=0}^{t-1} \mathbb{E}_x 1(X_s = y) \; = \; \sum_{s=0}^{t-1} \mathbb{P}_x(X_s = y)$$

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) \;=\; \mathbb{E}_x \sum_{s=0}^{t-1} 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{E}_x 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{P}_x(X_s = y)$$

$$\implies \quad M_t(x, y) \;=$$

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) \;=\; \mathbb{E}_x \sum_{s=0}^{t-1} 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{E}_x 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{P}_x(X_s = y)$$

$$\implies \quad M_t(x,y) \;=\; \mathbb{E}_x N_t(y)$$

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) \;=\; \mathbb{E}_x \sum_{s=0}^{t-1} 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{E}_x 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{P}_x(X_s = y)$$

$$\implies \quad M_t(x,y) \;=\; \mathbb{E}_x N_t(y) \;=\; \sum_{s=0}^{t-1} \mathbb{P}_x(X_s = y)$$

Computing the occupancy matrix

Theorem

The occupancy matrix can be computed as $M_t = \sum_{s=0}^{t-1} P^s$.

Proof.

Let $\mathbb{P}_x, \mathbb{E}_x$ be the conditional probability and expectation given $X_0 = x$.

$$\mathbb{E}_x N_t(y) \;=\; \mathbb{E}_x \sum_{s=0}^{t-1} 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{E}_x 1(X_s = y) \;=\; \sum_{s=0}^{t-1} \mathbb{P}_x(X_s = y)$$

$$\Longrightarrow \quad M_t(x,y) \;=\; \mathbb{E}_x N_t(y) \;=\; \sum_{s=0}^{t-1} \mathbb{P}_x(X_s = y) \;=\; \sum_{s=0}^{t-1} P^s(x,y).$$

$\square$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The 7-day occupancy matrix equals

$$M_7 =$$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The 7-day occupancy matrix equals

$$M_7 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^1 + \cdots + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^6$$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The 7-day occupancy matrix equals

$$M_7 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^1 + \cdots + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^6 = \begin{bmatrix} 5.408 & 1.592 \\ 3.980 & 3.020 \end{bmatrix}$$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The 7-day occupancy matrix equals

$$M_7 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^1 + \cdots + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^6 = \begin{bmatrix} 5.408 & 1.592 \\ 3.980 & 3.020 \end{bmatrix}$$

According to the prediction, the expected number of cloudy days equals

$$M_7(\ ,\ ) =$$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The 7-day occupancy matrix equals

$$M_7 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^1 + \cdots + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^6 = \begin{bmatrix} 5.408 & 1.592 \\ 3.980 & 3.020 \end{bmatrix}$$

According to the prediction, the expected number of cloudy days equals

$$M_7(2, \quad) =$$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The 7-day occupancy matrix equals

$$M_7 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^1 + \cdots + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^6 = \begin{bmatrix} 5.408 & 1.592 \\ 3.980 & 3.020 \end{bmatrix}$$

According to the prediction, the expected number of cloudy days equals

$$M_7(2, 1) =$$

Example: Weather model

Predict the expected number of cloudy days during a week starting with a sunny day.
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

The 7-day occupancy matrix equals

$$M_7 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^1 + \cdots + \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}^6 = \begin{bmatrix} 5.408 & 1.592 \\ 3.980 & 3.020 \end{bmatrix}$$

According to the prediction, the expected number of cloudy days equals

$$M_7(2, 1) = 3.980$$

Contents

Simulation

A stochastic representation of transition matrix $P =$

Simulation

A stochastic representation of transition matrix $P =$
Pair (f, U) where $f : S \times S' \to S$ is a deterministic function and $U \in S'$ is a random variable such that

$$\mathbb{P}(f(x, U) = y) \;=\; P(x, y) \quad \text{for all } x, y \in S.$$

Simulation

A stochastic representation of transition matrix $P =$
Pair (f, U) where $f : S \times S' \to S$ is a deterministic function and $U \in S'$ is a random variable such that

$$\mathbb{P}(f(x, U) = y) \;=\; P(x, y) \quad \text{for all } x, y \in S.$$

Simulation:

(i) Find a stochastic representation (f, U) of P.

Simulation

A stochastic representation of transition matrix $P =$
Pair (f, U) where $f : S \times S' \to S$ is a deterministic function and $U \in S'$ is a random variable such that

$$\mathbb{P}(f(x, U) = y) \;=\; P(x, y) \quad \text{for all } x, y \in S.$$

Simulation:

(i) Find a stochastic representation (f, U) of P.

(ii) Find a random number generator which produces independent random variables $U_1, U_2, \dots$ with the same distribution as U.

Simulation

A stochastic representation of transition matrix $P =$
Pair (f, U) where $f : S \times S' \to S$ is a deterministic function and $U \in S'$ is a random variable such that

$$\mathbb{P}(f(x, U) = y) \;=\; P(x, y) \quad \text{for all } x, y \in S.$$

Simulation:

(i) Find a stochastic representation (f, U) of P.

(ii) Find a random number generator which produces independent random variables $U_1, U_2, \dots$ with the same distribution as U.

(iii) Compute $X_{t+1} = f(X_t, U_{t+1})$ for $t = 0, 1, \dots$

Simulation

A stochastic representation of transition matrix $P =$
Pair (f, U) where $f : S \times S' \to S$ is a deterministic function and $U \in S'$ is a random variable such that

$$\mathbb{P}(f(x, U) = y) \;=\; P(x, y) \quad \text{for all } x, y \in S.$$

Simulation:

(i) Find a stochastic representation (f, U) of P.

(ii) Find a random number generator which produces independent random variables $U_1, U_2, \dots$ with the same distribution as U.

(iii) Compute $X_{t+1} = f(X_t, U_{t+1})$ for $t = 0, 1, \dots$

Then $(X_0, X_1, X_2, \dots)$ is a Markov chain with transition matrix P.

Example: Weather simulator

Develop a simulator for the weather model ($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \ldots, 10\}$)

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1,2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1,\dots,10\}$) and define a function $f : \{1,2\} \times S' \to \{1,2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1										
2										

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \dots, 10\}$) and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1		
2										

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \ldots, 10\}$)
and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2										

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \ldots, 10\}$) and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1					

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1,2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1,\dots,10\}$) and define a function $f : \{1,2\} \times S' \to \{1,2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \ldots, 10\}$) and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Stochastic representation 2:

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \ldots, 10\}$) and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Stochastic representation 2:
Let U be uniformly distributed in the continuous interval $S' = [0, 1]$

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \dots, 10\}$) and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Stochastic representation 2:
Let U be uniformly distributed in the continuous interval $S' = [0, 1]$ and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ by

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \dots, 10\}$)
and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Stochastic representation 2:
Let U be uniformly distributed in the continuous interval $S' = [0, 1]$
and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ by

$$f(x, u) =$$

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1,2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1,\dots,10\}$)
and define a function $f : \{1,2\} \times S' \to \{1,2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Stochastic representation 2:
Let U be uniformly distributed in the continuous interval $S' = [0,1]$
and define a function $f : \{1,2\} \times S' \to \{1,2\}$ by

$$f(x,u) = \begin{cases} 1 & \text{if } x = 1 \text{ and } u \le 0.8 \\ & \\ & \end{cases}$$

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \ldots, 10\}$)
and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Stochastic representation 2:
Let U be uniformly distributed in the continuous interval $S' = [0, 1]$
and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ by

$$f(x, u) = \begin{cases} 1 & \text{if } x = 1 \text{ and } u \leq 0.8 \\ 1 & \text{if } x = 2 \text{ and } u \leq 0.5 \\ & \end{cases}$$

Example: Weather simulator

Develop a simulator for the weather model
($S = \{1, 2\}$ with 1=cloudy, 2=sunny)

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$$

Stochastic representation:
Let U be a 10-sided die (uniformly distributed on $S' = \{1, \ldots, 10\}$)
and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ via the table

	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	2	2
2	1	1	1	1	1	2	2	2	2	2

Stochastic representation 2:
Let U be uniformly distributed in the continuous interval $S' = [0, 1]$
and define a function $f : \{1, 2\} \times S' \to \{1, 2\}$ by

$$f(x, u) = \begin{cases} 1 & \text{if } x = 1 \text{ and } u \leq 0.8 \\ 1 & \text{if } x = 2 \text{ and } u \leq 0.5 \\ 2 & \text{else} \end{cases}$$

Discussion: Dynamical systems vs. Markov chains

	Dynamical system	**Markov chain**
Initial state	Deterministic x_0	Random X_0 with law μ_0
State evolution	$x_{t+1} = f(x_t)$	$X_{t+1} = f(X_t, U_{t+1})$
Law at time t	δ_{x_t}	μ_t
Law evolution	$\mu_{t+1} = \delta_{f(x_t)}$	$\mu_{t+1} = \mu_t P$

Every Markov chain admits a (nonunique) stochastic representation $X_{t+1} = f(X_t, U_{t+1})$ for some $f : S \times S' \to S$ and some $U_1, U_2, \ldots$

Next time we discuss Markov chains in the long run as $t \to \infty$

Kirjallisuutta

- R Durrett.
  Essentials of Stochastic Processes.
  2nd edition, Springer 2012.
- D Williams.
  Probability with Martingales.
  Cambridge University Press 1991.

Aineistolähteet

Esityksessä käytetyt kuvat (esiintymisjärjestyksessä)

1. Image courtesy of think4photop at FreeDigitalPhotos.net
2. Image courtesy of Lisa Gansky from New York, NY, USA [CC BY-SA 2.0 (https://creativecommons.org/licenses/by-sa/2.0)]