Markov Chains

Y. İlker TOPCU, Ph.D.

www.ilkertopcu.net

Özgür KABAK, Ph.D.

web.itu.edu.tr/kabak/

We would like to acknowledge Prof. W.L. Winston's "Operations Research: Applications and Algorithms" (slides submitted by Brooks/Cole, a division of Thomson Learning, Inc.) and Prof. J.E. Beasley's lecture notes which greatly influence these notes...

We retain responsibility for all errors and would love to hear from readers...

Stochastic Process

- Sometimes we are interested in how a random variable changes over time: how the price of a share of stock or a firm's market share evolves
- The study of how a random variable evolves over time includes **Stochastic Process**es
- In particular, we focus on a special type of stochastic process known as a **Markov Chain.**

Discrete-Time Stochastic Process

- Suppose we observe some characteristic of a system at discrete points in time.
- Let $\mathbf{X}_t$ be the value of the system characteristic at time t.
- In most situations, $\mathbf{X}_t$ is not known with certainty before time t and may be viewed as a random variable.
- A **discrete-time stochastic process** is simply a description of the relation between the random variables $\mathbf{X}_{0}, \mathbf{X}_{1}, \mathbf{X}_2$ …..

An example of discrete-time stochastic process

- An urn contains two unpainted balls at present.
- We choose a ball at random and flip a coin.
- If the chosen ball is unpainted and the coin comes up heads, we paint the chosen unpainted ball red; if the chosen ball is unpainted and the coin comes up tails, we paint the chosen unpainted ball black.
- If the ball has already been painted, then (whether heads or tails has been tossed) we change the color of the ball (from red to black or from black to red).
- To model this situation as a stochastic process, we define time t to be the time after the coin has been flipped for the tth time and the chosen ball has been painted.

- The state at any time may be described by the vector [*u r b*], where
  - *u* is the number of unpainted balls in the urn,
  - *r* is the number of red balls in the urn, and
  - *b* is the number of black balls in the urn.
- We are given that $\mathbf{X}_0$ = [2 0 0].
- After the first coin toss, one ball will have been painted either red or black, and the state will be either [1 1 0] or [1 0 1]. Hence, $\mathbf{X}_1$ = [1 1 0] or $\mathbf{X}_1$ = [1 0 1].
- There must be some sort of relation between the $\mathbf{X}_t$'s.
- For example, if $\mathbf{X}_t$ = [0 2 0],

  we can be sure that $\mathbf{X}_{t+1}$ will be [0 1 1].

Continuous-Time Stochastic Process

- A **continuous–time stochastic process** is simply the stochastic process in which the state of the system can be viewed at any time, not just at discrete instants in time.
- EXAMPLE: The number of people in a supermarket *t* minutes after the store opens for business may be viewed as a continuous-time stochastic process.

What is a Markov Chain?

- A discrete-time stochastic process is a **Markov chain** if the probability distribution of the state at time t+1 depends on the state at time t(i_t) and does not depend on the states the chain passed through on the way to i_t at time t.
- $P(\mathbf{X}_{t+1} = i_{t+1} | \mathbf{X}_t = i_t, \mathbf{X}_{t-1} = i_{t-1}, \ldots, \mathbf{X}_1 = i_1, \mathbf{X}_0 = i_0)$

$$= P(\mathbf{X}_{t+1} = i_{t+1} | \mathbf{X}_t = i_t)$$

for t = 0,1,2… and all states

Transition

- We make further assumption that for all states *i* and *j* and all *t*, $P(\mathbf{X}_{t+1} = j|\mathbf{X}_t = i)$ is independent of *t*.
- This assumption allows us to write
  $P(\mathbf{X}_{t+1} = j|\mathbf{X}_t = i) = p_{ij}$
  where p_{ij} is the probability that given the system is in state *i* at time *t*, it will be in a state *j* at time *t*+1.
- If the system moves from state *i* during one period to state *j* during the next period, we say that a **transition** from *i* to *j* has occurred.

Stationary Markov Chain

- The p_{ij}'s are often referred to as the **transition probabilities** for the Markov chain.
- $P(\mathbf{X}_{t+1} = j|\mathbf{X}_t = i) = p_{ij}$ implies that the probability law relating the next period's state to the current state does not change over time.
- This equation is called the **Stationary Assumption** and any Markov chain that satisfies it is called a **stationary Markov chain**.

Transition Matrix

- We also must define q_i to be the probability that the chain is in state *i* at the time 0; $P(\mathbf{X}_0=i) = q_i$.
- The vector **q**= $[q_1, q_2, \ldots q_s]$ is called as the **initial probability distribution** for the Markov chain.
- The transition probabilities are displayed as an *s* x *s* **transition probability matrix** *P:*

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{11} & \cdots & p_{11} \\ \vdots & \vdots & & \vdots \\ p_{s1} & p_{s2} & \cdots & p_{ss} \end{bmatrix}$$

- Given that the state at time *t* is *i*, the process must be somewhere at time *t*+1.
- For each *i*

$$\sum_{j=1}^{j=s} p_{ij} = 1$$

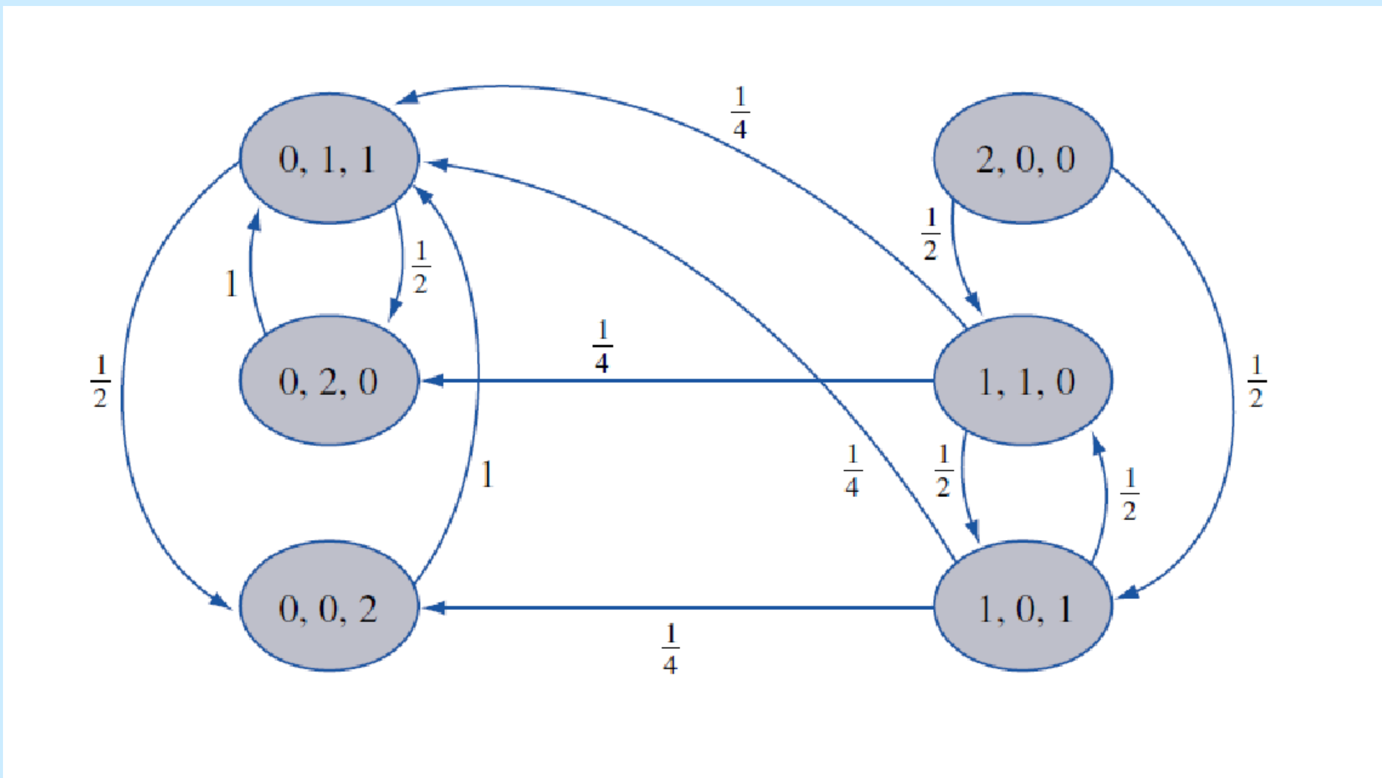
- We also know that each entry in the *P* matrix must be nonnegative.
- Hence, all entries in the transition probability matrix are nonnegative, and the entries in each row must sum to 1.

Ball choosing example

- Since the state of the urn after the next coin toss only depends on the past history of the process through the state of the urn after the current coin toss, we have a Markov chain. The rules don't change over time, so it's a stationary Markov chain
- To illustrate the formation of the transition matrix, let's determine the [1 1 0] row of this matrix.
  If current state is [1 1 0], transition probabilities will be as
  [# of unpainted balls, # or red balls, # of black balls]

Event	Probability	New state
Flip heads and choose unpainted ball	1/4	[0 2 0]
Flip tails and choose unpainted ball	1/4	[0 1 1]
Choose red ball	1/2	[1 0 1]

$$
P = \begin{array}{c|cccccc|}
 & \multicolumn{6}{c}{\text{State}} \\
 & [0\ 1\ 1] & [0\ 2\ 0] & [0\ 0\ 2] & [2\ 0\ 0] & [1\ 1\ 0] & [1\ 0\ 1] \\
\hline
[0\ 1\ 1] & 0 & 1/2 & 1/2 & 0 & 0 & 0 \\
[0\ 2\ 0] & 1 & 0 & 0 & 0 & 0 & 0 \\
[0\ 0\ 2] & 1 & 0 & 0 & 0 & 0 & 0 \\
[2\ 0\ 0] & 0 & 0 & 0 & 0 & 1/2 & 1/2 \\
[1\ 1\ 0] & 1/4 & 1/4 & 0 & 0 & 0 & 1/2 \\
[1\ 0\ 1] & 1/4 & 0 & 1/4 & 0 & 1/2 & 0
\end{array}
$$



[# of unpainted balls, # or red balls, # of black balls]

n-Step Transition Probabilities

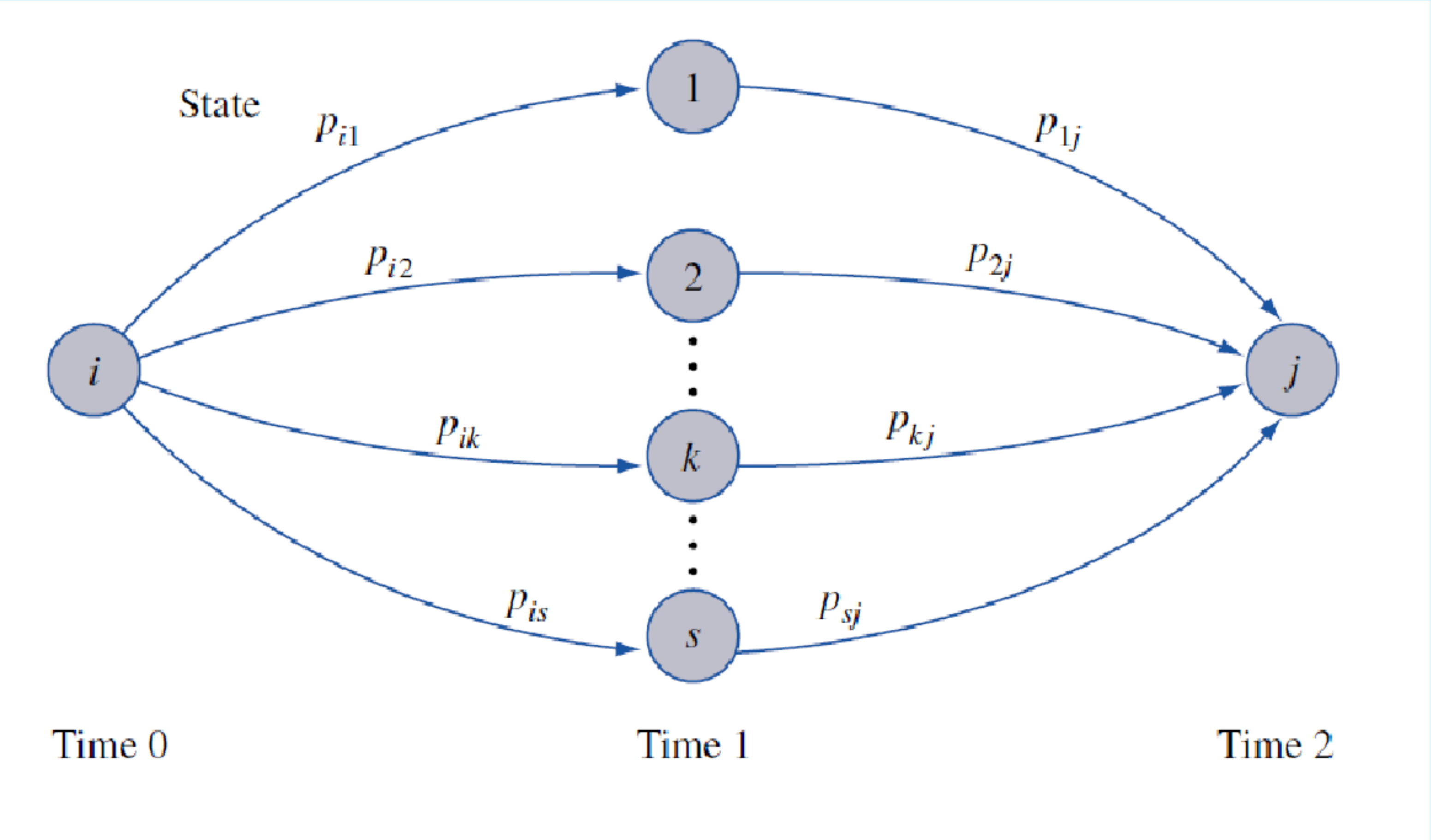
- Suppose we are studying a Markov chain with a known transition probability matrix P.
- If a Markov chain is in a state i at time m, what is the probability that n periods later than the Markov chain will be in state j?
- This probability will be independent of m, so:

$$P(\mathbf{X}_{m+n} = j | \mathbf{X}_m = i) = P(\mathbf{X}_n = j | \mathbf{X}_0 = i) = P_{ij}(n)$$

- $P_{ij}(n)$ is called the ***n*-step probability** of a transition from state i to state j.
- $P_{ij}(1) = p_{ij}$

n-Step Transition Probabilities

- $P_{ij}(2) = \sum_{k=1}^{s} p_{ik}p_{kj}$



n-Step Transition Probabilities

- The right-hand side of $P_{ij}(2) = \sum_{k=1}^{s} p_{ik}p_{kj}$ is just the scalar product of row i of the P matrix with column j of the P matrix.
- Hence, $P_{ij}(2)$ is the ij th element of the matrix P^2.
- By extending this reasoning, it can be shown that for $n > 1$

$$P_{ij}(n) = ij\text{th element of } P^n$$

The Cola Example

- Suppose the entire cola industry produces only two colas.
- Given that a person last purchased cola A, there is a 90% chance that her next purchase will be cola A.
- Given that a person last purchased cola B, there is an 80% chance that her next purchase will be cola B.
- Answer the following questions:
    1. If a person is currently a cola B purchaser, what is the probability that she will purchase cola A two purchases from now?
    2. If a person is currently a cola A purchaser, what is the probability that she will purchase cola A three purchases from now?

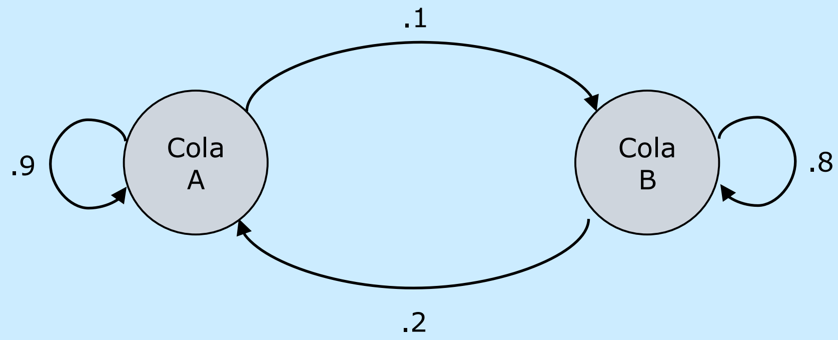
The states for cola example

- We view each person’s purchases as a Markov chain with the state at any given time being the type of cola the person last purchased.
- Hence, each person’s cola purchases may be represented by a two-state Markov chain, where
  - State 1 = person has last purchased cola A
  - State 2 = person has last purchased cola B

The transition matrix and graphical representation

- If we define $\mathbf{X}_n$ to be the type of cola purchased by a person on her nth future cola purchase, then $\mathbf{X}_0$, $\mathbf{X}_1$, … may be described as the Markov chain with the following transition matrix and graphical representation:

$$P = \begin{array}{cc} & \begin{array}{cc}\text{Cola A} & \text{Cola B}\end{array} \\ \begin{array}{c}\text{Cola A} \\ \text{Cola B}\end{array} & \begin{bmatrix} 90\% & 10\% \\ 20\% & 80\% \end{bmatrix} \end{array}$$



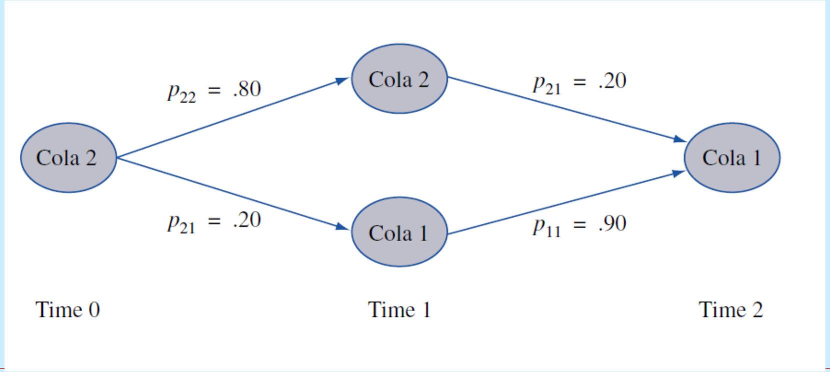
Answer to question 1

- If a person is currently a cola B purchaser, what is the probability that she will purchase cola A two purchases from now?

$$P^2 = \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix} \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix} = \begin{bmatrix} .83 & .17 \\ .34 & .66 \end{bmatrix}$$

$P(\mathbf{X}_2 = 1 | \mathbf{X}_0 = 2) = P_{21}(2) = 34\%$

- We may obtain this answer in a different way
$P_{21}(2)$ = prob. that 1st purchase is cola B and 2nd is cola A + (prob. that 1st purchase is cola A and 2nd is cola A)
$= p_{22}p_{21} + p_{21}p_{11} = (.80)(.20) + (.20)(.90) + = .34.$



Answer to question 2

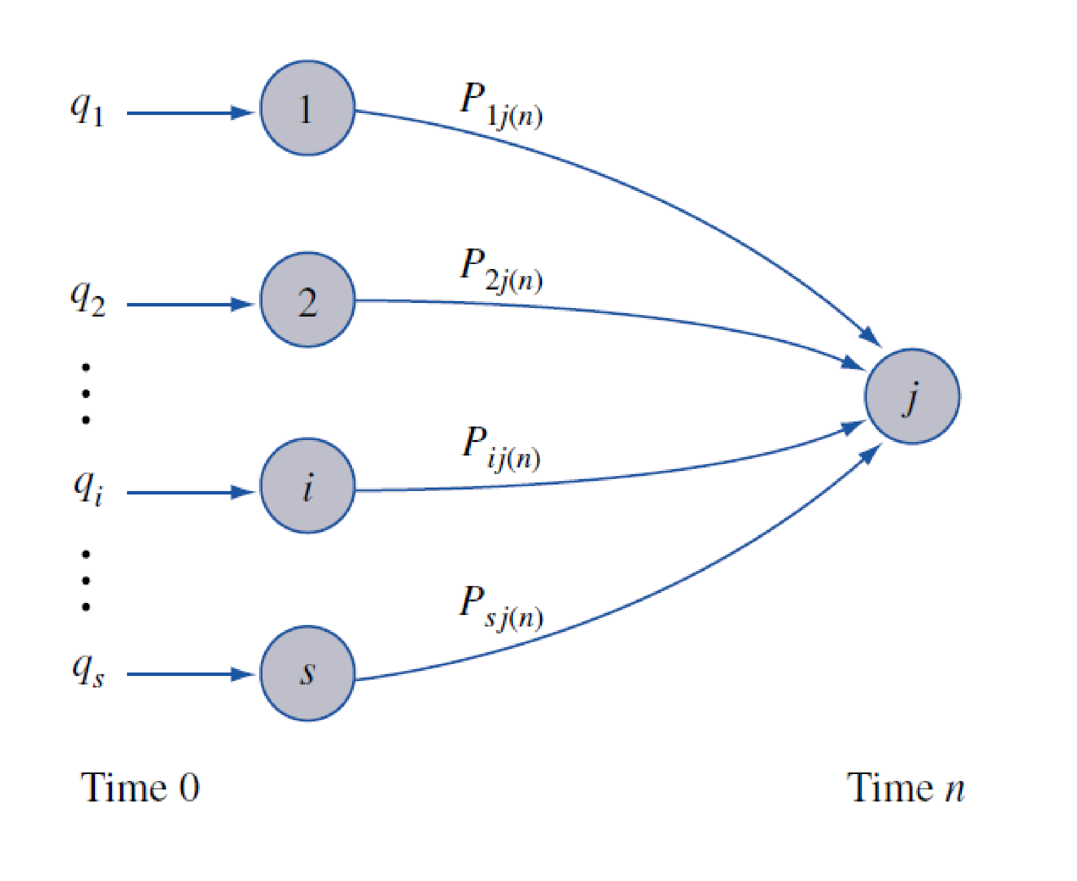
- If a person is currently a cola A purchaser, what is the probability that she will purchase cola A three purchases from now?

$$P^3 = P(P^2) = \begin{bmatrix} .90 & .10 \\ .20 & .80 \end{bmatrix} \begin{bmatrix} .83 & .17 \\ .34 & .66 \end{bmatrix} = \begin{bmatrix} .781 & .219 \\ .438 & .562 \end{bmatrix}$$

Therefore, $P_{11}(3)$ = 78.1 %

Behavior of *n*-Step Transition Probabilities

- In many situations, we do not know the state of the Markov chain at time 0.



Behavior of *n*-Step Transition Probabilities

- Probability of being in state *j* at time *n*

$$= \sum_{i=1}^{i=s} q_i P_{ij}(n)$$

$$= \mathbf{q} \text{ (column } j \text{ of } P^n)$$

where

q_i is the probability that state is originally *i*

$P_{ij}(n)$ is the probability of going from *i* to *j* in *n* transitions

$\mathbf{q} = [q_1 \ \ q_2 \ \ \ldots \ \ q_s]$

The Cola Example extended

- Suppose 60% of all people now drink cola A, and 40% now drink cola B.
- Three purchases from now, what fraction of all purchasers will be drinking cola A?

ANSWER

- The desired probability is

$$[.60 \quad .40] \begin{bmatrix} .781 \\ .438 \end{bmatrix} = 64.38\%$$

Current position

Column 1 of P^3

- Hence, three purchases from now, 64% of all purchasers will be purchasing cola A

The behavior of the *n*-step transition probabilities for large values of *n*

- *n*-step transition probabilities for cola drinkers

n	$P_{11}(n)$	$P_{12}(n)$	$P_{21}(n)$	$P_{22}(n)$
1	0.90	0.10	0.20	0.80
2	0.83	0.17	0.34	0.66
3	0.78	0.22	0.44	0.56
4	0.75	0.25	0.51	0.49
5	0.72	0.28	0.56	0.44
10	0.68	0.32	0.65	0.35
20	0.67	0.33	0.67	0.33
30	0.67	0.33	0.67	0.33

- This means that for large *n,* no matter what the initial state, there is a 67% chance that a person will be a cola A purchaser.

Classification of States in a Markov Chain

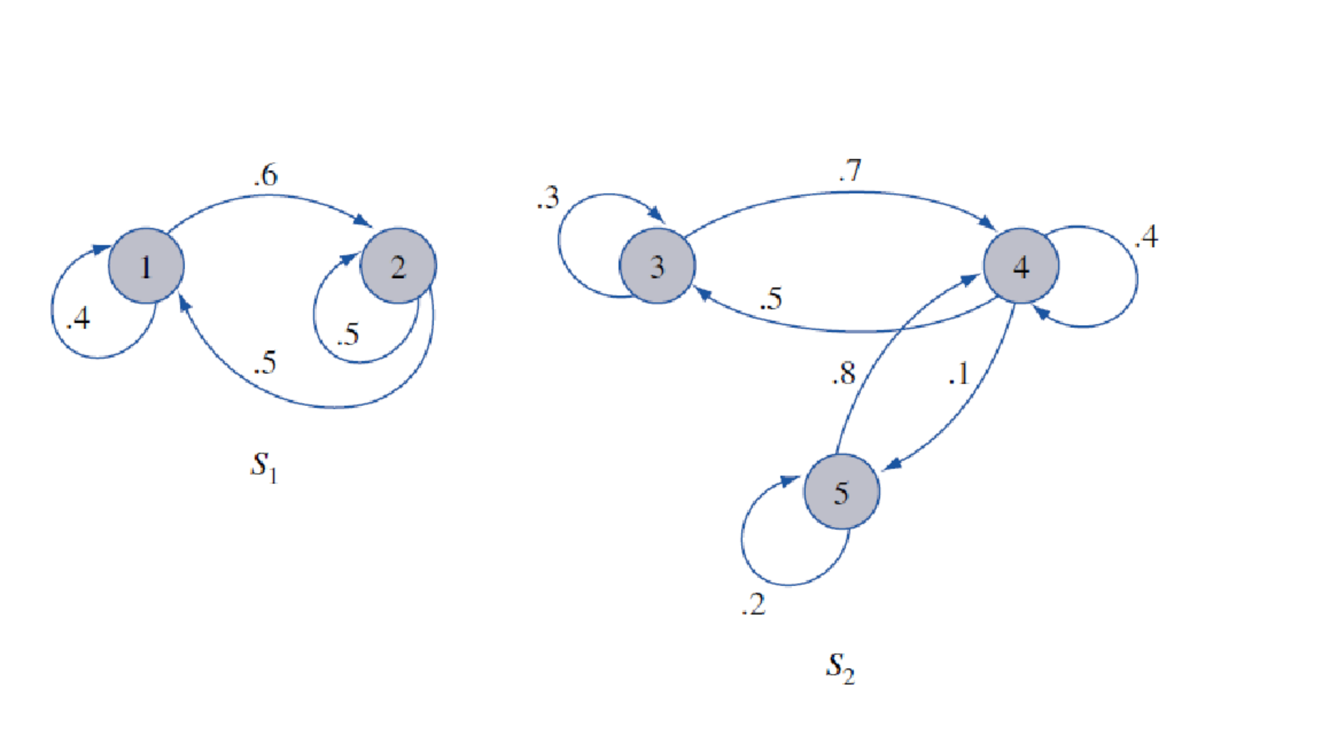
- To understand the n-step transition in more detail, we need to study how mathematicians classify the states of a Markov chain.

Path, Reach, Communnication

- Given two states of *i* and *j*, a **path** from *i* to *j* is a sequence of transitions that beging in *i* and ends in *j*, such that each transition in the sequence has a positive probability of occurring.
- A state *j* is **reachable** from state *i* if there is a path leading from *i* to *j*.
- Two states *i* and *j* are said to **communicate** if *j* is reachable from *i*, and *i* is reachable from *j*.

Example

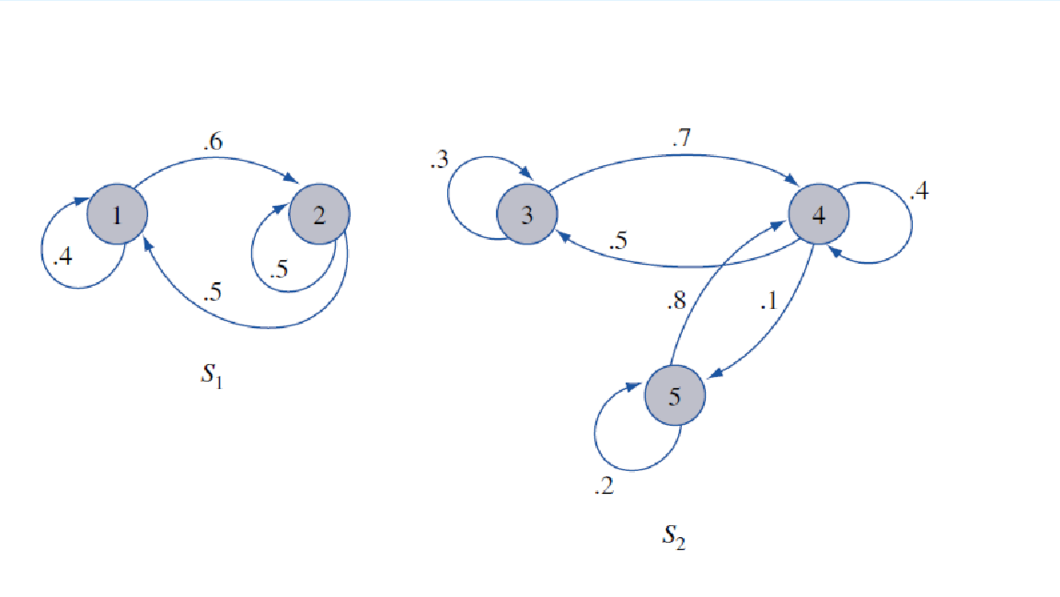
$$P = \begin{bmatrix} .4 & .6 & 0 & 0 & 0 \\ .5 & .5 & 0 & 0 & 0 \\ 0 & 0 & .3 & .7 & 0 \\ 0 & 0 & .5 & .4 & .1 \\ 0 & 0 & 0 & .8 & .2 \end{bmatrix}$$



- State 5 is <u>reachable</u> from state 3 (via the <u>path</u> 3–4–5)
- State 5 is not reachable from state 1 (there is no path)
- States 1 and 2 <u>communicate</u> (one can go from 1 to 2 and from 2 to 1).

Closed set

- A set of states *S* in a Markov chain is a **closed set** if no state outside of *S* is reachable from any state in *S*.

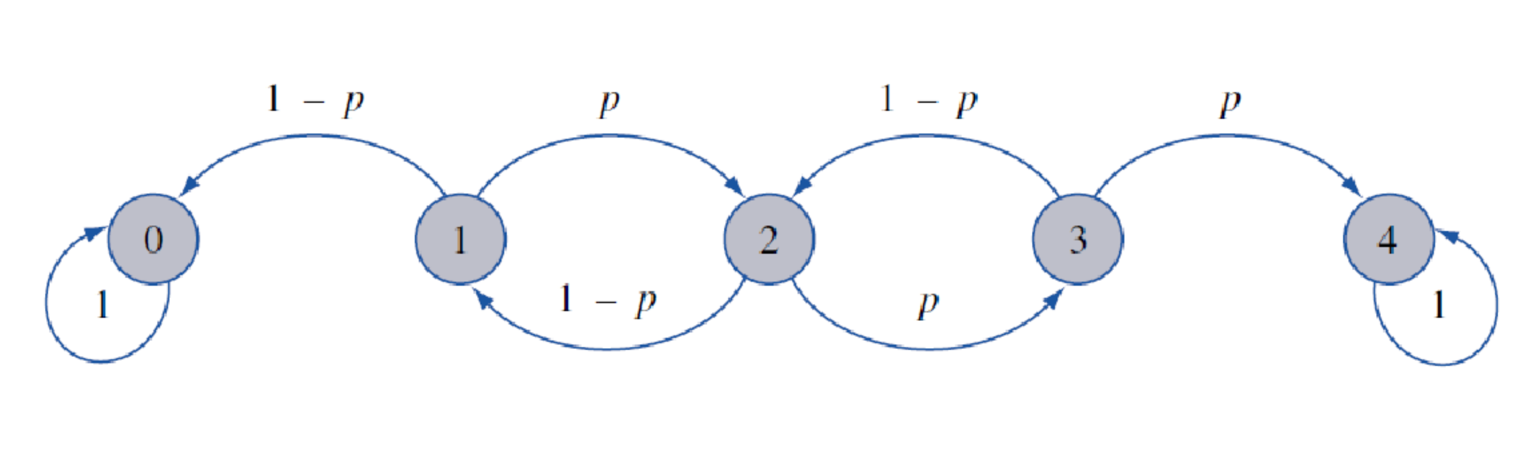


*S*1={1, 2} and *S*2={3,4, 5} are both <u>closed sets</u>

Absorption

- A state i is an **absorbing state** if p_{ii}=1.

Whenever we enter an absorbing state, we never leave the state. Of course, an absorbing state is a closed set containing only one state.

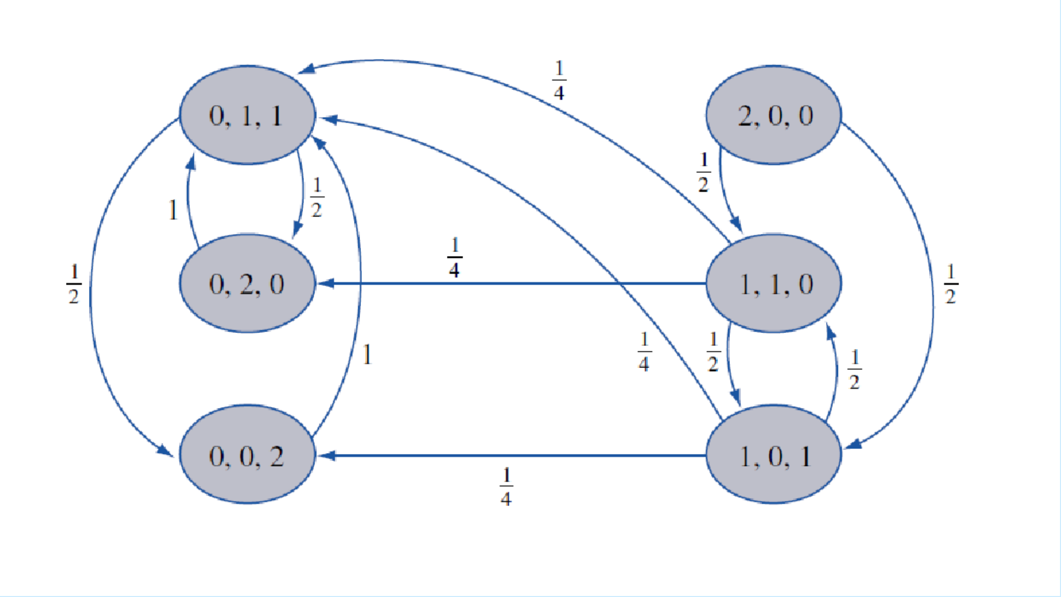


States 0 and 4 are <u>absorbing states</u>

Transition

- A state *i* is a **transient state** if there exists a state *j* that is reachable from *i*, but the state *i* is not reachable from state *j*.

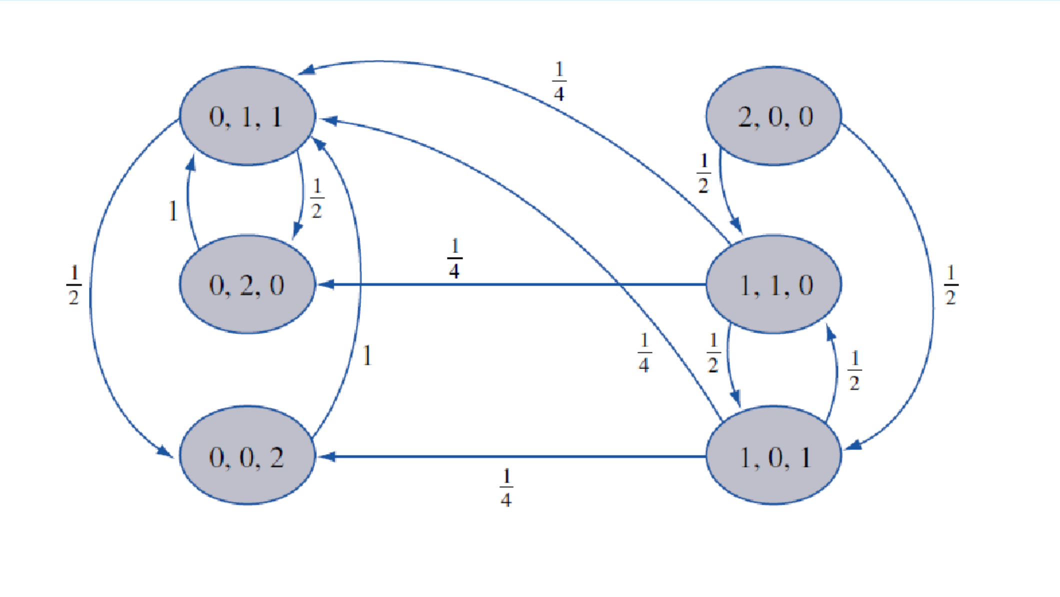
i.e. a state *i* is transient if there is a way to leave state *i* that never returns to state *i*



[2 0 0], [1 1 0], and [1 0 1] are all <u>transient states</u>

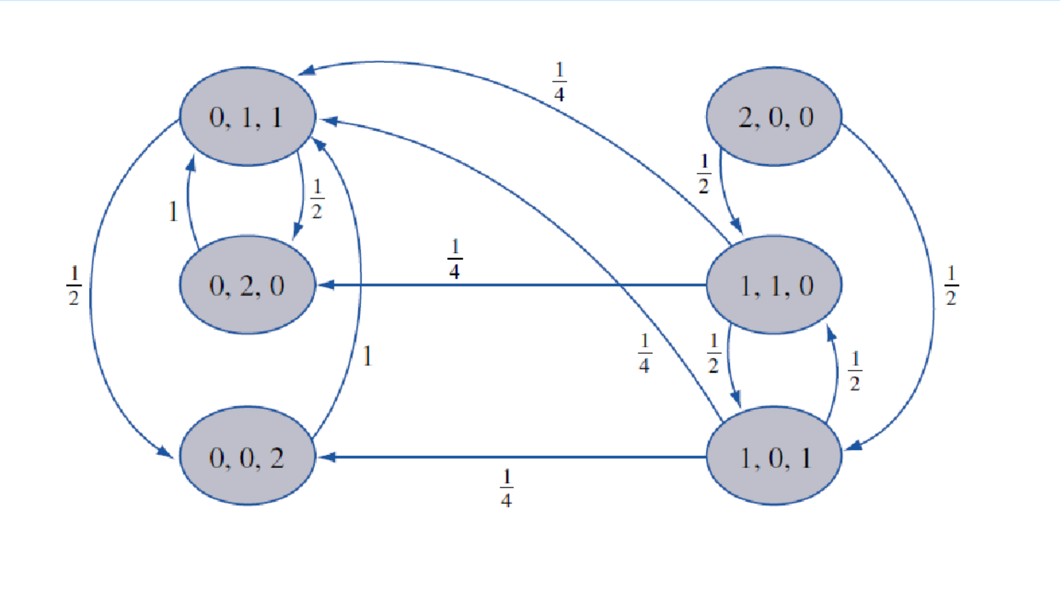
- After a large number of periods, the probability of being in any transient state *i* is zero.
- Each time we enter a transient state *i*, there is a positive probability that we will leave *i* forever and end up in the state *j* described in the definition of a transient state.
- Thus, eventually we are sure to enter state *j* (and then we will never return to state *i*).

- Suppose we are in the transient state [1 0 1].
- With probability 1, the unpainted ball will eventually be painted, and we will never reenter state [1 0 1]

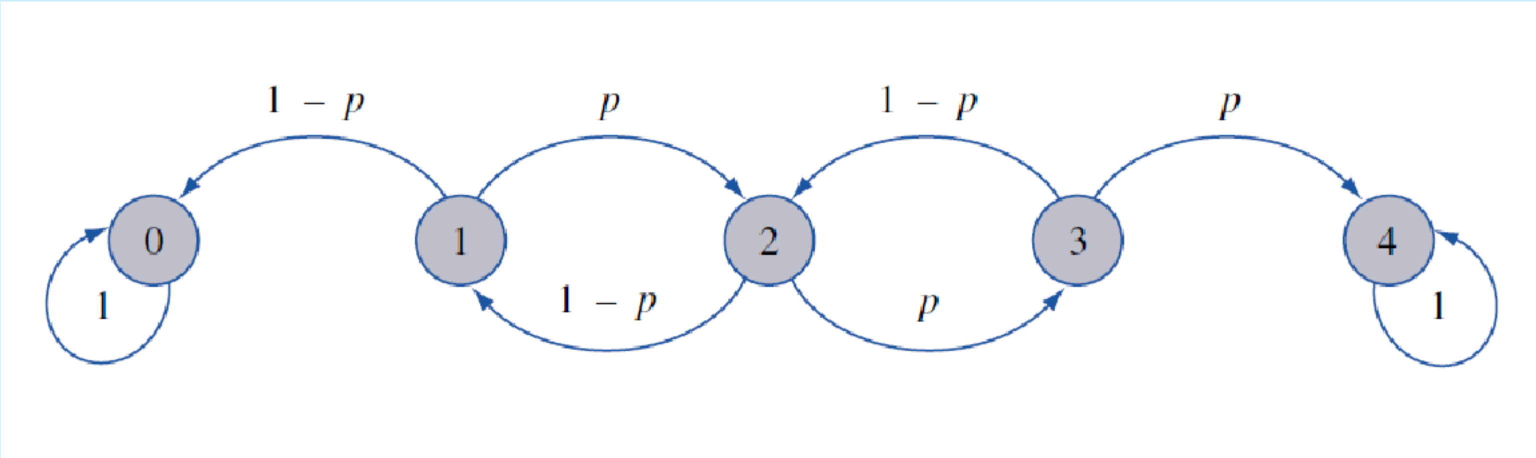


Recurrence

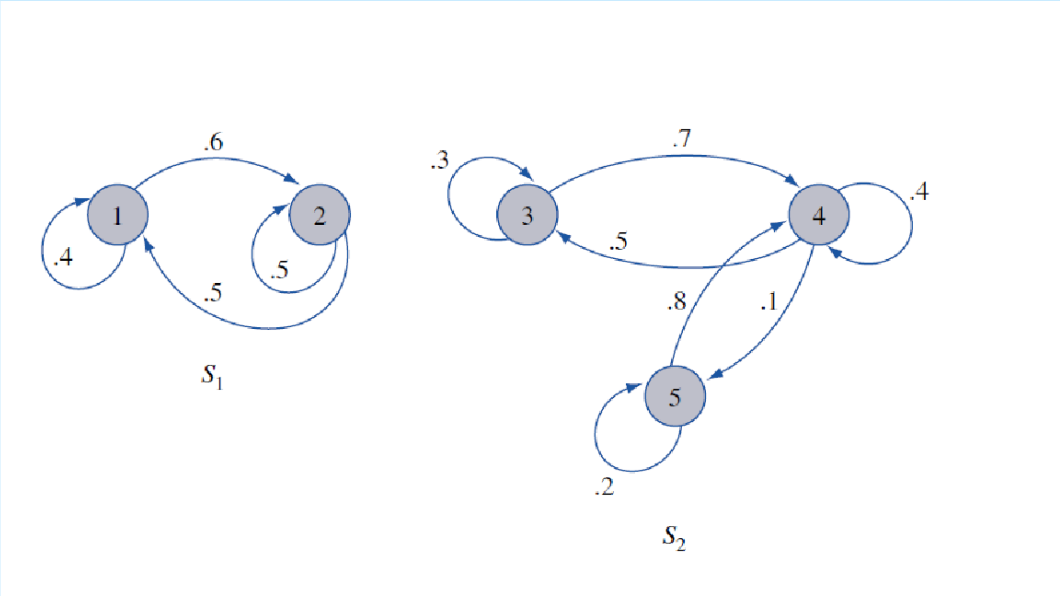
- A state *i* is a **recurrent state** if it is not transient



[0 2 0], [0 0 2], and [0 1 1] are <u>recurrent states</u>



State 0 & State 4 are recurrent

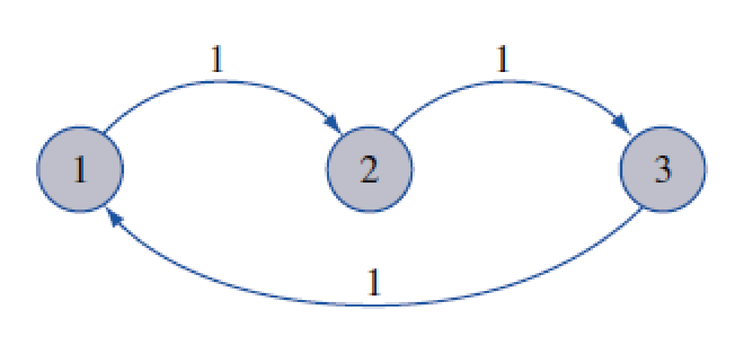


All states are recurrent

Periodic

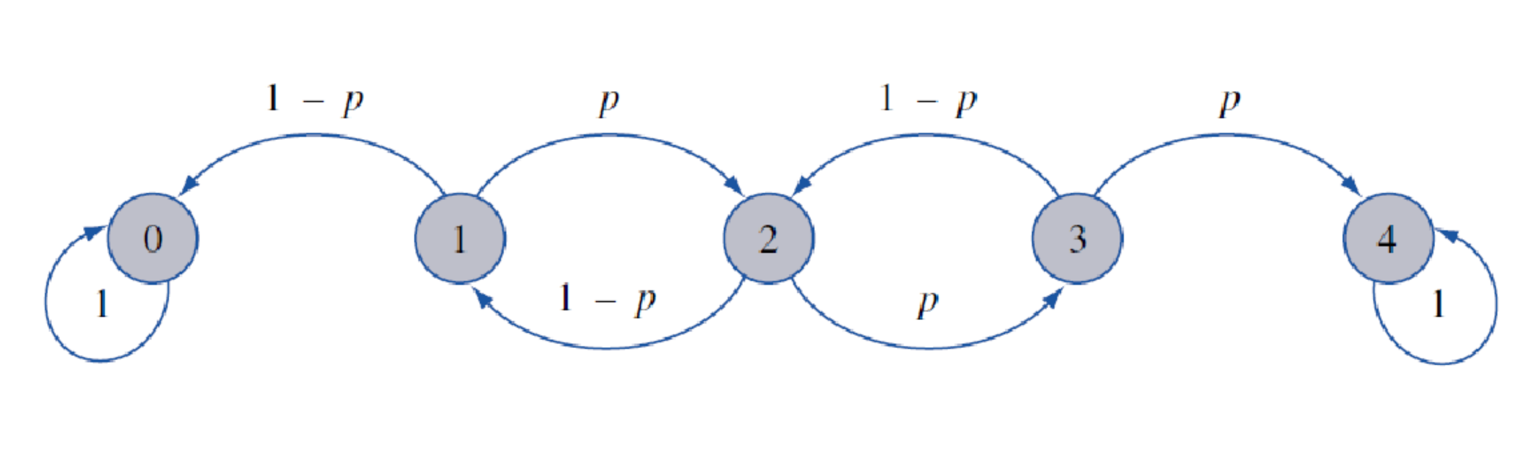
- A state *i* is **periodic** with period *k* > 1 if *k* is the smallest number such that all paths leading from state *i* back to state *i* have a length that is a multiple of *k*.

- For the Markov chain with transition matrix Q = $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ each state has period 3 (e.g. 1-2-3-1; hence if we begin in state 1, any return to state 1 will take 3*m* transitions)

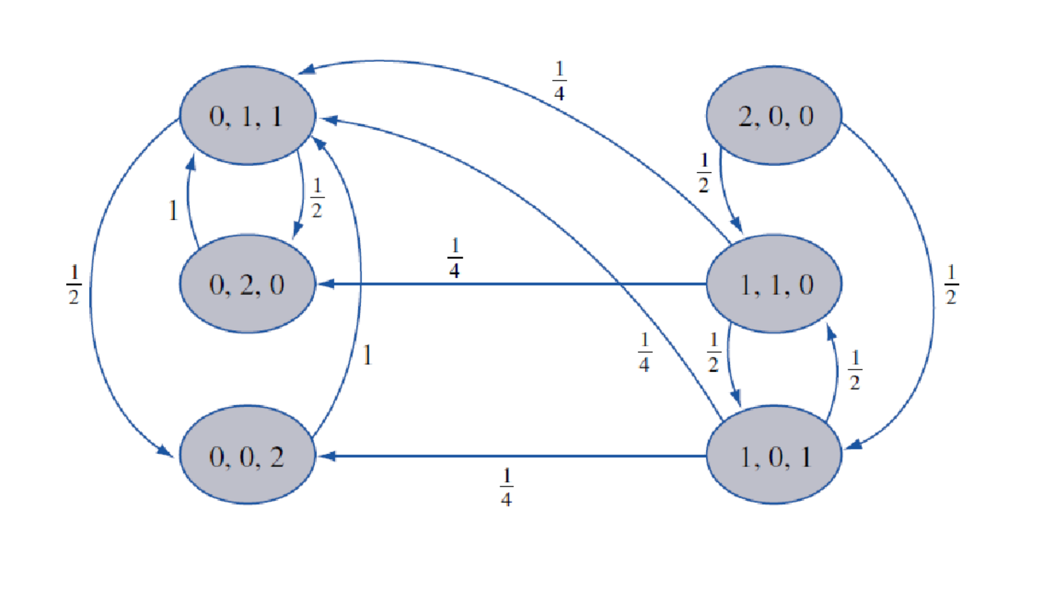


Aperiodic, Ergodic

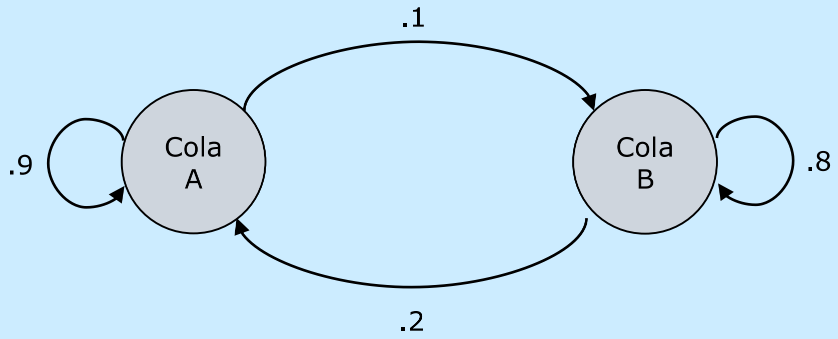
- If a recurrent state is not periodic, it is referred to as **aperiodic.**
- If all states in a chain are recurrent, aperiodic, and communicate with each other, the chain is said to be **ergodic.**



Nonergodic chain (e.g. States 3 and 4 do not communicate)



Nonergodic Chain
(e.g. [2 0 0] and [0 1 1] do not communicate)



Ergodic Chain
(all states are recurrent, aperiodic, and communicate with each other

Steady-State Probabilities

- Steady-state probabilities are used to describe the long-run behavior of a Markov chain.
- **Theorem 1:** Let P be the transition matrix for an s-state ergodic chain. Then there exists a vector $\pi = [\pi_1 \ \pi_2 \ \ldots \ \pi_s]$ such that

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_s \\ \pi_1 & \pi_2 & \cdots & \pi_s \\ \vdots & \vdots & & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_s \end{bmatrix}$$

- For large n, P^n approaches a matrix with identical rows

Steady-State Distribution

- Theorem 1 tells us that for any initial state *i*,

$$\lim_{n\to\infty} P_{ij}(n) = \pi_j$$

- This means that after a long time, the Markov Chain settles down, and independent of the initial state *i*, there is a probability π_j that we are in state j.

- The vector $\pi = [\pi_1 \ \pi_2 \ \ldots \ \pi_s]$ is often called **steady-state distribution** or **equilibrium distribution** for the Markov chain.

Finding steady-state probabilities

- From Theorem 1, for large n and all i

  $$P_{ij}(n+1) \cong P_{ij}(n) \cong \pi_j$$

- Since $P_{ij}(n+1)$ =(row i of P^n) (column j of P)

  $$P_{ij}(n+1) = \sum_{k=1}^{s} P_{ik}(n)p_{kj}$$

- If n is large

  $$\pi_j = \sum_{k=1}^{s} \pi_k p_{kj}$$

- In matrix form:

  $$\pi = \pi P$$

Finding steady-state probabilities

- Unfortunately, the system of equations in $\pi_j = \sum_{k=1}^{s} \pi_k p_{kj}$ has an infinite number of solutions because the rank of the P matrix always turns out to be ≤ s-1
- To obtain unique values, we may replace any of the equations in $\pi_j = \sum_{k=1}^{s} \pi_k p_{kj}$ with $\pi_1 + \pi_2 + \ldots + \pi_s = 1$

Example

- Find the steady-state probabilities for cola example

ANSWER

$$[\pi_1 \quad \pi_2] = [\pi_1 \quad \pi_2] \begin{bmatrix} .9 & .1 \\ .2 & .8 \end{bmatrix} \rightarrow$$

$\pi_1 = .9\ \pi_1 + .2\ \pi_2$ and $\pi_2 = .1\ \pi_1 + .8\ \pi_2$

Replacing any of the equations above with $\pi_1 + \pi_2 = 1 \rightarrow$

$\pi_1 = 2/3,\ \pi_2 = 1/3$

- Hence, after a long time, there is a 2/3 probability that a given person will purchase cola A and a 1/3 probability that a given person will purchase cola B.

Intuitive interpretation of steady-state probabilities

- An intuitive interpretation can be given to the steady-state probability equations
  $$\pi_j = \sum_{k=1}^{s} \pi_k p_{kj}$$
- By subtracting $\pi_j p_{ij}$ from both sides, we obtain:
  $$\pi_j (1 - p_{ij}) = \sum_{k \neq j} \pi_k p_{kj}$$
- This equation states that in the steady-state, the "flow" of probability into each state must equal the flow of probability out of each state.
- probality that a particular transition leaves state j = prob. that a particular transition enters state j

Example

- In the cola example suppose that each customer makes one purchase of cola during any week.
- Suppose there are 100 million cola customers.
- One selling unit of cola costs the company $1 to produce and is sold for $2.
- For $500 million per year (52 weeks), an advertising firm guarantees to decrease from 10% to 5% the fraction of cola A customers who switch to cola B after a purchase. Should the company that makes cola A hire the advertising firm?

Answer

- At present, a fraction π_1 = 2/3 of all purchases are cola A purchases.
- Each purchase of cola A earns the company a $1 profit.
- The annual profit would be 3,466.67 million dollars
- The advertising firm is offering to change the *P* matrix to

$$P_1 = \begin{bmatrix} .95 & .05 \\ .20 & .80 \end{bmatrix}$$

- For *P*1, the steady-state equations become

$$\pi_1 = .95\pi_1 + .20\pi_2$$
$$\pi_2 = .05\pi_1 + .80\pi_2$$

- As $\pi_1 + \pi_2 = 1$, we obtain

$$\pi_1 = .8 \text{ and } \pi_{12} = .2$$

- Now the cola A company's annual profit will be

  4,160 - 500 = 3,660 million dollars.
- Hence, the cola A company should hire the ad agency.

Mean First Passage Times

- For an <u>ergodic</u> chain,
  let m_{ij} = expected number of transitions
  before we first reach state j,
  given that we are currently in state i;
  m_{ij} is called the **mean first passage time**
  from state i to state j.

EXAMPLE

- In the cola example, m_{12} would be the expected number of bottles of cola purchased by a person who just bought cola A before first buying a bottle of cola B.

Mean First Passage Times

- Assume that we are currently in state i.
- Then with probability p_{ij}, it will take one transition to go from state i to state j.
- For $k \neq j$, we next go with probability p_{ik} to state k.
- In this case, it will take an average of $1 + m_{kj}$ transitions to go from i and j.
- This reasoning implies that

$$m_{ij} = p_{ij}(1) + \sum_{k \neq j} p_{ik}(1 + m_{kj})$$

$$m_{ij} = p_{ij}(1) + \sum_{k \neq j} p_{ik}(1 + m_{kj})$$

- Since $p_{ij} + \sum_{k \neq j} p_{ik} = 1$, we may rewrite m_{ij} as

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj}$$

- By solving these linear equations, we may find all the mean first passage times.
- It can be shown that

$$m_{ii} = \frac{1}{\pi_i}$$

Example

- What are the mean first passage times in the cola example?

ANSWERS

Recall that π_1 = 2/3 and π_2 = 1/3. Then

m_{11} = 3/2 and m_{22} = 3

As $m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj}$

$m_{12} = 1 + p_{11}\, m_{12} = 1 + 0.9\, m_{12}$

$m_{21} = 1 + p_{22}\, m_{21} = 1 + 0.8\, m_{21}$

$\rightarrow$ $m_{12} = 10$ and $m_{21} = 5$

- This means, i.e., that a person who last drank cola A will drink an average of 10 bottles of soda before switching to cola B.