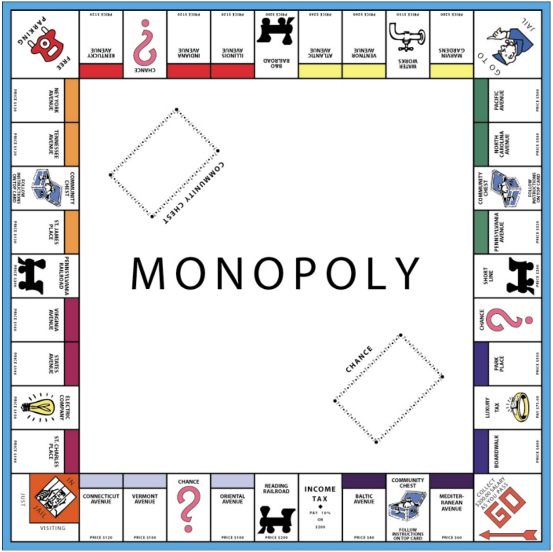
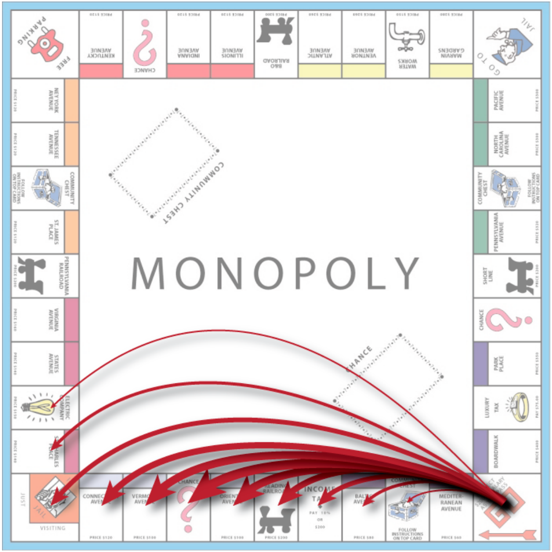
Chapter 2: Discrete-Time Markov Models (Part i)

Andrey Markov (1856-1922, Russian mathematician)

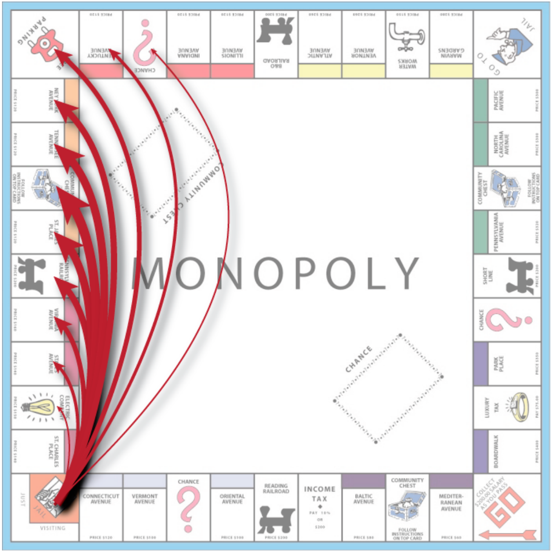
life is a *game* with uncertainty:



Game of Monopoly: at least 40 states where you live in



At least 40 × 11 transitions between states.



- Fate (next state) depends on where you are (current state).
- Fate does *not* depend on how you got there (history).

What we learn from Game of Monopoly

What we learn from Game of Monopoly

YESTERDAY IS HISTORY. TOMORROW IS A MYSTERY.
TODAY IS A GIFT.

What we learn from Game of Monopoly

YESTERDAY IS HISTORY. TOMORROW IS A MYSTERY.
TODAY IS A GIFT.

— the Turtle (*Kung Fu Panda (movie)*)

What we learn from Game of Monopoly

YESTERDAY IS HISTORY. TOMORROW IS A MYSTERY.
TODAY IS A GIFT.

— the Turtle (*Kung Fu Panda (movie)*)

Markovian property:

- X_{n+1} is a random variable, whose law (distribution) is determined only by the current value of X_n and independent of all history $X_{n-1}, X_{n-2}, \cdots$.
- The law governing the distribution of X_{n+1} (given the value of X_n) is called *transition probability*: $\Pr(X_{n+1} = ?|X_n)$

What we learn from Game of Monopoly

YESTERDAY IS HISTORY. TOMORROW IS A MYSTERY.
TODAY IS A GIFT.

— the Turtle (*Kung Fu Panda (movie)*)

Markovian property:

- X_{n+1} is a random variable, whose law (distribution) is determined only by the current value of X_n and independent of all history $X_{n-1}, X_{n-2}, \cdots$.
- The law governing the distribution of X_{n+1} (given the value of X_n) is called *transition probability*: $\Pr(X_{n+1} = ?|X_n)$

Corollary:

- What is the joint distribution of (X_{n+1}, X_n) ? $\Pr(X_{n+1}, X_n) = \Pr(X_{n+1}|X_n)\Pr(X_n)$
- What is the (marginal) distribution of X_{n+1} ?
  $\Pr(X_{n+1}) = \sum_x \Pr(X_{n+1}, X_n = x) = \sum_x \Pr(X_{n+1}|X_n = x)\Pr(X_n = x)$

Markov model of the weather

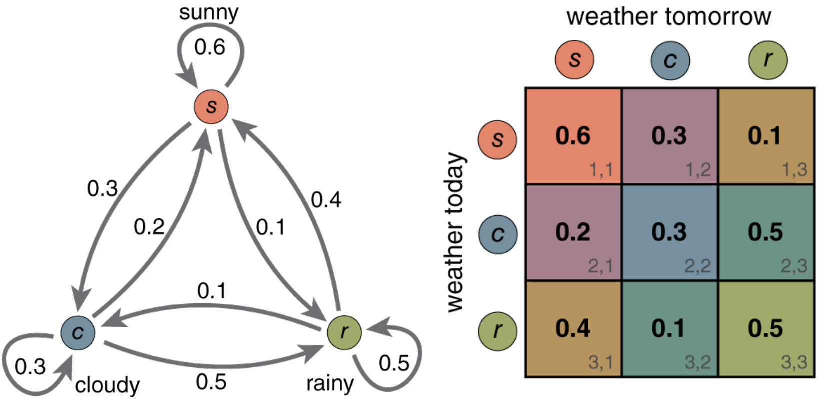


Figure: transition diagram and transition matrix

Theory of DTMC

Definition

A stochastic process $\{X_n, n \geq 0\}$ on state space S is said to be a discrete-time Markov chain (DTMC) if , for all i and j in S ,

$$\Pr(X_{n+1} = j \,|\, X_n = i, X_{n-1}, \ldots X_0) = \Pr(X_{n+1} = j \,|\, X_n = i)$$

A DTMC $\{X_n\}$ is said to be *time-homogeneous* if, for all $n = 0, 1, 2, \ldots$,

$$\Pr(X_{n+1} = j \,|\, X_n = i)$$

is *independent* of the time n.

In this chapter we mainly consider time-homogeneous DTMCs with finite state space $S = \{1, 2, 3, \cdots, N\}$, and occasionally we may consider state space with countable infinite elements labelled as $S = \mathbb{N} := \{1, 2, 3, \cdots\}$ or $S = \mathbb{Z} := \{\cdots, -2, -1, 0, 1, 2, \cdots\}$.

Transition Probability Matrix

One-step transition probability

$$p_{ij} = \Pr(X_{n+1} = j \,|\, X_n = i)$$

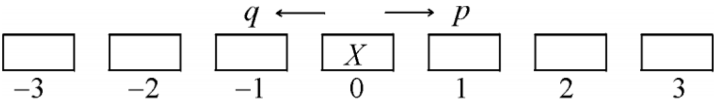
Remark: Strictly speaking, the above notation p_{ij} also depends on the time n. We only consider the time-homogeneous DTMC. So we only have one transition matrix which works for all time.

transition matrix

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}$$

Example: Random Walk

random walk on $\mathbb{Z}$



Let $X_0 = 0$ and $X_n = \sum_{i=1}^{n} Z_i$, where the iid random variable Z_i is defined as follows

$$Z_i = \begin{cases} +1, & \text{with prob } p = 0.5 \\ -1, & \text{with prob } 1-p = 0.5 \end{cases}$$

The case of $p = q = 1/2$ is called symmetric random walk.

The random walk is a DTMC on $S = \mathbb{Z}$

Proof of Markovian property:

The random walk is a DTMC on $S = \mathbb{Z}$

Proof of Markovian property:

$$
\begin{aligned}
&\Pr(X_{n+1} = j | X_n = i, X_{n-1}, \cdots, X_0) \\
=&\Pr(X_n + Z_{n+1} = j | X_n = i, X_{n-1}, \cdots, X_0) \\
=&\Pr(Z_{n+1} = j - X_n | X_n = i, X_{n-1}, \cdots, X_0) \\
=&\Pr(Z_{n+1} = j - i | X_n = i) \quad \because Z_{n+1} \text{ indpt. of } \{X_n, X_{n-1}, \ldots\} \\
=&\begin{cases} p & \text{if } j = i+1; \\ q & \text{if } j = i-1; \\ 0 & \text{else.} \end{cases}
\end{aligned}
$$

The random walk is a DTMC on $S = \mathbb{Z}$

Proof of Markovian property:

$$
\begin{aligned}
&\Pr(X_{n+1} = j | X_n = i, X_{n-1}, \cdots, X_0) \\
=&\Pr(X_n + Z_{n+1} = j | X_n = i, X_{n-1}, \cdots, X_0) \\
=&\Pr(Z_{n+1} = j - X_n | X_n = i, X_{n-1}, \cdots, X_0) \\
=&\Pr(Z_{n+1} = j - i | X_n = i) \quad \because Z_{n+1} \text{ indpt. of } \{X_n, X_{n-1}, \ldots\} \\
=&\begin{cases} p & \text{if } j = i+1; \\ q & \text{if } j = i-1; \\ 0 & \text{else.} \end{cases}
\end{aligned}
$$

Question: What is the transition diagram and transition matrix ?

The random walk is a DTMC on $S = \mathbb{Z}$

Proof of Markovian property:

$$
\begin{aligned}
& \Pr(X_{n+1} = j | X_n = i, X_{n-1}, \cdots, X_0) \\
=& \Pr(X_n + Z_{n+1} = j | X_n = i, X_{n-1}, \cdots, X_0) \\
=& \Pr(Z_{n+1} = j - X_n | X_n = i, X_{n-1}, \cdots, X_0) \\
=& \Pr(Z_{n+1} = j - i | X_n = i) \quad \because Z_{n+1} \text{ indpt. of } \{X_n, X_{n-1}, \ldots\} \\
=& \begin{cases} p & \text{if } j = i+1; \\ q & \text{if } j = i-1; \\ 0 & \text{else.} \end{cases}
\end{aligned}
$$

Question: What is the transition diagram and transition matrix ?

$$
\mathbf{P} = \begin{bmatrix} \cdots & q & 0 & p & \cdots & \cdots \\ \cdots & \cdots & q & 0 & p & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}
$$

Random Walk

random walk on finite interval of $\mathbb{Z}$

Let $-M, N$ be two positive integers, define a random walk on

$$S = [-M, N] \cap \mathbb{Z} = \{-M, -M+1, \cdots, N-1, N\}.$$

The particle X_n will randomly jump to one of its two neighbors, according to probability p and q where $p+q=1$. Define $\Pr(X_{n+1} = i+1 | X_n = i) = p$, $\Pr(X_{n+1} = i-1 | X_n = i) = q$ for any *interior* point i.

Let $-M, N$ be two positive integers, define a random walk on

$$S = [-M, N] \cap \mathbb{Z} = \{-M, -M+1, \cdots, N-1, N\}.$$

The particle X_n will randomly jump to one of its two neighbors, according to probability p and q where $p+q=1$. Define $\Pr(X_{n+1} = i+1 | X_n = i) = p$, $\Pr(X_{n+1} = i-1 | X_n = i) = q$ for any *interior* point i.

For finite interval $[-M, N]$, we need specify the boundary condition

1. absorbing (gambler's ruin problem)

$\Pr(X_{n+1} = N | X_n = N) = 1$ and $\Pr(X_{n+1} = -M | X_n = -M) = 1$

Random Walk

random walk on finite interval of $\mathbb{Z}$

Let $-M, N$ be two positive integers, define a random walk on

$$S = [-M, N] \cap \mathbb{Z} = \{-M, -M+1, \cdots, N-1, N\}.$$

The particle X_n will randomly jump to one of its two neighbors, according to probability p and q where $p+q=1$. Define $\Pr(X_{n+1} = i+1 | X_n = i) = p$, $\Pr(X_{n+1} = i-1 | X_n = i) = q$ for any *interior* point i.

For finite interval $[-M, N]$, we need specify the boundary condition

1. absorbing (gambler's ruin problem)

$\Pr(X_{n+1} = N | X_n = N) = 1$ and $\Pr(X_{n+1} = -M | X_n = -M) = 1$

2. reflection (elastic wall)

$\Pr(X_{n+1} = N-1 | X_n = N) = 1$ and $\Pr(X_{n+1} = -M+1 | X_n = -M) = 1$

Random Walk

random walk on finite interval of $\mathbb{Z}$

Let $-M, N$ be two positive integers, define a random walk on

$$S = [-M, N] \cap \mathbb{Z} = \{-M, -M+1, \cdots, N-1, N\}.$$

The particle X_n will randomly jump to one of its two neighbors, according to probability p and q where $p + q = 1$. Define $\Pr(X_{n+1} = i+1 | X_n = i) = p$, $\Pr(X_{n+1} = i-1 | X_n = i) = q$ for any *interior* point i.

For finite interval $[-M, N]$, we need specify the boundary condition

1. absorbing (gambler's ruin problem)

$\Pr(X_{n+1} = N | X_n = N) = 1$ and $\Pr(X_{n+1} = -M | X_n = -M) = 1$

2. reflection (elastic wall)

$\Pr(X_{n+1} = N-1 | X_n = N) = 1$ and $\Pr(X_{n+1} = -M+1 | X_n = -M) = 1$

3. periodic (walk on a circle)

$\Pr(X_{n+1} = -M | X_n = N) = p$ and $\Pr(X_{n+1} = N | X_n = -M) = q$

- absorbing boundary condition

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- absorbing boundary condition

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- reflection boundary condition

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

- absorbing boundary condition

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- reflection boundary condition

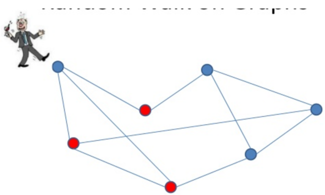
$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

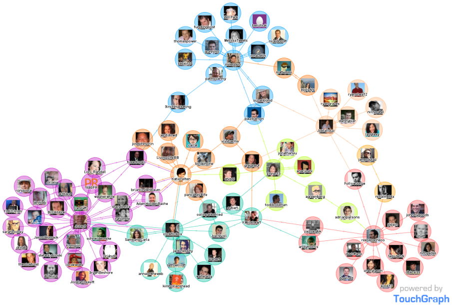
- periodic boundary condition

$$\mathbf{P} = \begin{bmatrix} 0 & p & 0 & 0 & q \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ p & 0 & 0 & q & 0 \end{bmatrix}$$

Random Walk on Graph

Consider a connected graph without self-loop with node $S = \{1, 2, \ldots, N\}$. The walker at state i goes to one of its direct neighbours with equal probability $1/d_i$, where d_i is the degree of node i. For undirected graph, the movement direction is bi-directional.





Transient Distribution: $\Pr(X_n = j) =?$

For the symmetric random walk, we worked very hard to obtain if $X_0 = 0$, then

$$\Pr(X_1 = 1) = p, \Pr(X_1 = -1) = q,$$
$$\Pr(X_2 = 2) = p^2, \Pr(X_2 = 0) = 2pq, \Pr(X_2 = -2) = q^2,$$
$$\Pr(X_3 = 3) = p^3, \Pr(X_3 = 1) = 3p^2q, \Pr(X_3 = -1) = 3pq^2, \Pr(X_3 = -3) = q^3,$$
$$\cdots$$

there is an easy way to calculate for DTMC

you need know matrix-vector multiplication and matrix power.

- Let the state space $S = \{1, 2, \ldots, N\}$. Specify an initial distribution at time $t = 0$: $\Pr(X_0 = j) = a_j$ where $a = (a_1, a_2, \cdots, a_N)$ satisfies $a_j \geq 0 \ \forall j$ and $\sum_{j=1}^{N} a_j = 1$. (called a *stochastic vector* or *probability vector*.)

there is an easy way to calculate for DTMC

you need know matrix-vector multiplication and matrix power.

- Let the state space $S = \{1, 2, \ldots, N\}$. Specify an initial distribution at time $t = 0$: $\Pr(X_0 = j) = a_j$ where $a = (a_1, a_2, \cdots, a_N)$ satisfies $a_j \geq 0 \ \forall j$ and $\sum_{j=1}^{N} a_j = 1$. (called a *stochastic vector* or *probability vector*.)
- Define n-step transition probability in matrix form $\mathbf{P}^{(n)} = [p_{ij}^{(n)}]$:

$$p_{ij}^{(n)} \triangleq \Pr(X_n = j | X_0 = i), \quad \mathbf{P}^{(0)} = \mathbf{I}, \quad \mathbf{P}^{(1)} = \mathbf{P}.$$

there is an easy way to calculate for DTMC

you need know matrix-vector multiplication and matrix power.

- Let the state space $S = \{1, 2, \dots, N\}$. Specify an initial distribution at time $t = 0$: $\Pr(X_0 = j) = a_j$ where $a = (a_1, a_2, \cdots, a_N)$ satisfies $a_j \geq 0\ \forall j$ and $\sum_{j=1}^{N} a_j = 1$. (called a *stochastic vector* or *probability vector*.)
- Define n-step transition probability in matrix form $\mathbf{P}^{(n)} = [p_{ij}^{(n)}]$:

$$p_{ij}^{(n)} \triangleq \Pr(X_n = j | X_0 = i), \quad \mathbf{P}^{(0)} = \mathbf{I}, \quad \mathbf{P}^{(1)} = \mathbf{P}.$$

- From Law of Total Probability, for any $j \in S$,

$$\Pr(X_n = j) = \sum_{i \in S} \Pr(X_n = j | X_0 = i) \Pr(X_0 = i) = \sum_i p_{ij}^{(n)} a_i = (a\mathbf{P}^{(n)})_j \tag{1}$$

  here $a\mathbf{P}^{(n)}$ is a (row)vector-matrix multiplication.

Joint distribution (path distribution)

What is the probability that the sample path of the (time homogeneous) DTMC is $(X_0, X_1, X_2, \cdots, X_n) = (i_0, i_1, i_2, \cdots, i_n)$? i.e., the joint distribution of $(X_0, X_1, X_2, \cdots, X_n)$?

*In the continuous limit (diffusion), these "sums" become the so-called *path-integral*.

Joint distribution (path distribution)

What is the probability that the sample path of the (time homogeneous) DTMC is $(X_0, X_1, X_2, \cdots, X_n) = (i_0, i_1, i_2, \cdots, i_n)$? i.e., the joint distribution of $(X_0, X_1, X_2, \cdots, X_n)$?

$$
\begin{aligned}
&\Pr\Big((X_0, X_1, X_2, \cdots, X_n) = (i_0, i_1, i_2, \cdots, i_n)\Big) \\
=&\Pr(X_0 = i_0) \times \Pr(X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n \mid X_0 = i_0) \\
=&a_{i_0}\Pr(X_1 = i_1 \mid X_0 = i_0) \times \Pr(X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n \mid X_0 = i_0, X_1 = i_1) \qquad (\triangle) \\
=&a_{i_0} p_{i_0,i_1} \Pr(X_2 = i_2, \cdots, X_n = i_n \mid X_1 = i_1) \\
=&\cdots \\
=&a_{i_0} \times p_{i_0,i_1} \times p_{i_1,i_2} \times p_{i_{n-1},i_n}
\end{aligned}
$$

So, the marginal distribution for X_n is

*In the continuous limit (diffusion), these "sums" become the so-called *path-integral*.

Joint distribution (path distribution)

What is the probability that the sample path of the (time homogeneous) DTMC is $(X_0, X_1, X_2, \cdots, X_n) = (i_0, i_1, i_2, \cdots, i_n)$? i.e., the joint distribution of $(X_0, X_1, X_2, \cdots, X_n)$?

$$\begin{aligned}
&\Pr\Big((X_0, X_1, X_2, \cdots, X_n) = (i_0, i_1, i_2, \cdots, i_n)\Big) \\
=&\Pr(X_0 = i_0) \times \Pr(X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n \mid X_0 = i_0) \\
=&a_{i_0} \Pr(X_1 = i_1 \mid X_0 = i_0) \times \Pr(X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n \mid X_0 = i_0, X_1 = i_1) \\
=&a_{i_0} p_{i_0,i_1} \Pr(X_2 = i_2, \cdots, X_n = i_n \mid X_1 = i_1) \\
=&\cdots \\
=&a_{i_0} \times p_{i_0,i_1} \times p_{i_1,i_2} \times p_{i_{n-1},i_n}
\end{aligned} \tag{$\triangle$}$$

So, the marginal distribution for X_n is

$$\Pr(X_n = i_n) =^{*} \sum_{i_0} \sum_{i_1} \cdots \sum_{i_{n-1}} (\triangle) = (a\mathbf{P}^n)_{i_n} \tag{2}$$

---

*In the continuous limit (diffusion), these "sums" become the so-called *path-integral*.

Joint distribution (path distribution)

What is the probability that the sample path of the (time homogeneous) DTMC is $(X_0, X_1, X_2, \cdots, X_n) = (i_0, i_1, i_2, \cdots, i_n)$? i.e., the joint distribution of $(X_0, X_1, X_2, \cdots, X_n)$?

$$
\begin{aligned}
&\Pr\Big((X_0, X_1, X_2, \cdots, X_n) = (i_0, i_1, i_2, \cdots, i_n)\Big) \\
=&\Pr(X_0 = i_0) \times \Pr(X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n \mid X_0 = i_0) \\
=&a_{i_0} \Pr(X_1 = i_1 \mid X_0 = i_0) \times \Pr(X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n \mid X_0 = i_0, X_1 = i_1) \\
=&a_{i_0} p_{i_0,i_1} \Pr(X_2 = i_2, \cdots, X_n = i_n \mid X_1 = i_1) \\
=&\cdots \\
=&a_{i_0} \times p_{i_0,i_1} \times p_{i_1,i_2} \times p_{i_{n-1},i_n}
\end{aligned} \qquad (\triangle)
$$

So, the marginal distribution for X_n is

$$
\Pr(X_n = i_n) =^* \sum_{i_0} \sum_{i_1} \cdots \sum_{i_{n-1}} (\triangle) = (a\mathbf{P}^n)_{i_n} \qquad (2)
$$

Compare (1) and (2) which hold for any vector a, then

$$
a\mathbf{P}^{(n)} = a\mathbf{P}^n \Longrightarrow \mathbf{P}^{(n)} = \mathbf{P}^n.
$$

*In the continuous limit (diffusion), these “sums” become the so-called *path-integral*.

n-step transition matrix

Theorem ((for time-homogeneous MC). Thm 2.2)

$$\mathbf{P}^{(n)} = \mathbf{P}^n.$$

where $\mathbf{P}^n$ *is the nth power of the one-step transition matrix* $\mathbf{P}$*.*

n-step transition matrix

Theorem ((for time-homogeneous MC). Thm 2.2)

$$\mathbf{P}^{(n)} = \mathbf{P}^n.$$

where $\mathbf{P}^n$ *is the nth power of the one-step transition matrix* $\mathbf{P}$*.*

We also have the following important property due to time homogeneity (invariant under time shift)

(Cor. 2.2)

$$\Pr(X_{n+k} = j | X_k = i) = \Pr(X_n = j | X_0 = i) = p_{ij}^{(n)}, \forall k$$

Properties of transition matrix

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij \geq 0}$ for all i, j.

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij \geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij\geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij \geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

$\lambda_1 = 1$ is one of the eigenvalues of any stochastic matrix.

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij \geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

$\lambda_1 = 1$ is one of the eigenvalues of any stochastic matrix.

Theorem

- *The transition matrix* $\mathbf{P}$ *is a stochastic matrix.*

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij\geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

$\lambda_1 = 1$ is one of the eigenvalues of any stochastic matrix.

Theorem

- *The transition matrix* $\mathbf{P}$ *is a stochastic matrix.*
- $\mathbf{P}^n$ *for any integer* $n \geq 0$ *is a stochastic matrix.*

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij \geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

$\lambda_1 = 1$ is one of the eigenvalues of any stochastic matrix.

Theorem

- *The transition matrix* $\mathbf{P}$ *is a stochastic matrix.*
- $\mathbf{P}^n$ *for any integer* $n \geq 0$ *is a stochastic matrix.*

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij \geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

$\lambda_1 = 1$ is one of the eigenvalues of any stochastic matrix.

Theorem

- *The transition matrix* $\mathbf{P}$ *is a stochastic matrix.*
- $\mathbf{P}^n$ *for any integer* $n \geq 0$ *is a stochastic matrix.*

Furthermore, $\mathbf{P}$ satisfies (*optional*)

- All eigenvalues (complex value possibly) satisfy $|\lambda_i| \leq 1$ (Perron-Frobenius theorem). So, the spectral radius

$$\rho(\mathbf{P}) \triangleq \max_i |\lambda| = 1.$$

Definition

An $N \times N$ matrix $P = [p_{ij}]$ is called a stochastic matrix if

1. $p_{ij \geq 0}$ for all i, j.
2. $\sum_{j=1}^{N} p_{ij} = 1$ for all i. That is the sum of each row is 1. In matrix form, $P\mathbb{1} = \mathbb{1}$, where $\mathbb{1} = (1, 1, \cdots, 1)^{\mathsf{T}}$

$\lambda_1 = 1$ is one of the eigenvalues of any stochastic matrix.

Theorem

- *The transition matrix* $\mathbf{P}$ *is a stochastic matrix.*
- $\mathbf{P}^n$ *for any integer* $n \geq 0$ *is a stochastic matrix.*

Furthermore, $\mathbf{P}$ satisfies (*optional*)

- All eigenvalues (complex value possibly) satisfy $|\lambda_i| \leq 1$ (Perron-Frobenius theorem). So, the spectral radius

$$\rho(\mathbf{P}) \triangleq \max_i |\lambda| = 1.$$

- $\lambda_1 = 1$ *might not be the only eigenvalue on the unit circle and the associated eigenspace can be multi-dimensional.*

Chapman–Kolmogorov Equation

This is the most fundamental equation for Markov process (not limited to time-homogeneous case, even not to the discrete time or discrete state space).

Theorem (Thm 2.3)

The n-step transition probabilities $\mathbf{P}^{(n)}$ *satisfy the following equation, called the Chapman–Kolmogorov equation:*

$$p_{ij}^{(n+m)} = \sum_{k=1}^{N} p_{ik}^{(n)} p_{kj}^{(m)}.$$

Remark : For time homogeneous DTMC, this is the simple fact for any square matrix $\mathbf{P}$: $\mathbf{P}^{n+m} = \mathbf{P}^n \mathbf{P}^m$

Proof: textbook page 18.

Occupancy Times

Occupancy time

is the expected amount of time that the Markov chain spends in a given state during a given interval of time.

Occupancy Times

- Let $N_j^{(n)}$ be the number of times the DTMC visits state j over the time span $\{0,1,2,\cdots,n\}$. That is $N_j^{(n)} = \sum_{t=0}^{n} 1_{\{X_t=j\}}$, where $1_{\{\cdot\}}$ is the indicator function.

Occupancy time

is the expected amount of time that the Markov chain spends in a given state during a given interval of time.

Occupancy Times

- Let $N_j^{(n)}$ be the number of times the DTMC visits state j over the time span $\{0,1,2,\cdots,n\}$. That is $N_j^{(n)} = \sum_{t=0}^{n} 1_{\{X_t=j\}}$, where $1_{\{\cdot\}}$ is the indicator function.
- $m_{i,j}^{(n)} = \mathrm{E}(N_j^{(n)}|X_0 = i)$ is the occupancy time up to n of state j starting from state i.

Occupancy time

is the expected amount of time that the Markov chain spends in a given state during a given interval of time.

Occupancy Times

- Let $N_j^{(n)}$ be the number of times the DTMC visits state j over the time span $\{0, 1, 2, \cdots, n\}$. That is $N_j^{(n)} = \sum_{t=0}^{n} 1_{\{X_t = j\}}$, where $1_{\{\cdot\}}$ is the indicator function.
- $m_{i,j}^{(n)} = \mathrm{E}(N_j^{(n)} | X_0 = i)$ is the occupancy time up to n of state j starting from state i.
- occupancy times matrix: $\mathbf{M}^{(n)} = [m_{i,j}^{(n)}]$.

Theorem (Thm 2.4)

$$\mathbf{M}^{(n)} = \sum_{t=0}^{n} \mathbf{P}^{t}.$$

Occupancy time is the sum of all t-step transition matrices.

Proof : see textbook page 22. (or see next slide for proof of a general case.)

Generalization: Cost Models over a finite time *

Let $c(x) : S \to \mathbb{R}$ be a cost function. The expectation of the total cost up to time n is

$$C_i^{(n)} \triangleq \mathrm{E}\left[\sum_{t=0}^{n} c(X_t) | X_0 = i\right].$$

For a general $c(x)$, the calculation is below (Theorem 2.11 (page 35)).

$$\begin{aligned}
C_i^{(n)} &= \sum_{t=0}^{n} \mathrm{E}\left[c(X_t)(\sum_{j\in S} 1_{\{X_t=j\}}) \mid X_0 = i\right] (\because \text{ law of total prob.}) \\
&= \sum_{t=0}^{n} \sum_{j\in S} \mathrm{E}\left[c(X_t) \mid X_t = j, X_0 = i\right] \Pr\left(X_t = j \mid X_0 = i\right) (\because \text{ cond. prob.}) \\
&= \sum_{t=0}^{n} \sum_{j\in S} c(j) p_{ij}^{(t)} = \sum_{j\in S} c(j) \sum_{t=0}^{n} p_{ij}^{(t)}
\end{aligned}$$

*Section 2.6.1

Generalization: Cost Models over a finite time *

Let $c(x): S \to \mathbb{R}$ be a cost function. The expectation of the total cost up to time n is

$$C_i^{(n)} \triangleq \mathrm{E}\left[\sum_{t=0}^{n} c(X_t) | X_0 = i\right].$$

For a general $c(x)$, the calculation is below (Theorem 2.11 (page 35)).

$$\begin{aligned} C_i^{(n)} &= \sum_{t=0}^{n} \mathrm{E}\left[c(X_t)(\sum_{j\in S} 1_{\{X_t=j\}}) \mid X_0 = i\right] (\because \text{ law of total prob.}) \\ &= \sum_{t=0}^{n} \sum_{j\in S} \mathrm{E}\left[c(X_t) \mid X_t = j, X_0 = i\right] \Pr\left(X_t = j \mid X_0 = i\right) (\because \text{ cond. prob.}) \\ &= \sum_{t=0}^{n} \sum_{j\in S} c(j) p_{ij}^{(t)} = \sum_{j\in S} c(j) \sum_{t=0}^{n} p_{ij}^{(t)} \end{aligned}$$

The equivalent matrix-vector multiplication form is ($c = (c(1), c(2), \cdots, c(N))^{\mathsf{T}}$ is column vector)

$$\boxed{C^{(n)} = \left(\sum_{t=0}^{n} \mathbf{P}^t\right) c} \qquad (3)$$

*Section 2.6.1

In particular, if $c(\cdot) : S \to \mathbb{R}$ is the indicator function $1_{\{j\}}(\cdot)$, then by definition,

$$C_i^{(n)} = \mathrm{E}\left[\sum_{t=0}^{n} 1_j(X_t) | X_0 = i\right] = \mathrm{E}\left[N_j^{(n)} | X_0 = i\right] = m_{i,j}^{(n)}$$

which is the occupancy time of state j. On the other hand, the previous slide told us that the above line is actually equal to

$$\sum_{j' \in S} c(j') \sum_{t=0}^{n} p_{ij'}^{(t)} = \sum_{j' \in S} 1_{\{j\}}(j') \sum_{t=0}^{n} p_{ij'}^{(t)} = \sum_{t=0}^{n} p_{ij}^{(t)}$$

So, we have shown [Thm 2.4]:

$$\mathbf{M}^{(n)} = \sum_{t=0}^{n} \mathbf{P}^t.$$

Then (3) can be written in terms of $\mathbf{M}^{(n)}$:

$$\boxed{C^{(n)} = \mathbf{M}^{(n)} c}$$

Homework

- Let X_n be the random walk on $S = \mathbb{Z}$ with transition probability $\Pr(X_{n+1} = i+1|X_n) = p$ and $\Pr(X_{n+1} = i-1|X_n) = q = 1-p$.
  Calculate the mean $\mathrm{E}\,X_n$, the variance $\mathrm{var}\,X_n$ and the autocovariance $\mathrm{E}[(X_n - \mathrm{E}\,X_n)(X_m - \mathrm{E}\,X_m)]$.
- Assume that $\{X_n\}$, $n = 0, 1, 2, \cdots$, are iid $\{-1, 1\}$-valued random variables. $\Pr(X_i = 1) = p$ and $\Pr(X_i = -1) = 1-p$. For any $n \geq 1$, define new random variables
  $$S_n \triangleq \sum_{i=0}^{n} X_i, \;\; Y_n \triangleq X_n + X_{n-1},$$
  and the two-dim random vector $Z_n \triangleq \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix}$. Discuss if the stochastic processes $\{S_n\}$, $\{Y_n\}$, $\{Z_n\}$, are Markov chains. Why? Write the corresponding transition matrices for Markov chains.
- textbook page 51: 2.13, 2.16
- textbook page 53-54: 2.10, 2.11, 2.15(b). *

*MATLAB command for power of a matrix is : A^n.