Chapter 6: Markov Processes and Discrete-Time Markov Chains



OUTLINE

- 6.1 Markov Processes
- 6.2 Chapman-Kolmogorov Equation
- 6.3 Basic Concepts of Markov Chains(MCs)
- 6.4 Classification of States
- 6.5 Ergodic MC and Stationary Distribution



- The behavior of a Markov chain depends on the structure of the transition matrix P.
- We can classify a state space into different subsets according to the character of states.

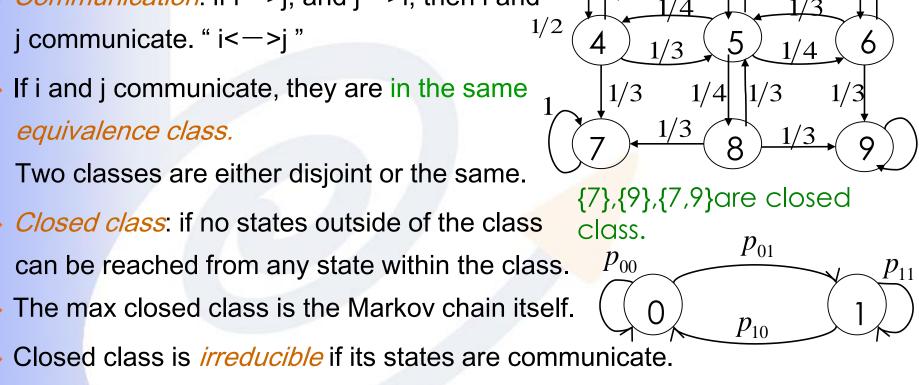


1/2

1/3

1/3

- Accessible: state j is said to be accessible from i if for some $n \ge 0, p_{ii}^{(n)} > 0$. "i->j"
- *Communication*: if i->j, and j->i, then i and j communicate. " i<->j "
- If i and j communicate, they are in the same equivalence class.
- Closed class: if no states outside of the class
- The max closed class is the Markov chain itself.
- Irreducible Markov chain: its only irreducible closed class is the set of states 3 in its state space S.





e.g. 1: Consider a Markov chain with transition probability matrix:

A random walk with two absorbing barriers.

$$P = \begin{bmatrix} 0 & 1 & 2 & \cdots & N-1 & N \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ N-1 & 0 & 0 & 0 & 0 & q & 0 & p \\ N & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where p + q = 1 and p > 0 and q > 0.

The chain has three classes {0}, {*N*}, and {1,2,...N-1}.

Closed classes: {0}, {*N*}, {0,*N*}



First passage time T_{ii} is the first passage time from *i* to *j*.

First entrance probability $f_{ij}^{(n)}$ is the probability mass function for T_{ij}

$$f_{ij}^{(n)} = P\{T_{ij} = n\} = P\{X_n = j, X_{n-1} \neq j, ..., X_1 \neq j | X_0 = i\}$$

$$n = 1, 2,$$

Recurrence time of state j: T_{ii}

Reaching probability from i to j f_{ij} is the probability that j is ever reached from i.

$$f_{ij} = P\{T_{ij} < \infty\} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$



If $f_{jj} = 1$ state j is called recurrent and $P\{T_{jj} < \infty\} = 1$

 $\mu_{jj} = E[T_{jj}]$ denote the mean recurrence time of state j.

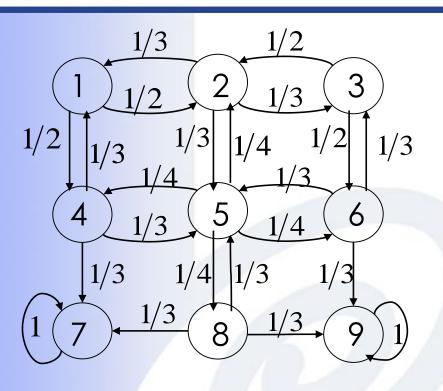
When the state space is infinite it is not necessary that $E[T_{ii}] < \infty$

If $\mu_{ii} < \infty$ state j is called positive recurrent.

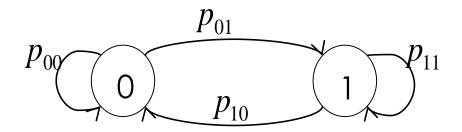
If $\mu_{ii} = \infty$ state j is called *null recurrent*.

If $f_{ii} < 1$ state j is called *transient*.





- ◆ 1,2,3,4,5,6,8 are transient
- 7,9 are positive recurrent



0,1 are positive recurrent

The states in an irreducible

Markov Chain with finite

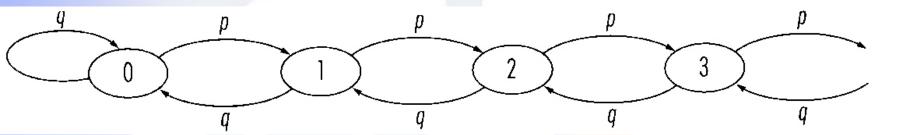
state

Ispace are positive recurrent.



Example: (example 4.3.1)

A random walk with a reflecting barrier where p>0,q>0,p+q=1



- All states are communicate, hence the chain is irreducible.
- q>p, the chain is positive recurrent. q <p, the chain is transient. q=p, the chain is null recurrent.

Chapter 6: Markov Processes and Discrete-Time Markov Chains



OUTLINE

- 6.1 Markov Processes
- 6.2 Chapman-Kolmogorov Equation
- 6.3 Basic Concepts of Markov Chains(MC)
- 6.4 Classification of States
- 6.5 Ergodic MC and Stationary Distribution



- Limiting distributions
- Periodic Markov chains
- 3. Ergodic Markov chains
- 4. Stationary distributions



Limiting distribution:

Example 1:

t = day index 0, 1, 2, ...

$$X_t = 0$$
 rainy on t^{th} day $= 1$ sunny on t^{th} day

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix}$$

$$\mathbf{P}^{(2)} = \begin{bmatrix} p_{00}^{(2)} & p_{01}^{(2)} \\ p_{10}^{(2)} & p_{11}^{(2)} \end{bmatrix} = \begin{bmatrix} 7/16 & 9/16 \\ 3/8 & 5/8 \end{bmatrix} = \begin{bmatrix} 0.4375 & 0.5625 \\ 0.375 & 0.625 \end{bmatrix}$$

$$\mathbf{P}^{(4)} = \begin{bmatrix} 0.40234375 & 0.59765625 \\ 0.3984375 & 0.6015625 \end{bmatrix} \quad \mathbf{P}^{(8)} = \begin{bmatrix} p_{00}^{(8)} & p_{01}^{(8)} \\ p_{10}^{(8)} & p_{11}^{(8)} \end{bmatrix} \approx \begin{bmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{bmatrix}$$

$$\mathbf{P}^{(8)} = \begin{bmatrix} p_{00}^{(8)} & p_{01}^{(8)} \\ p_{10}^{(8)} & p_{11}^{(8)} \end{bmatrix} \approx \begin{bmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{bmatrix}$$

$$\lim_{m \to \infty} \mathbf{P}^{(m)} = \begin{bmatrix} p_{00}^{(m)} & p_{01}^{(m)} \\ p_{10}^{(m)} & p_{11}^{(m)} \end{bmatrix} = \begin{bmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{bmatrix} = \begin{bmatrix} 1/\mu_0 & 1/\mu_1 \\ 1/\mu_0 & 1/\mu_1 \end{bmatrix}$$



Example 2:



Limiting distribution (independent with i):

$$\lim_{m\to\infty} p_{ij}^{(m)} = \frac{1}{\mu_j}, \ j \in S$$

Example 1:

$$\lim_{m \to \infty} p_{i0}^{(m)} = \frac{1}{\mu_0} = 0.4 \qquad \lim_{m \to \infty} p_{i1}^{(m)} = \frac{1}{\mu_1} = 0.6$$

Example 2:

$$\lim_{m \to \infty} p_{i0}^{(m)} = \frac{1}{\mu_0} = 0.5 \qquad \lim_{m \to \infty} p_{i1}^{(m)} = \frac{1}{\mu_1} = 0.5$$

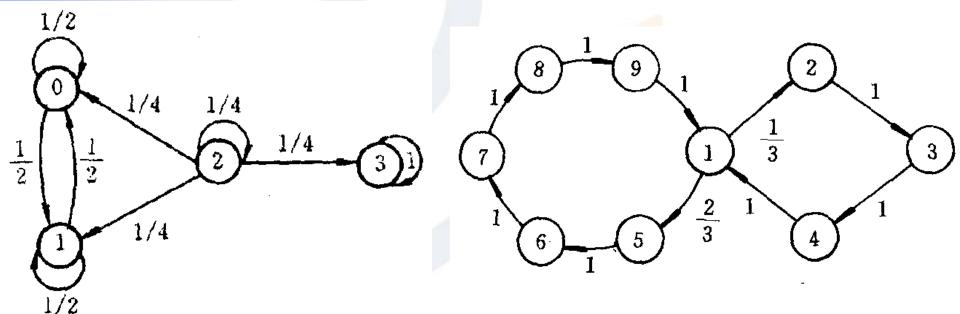


2. Periodic Markov chain: the period of state i:

d(i) denote the greatest common divisor of all integers $n \ge 1$ for which $p_{ii}^{(n)} > 0$.

The integer d(i) is called the period of state i.

- If d(i)=1, state i is called aperiodic.
- If $p_{ii}^{(1)} > 0$, then d(i)=1, state *i* is aperiodic.





♦ If d is the period of state i, for all $n(\text{mod})d \neq 0$

$$p_{ii}^{(n)} = 0$$

- If one state in a class is periodic, all states in the class are periodic.
- Aaperiodic Markov chain :

An irreducible Markov chain whose states have a period of 1 is called an aperiodic chain.

$$P_{00} \qquad P_{01} \qquad P_{11} \qquad P = \begin{bmatrix} 0 & 1 & 0 \\ 0.4 & 0 & 0.6 \\ 0 & 0.7 & 0.3 \end{bmatrix} \quad P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}$$



Ergodic Markov chain

- An irreducible aperiodic Markov chain with finite number of states is called an *Ergodic Markov chain*.
- An ergodic Markov chain is irreducible.
- All states in an ergodic Markov chain are Communicate and Aperiodic.

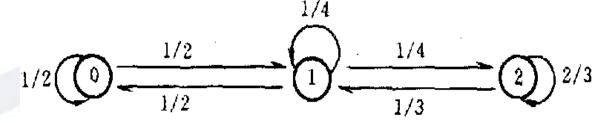
Communicate: For some $m \ge 1$, $p_{ij}^{(m)} \ne 0$, $i,j \in S$

$$\mathbf{P} = \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix}$$

Example 2:

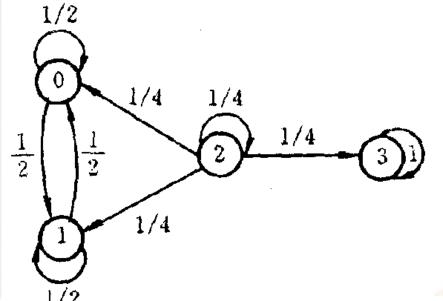
$$\mathbf{P} = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}$$

Example 3:
$$\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$





◆ Example 4:



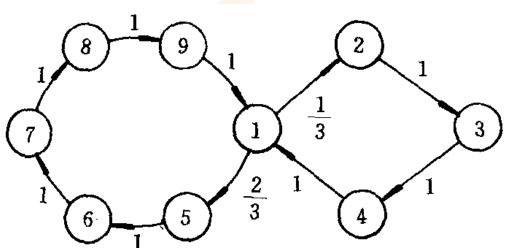
{0,1},{3},{0,1,3} are closed class.

{0,1} are irreducible closed class.

Example 5:

All states in state space are communicate.

A periodic irreducible Markov chain with d(i)=2





Theorem:

The limiting distribution exists if the Markov chain is ergodic.



Stationary distribution:

Finding the limiting distribution:

If the limiting distribution of a Markov chain exists, it can be write

as

$$\lim_{m o \infty} P^{(m)} = egin{bmatrix} \pi_0 & \cdots & \pi_i & \pi_N \ \pi_0 & \cdots & \pi_i & \pi_N \ \cdots & \cdots & \cdots \ \pi_0 & \cdots & \pi_i & \pi_N \end{bmatrix}$$

After a period of time the absolute distribution is not change with

$$\mathbf{s}(m) = \mathbf{s}(0)\mathbf{P}^{(m)} = \begin{bmatrix} \pi_0 & \cdots & \pi_i & \pi_N \end{bmatrix}$$

$$\mathbf{s}(m+1) = \mathbf{s}(0)\mathbf{P}^{(m+1)} = \begin{bmatrix} \pi_0 & \cdots & \pi_i & \pi_N \end{bmatrix}$$

$$\mathbf{s}(m+1) = \mathbf{s}(m)\mathbf{P} = s(m)$$



Stationary distribution:

The probability distribution $\{\pi_i, j \in S\}$ is called stationary distribution of Markov process, if it satisfies:

$$\begin{cases} \pi_{j} = \sum_{i \in S} \pi_{i} p_{ij} \\ \sum_{j \in S} \pi_{j} = 1, \ \pi_{j} \geq 0 \end{cases} \quad \text{or} \quad \begin{cases} \pi = \pi \mathbf{P} \\ \sum_{j \in S} \pi_{j} = 1, \ \pi_{j} \geq 0 \end{cases}$$

$$\begin{cases} \boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P} \\ \sum_{j \in S} \boldsymbol{\pi}_j = 1, \ \boldsymbol{\pi}_j \ge 0 \end{cases}$$

Theorem:

The unique stationary distribution exists if the Markov chain is ergodic. For any $i, j \in S$, the stationary

distribution is given by
$$\pi_j = \lim_{m \to \infty} p_{ij}^{(m)} = \frac{1}{\mu_j}, \quad i,j \in S$$



First passage time T_{ij} is the first passage time from *i* to *j*.

First entrance probability $f_{ij}^{(n)}$ is the probability mass function for T_{ii}

$$f_{ij}^{(n)} = P\{T_{ij} = n\} = P\{X_n = j, X_{n-1} \neq j, ..., X_1 \neq j | X_0 = i\}$$

Mean of fist passage time:

$$\mu_{ij} = E[T_{ij}] = \sum_{n=1}^{\infty} n f_{ij}^{(n)}$$

Recurrence time of state j: T_{jj}

Mean of Recurrence time of state j:

$$\mu_j = \mu_{jj} = E[T_{jj}] = \sum_{n=1}^{\infty} n f_{jj}^{(n)}$$



◆ Example 1:

$$\lim_{m \to \infty} p_{i0}^{(m)} = \frac{1}{\mu_0} = 0.4 \qquad \lim_{m \to \infty} p_{i1}^{(m)} = \frac{1}{\mu_1} = 0.6$$

$$\mathbf{P} = \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix} \quad \begin{bmatrix} 0.4 & 0.6 \end{bmatrix} \cdot \begin{bmatrix} 1/4 & 3/4 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.6 \end{bmatrix}$$

Example 2:

$$\lim_{m \to \infty} p_{i0}^{(m)} = \frac{1}{\mu_0} = 0.5 \qquad \lim_{m \to \infty} p_{i1}^{(m)} = \frac{1}{\mu_1} = 0.5$$

$$\mathbf{P} = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \quad \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \cdot \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$$



Example 3:

$$\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

$$1/2 \bigcirc 0 = \frac{1/2}{1/2} \bigcirc 1/4 = \frac{1/4}{1/3} \bigcirc 2/3$$

$$\begin{bmatrix} \pi_0 & \pi_1 & \pi_2 \end{bmatrix} = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 \end{bmatrix} \cdot \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

$$\sum_{j \in \mathbb{Z}} \pi_j = 1, \ \pi_j \ge 0$$

$$\begin{cases} \pi_0 = 1/2 \cdot \pi_0 + 1/2 \cdot \pi_1 & \pi_0 = 8/19 \\ \pi_1 = 1/2 \cdot \pi_0 + 1/4 \cdot \pi_1 + 1/3 \cdot \pi_2 & \pi_1 = 8/19 \\ \pi_2 = 1/4 \cdot \pi_1 + 2/3 \cdot \pi_2 & \pi_2 = 3/19 \\ \pi_0 + \pi_1 + \pi_2 = 1, \ \pi_i \ge 0 \end{cases}$$



A Markov chain is strictly stationary if and only if its absolute state probabilities do not depend on time.

An ergodic Markov chain is not a stationary process unless:

A Markov chain is strictly stationary if and only if its initial distribution is stationary.



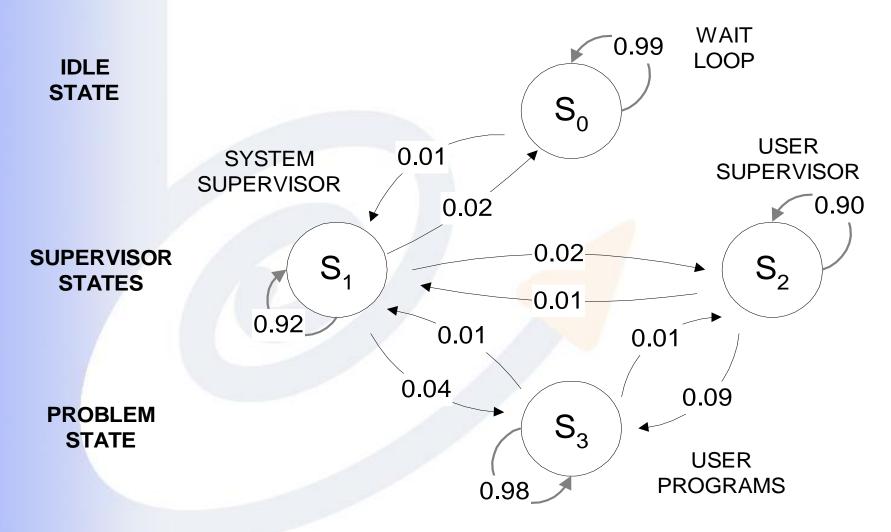
Example 4:

CPU of a multiprogramming system is at any time executing instructions from:

- User program ==> Problem State (S₃)
- OS routine explicitly called by a user program (S₂)
- OS routine performing system wide control task (S₁)
 ==> Supervisor State
- Wait loop ==> Idle State (S₀)



State Transition graph of discrete-time Markov of a CPU





	\sim 1
\Box	Stata
$\mathbf{I} \mathbf{U}$	State

		S_0	S ₁	S_2	S_3
	S ₀	0.99	0.01	0	0
From	S ₁	0.02	0.92	0.02	0.04
State	S ₂	0	0.01	0.90	0.09
	S ₃	0	0.01	0.01	0.98

Transition Probability Matrix



$$\begin{split} \pi_0 &= 0.99 \, \pi_0 + 0.02 \, \pi_1 \\ \pi_1 &= 0.01 \, \pi_0 + 0.92 \, \pi_1 + 0.01 \pi_2 + 0.01 \pi_3 \\ \pi_2 &= 0.02 \pi_1 + 0.90 \pi_2 + 0.01 \, \pi_3 \\ \pi_3 &= 0.04 \pi_1 + 0.09 \, \pi_2 + 0.98 \, \pi_3 \\ 1 &= \pi_0 + \pi_1 + \pi_2 + \pi_3 \end{split}$$

Stationary state probabilities can be computed by solving system of equations. So we have:

$$\pi_0 = 2/9$$
, $\pi_1 = 1/9$, $\pi_2 = 8/99$, $\pi_3 = 58/99$



$$S_0$$
: waiting loop; $\pi_0 = 2/9$

$$S_1$$
: system supervisor; $\pi_1 = 1/9$

$$S_2$$
: user supervisor; $\pi_2 = 8/99$

$$S_3$$
: user programs; $\pi_3 = 58/99$

- Utilization of CPU: 1 π_0 = 77.7%
- 58.6% of total time spent for processing users programs (S₃).
- ◆ 19.1% (77.7 58.6) of time spent in supervisor state.



Finding the n-step transition probability matrix

$$|P - \lambda I| = 0$$

$$P'' = Q \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} Q^{-1}$$

$$P'' = Q \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} Q^{-1}$$

Example: Two-state Markov chain

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \qquad \frac{\lambda_1 = 1}{\lambda_2 = p_{00} + p_{11} - 1}$$

$$Q = \begin{pmatrix} 1 & p_{01} \\ 1 & -p_{10} \end{pmatrix} \qquad Q^{-1} = \frac{1}{p_{01} + p_{10}} \begin{pmatrix} p_{10} & p_{01} \\ 1 & -1 \end{pmatrix}$$





$$P(n) = P^{n} = \frac{1}{p_{01} + p_{10}} \begin{pmatrix} 1 & p_{01} \\ 1 & -p_{10} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (p_{00} + p_{11} - 1)^{n} \end{pmatrix} \begin{pmatrix} p_{10} & p_{01} \\ 1 & 1 \end{pmatrix}$$

$$= \frac{1}{p_{01}^{2} + p_{10}} \begin{pmatrix} p_{10} & p_{01} \\ p_{10} & p_{01} \end{pmatrix} + \frac{(p_{00} + p_{11} - 1)^{n}}{p_{01} + p_{10}} \begin{pmatrix} p_{01} & -p_{01} \\ -p_{10} & -p_{10} \end{pmatrix}$$



Summary of Markov Chain

Important concepts (1):

One-step transition probabilities:

$$p_{ij}(n) = P\{X_{n+1} = j \mid X_n = i\}; \ n = 0,1,...$$

Homogeneity:

$$p_{ij}(n) = P\{X_1 = j \mid X_0 = i\} = p_{ij}$$
 for all $n = 0,1,...$

One-step transition probability matrix: P

M-step transition probabilities:

$$p_{ij}^{(m)} = P\{X_{n+m} = j \mid X_n = i\}, \quad m = 0,1,...$$

M-step transition probability matrix: $\mathbf{P}^{(m)}$



Important concepts (2):

Initial distribution s(0):

A probability distribution of $X(t_0)$

Absolute distribution $\mathbf{s}(m)$:

One-dimensional state probabilities of the Markov chain after m steps.

$$\mathbf{s}(m) = \{ p_j^{(m)} = P(X_m = j), j \in S, \sum_{j \in S} p_j = 1 \}$$

Limiting distribution:

$$\lim_{m\to\infty} p_{ij}^{(m)} = \frac{1}{\mu_j}, \ j\in S$$



Important concepts (3):

Ergodic Markov chain: Communicate irreducible aperiodic Markov chain with finite number of states is called an *Ergodic Markov chain*.

For some
$$m \ge 1$$
, $p_{ij}^{(m)} \ne 0$, $i,j \in S$

Stationary distribution: $\{\pi_j, j \in S\}$

Stationary distribution exists when the Markov chain is

ergodic.
$$\begin{cases} \pi_j = \sum_{i \in Z} \pi_i p_{ij} \\ \sum_{j \in Z} \pi_j = 1, \ \pi_j \geq 0 \end{cases}$$



Important relationship (1)

Chapman-Kolmogorov equations

$$p_{ij}^{(m)} = \sum_{k \in \mathbb{Z}} p_{ik}^{(r)} p_{kj}^{(m-r)}, \quad r = 0,1,...,m$$

$$\mathbf{P}^{(m)} = \mathbf{P}\mathbf{P}^{(m-1)}$$

or
$$\mathbf{P}^{(m)} = \mathbf{P}^{(r)}\mathbf{P}^{(m-r)}, r = 0,1,...,m$$

$$\mathbf{P}^{(m)} = \mathbf{P}^m$$

Joint distribution:

$$P\{X_0 = i_0, X_1 = i_1, ..., X_n = i_n\} = s_{i_0}(0) p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}$$

Absolute distribution:

$$p_{j}(m) = \sum_{i \in S} s_{i}(0) p_{ij}^{(m)} = \sum_{i \in S} s_{i}(m-1) p_{ij}, m = 1,2,...$$

$$\mathbf{s}(m) = \mathbf{s}(0)\mathbf{P}^{(m)} = \mathbf{s}(m-1)\mathbf{P}$$



Important relationship (2)

The unique stationary distribution and limit distribution exist if the Markov chain is ergodic. For any $i, j \in S$,

$$\pi_j = \lim_{m \to \infty} p_{ij}^{(m)} = \frac{1}{\mu_j}, \quad i, j \in S$$

$$m{m}
ightarrow \infty, \qquad \mathbf{P}^{(m)} = egin{bmatrix} \pi_0 & \cdots & \pi_i & \cdots & \pi_N \ \cdots & \ddots & \ddots & \ddots \ \pi_0 & \cdots & \pi_i & \cdots & \pi_N \end{bmatrix}$$



Let $\{X_i; i = 0,1,...\}$ be a sequence of independent, identically distributed binary random variables with

$$P(X_i = 1) = P(X_i = -1) = 1/2$$

Let moving averages Y_n be defined as follows

$$Y_n = \frac{1}{2}(X_n + X_{n-1}); \quad n = 1, 2, \dots$$

Is $\{Y_1, Y_2, ...\}$ a Markov chain?

Ex. 2



- ◆ A wireless packet communications channel suffers from clustered errors. That is, whenever a packet has an error, the next packet will have an error with probability 0.9. When a packet is error-free, the next packet is error-free with probability 0.99.
- In steady state, what is the probability that a packet has an error?

Ex. 3



- ◆ The state of a process changes daily according to a two-state Markov chain. If the process is in state *i* during one day, then it is in state *j* the following day with probability P_{ij} , where $p_{00} = 0.4$, $p_{01} = 0.6$, $p_{10} = 0.2$, $p_{11} = 0.8$
- Every day a message is sent. If the state of the Markov chain that day is i then the message sent is "good" with probability P_i and is "bad" with probability $q_i = 1 p_i$, i = 0,1
- ◆ If the process is in state 0 on Monday, what is the probability that a good message is sent on Tuesday?
- ◆ If the process is in state 0 on Monday, what is the probability that a good message is sent on Friday?
- ◆ In the long run, what proportion of messages is good?
- Let $Y_n = 1$ if a good message is sent on day n and let $Y_n = 2$ otherwise. Is $\{Y_n, n \ge 1\}$ a Markov chain? If so, give its transition probability matrix. If not, briefly explain why not.

Ex.4



- Every week a professor gives exams to his students. He can give three possible types of exams, and his class is graded as either having done well or badly. Let p_i denote the probability that the class does well on a type i exam, and suppose that $p_1 = 0.3$, $p_2 = 0.6$, and $p_3 = 0.9$.
- If the class does well on an exam, then the next exam is equally likely to be any of the three types. If the class does badly, then the next exam is always type 1. Let X_n , $n \ge 1$ denote the exam type of the nth week.
- Find its transition probability matrix.
- ◆ If the exam type is 2 at first week, compute

$$P(X_1=2, X_2=1, X_3=3).$$



End of Chapter 6