Markov Chains

Kazi Saeed Alam
Assistant Professor, Dept. of CSE, KUET
Courtesy: Prof. Dr. Sk. Md. Masudul Ahsan

Stochastic Process

- Many real-world systems contain uncertainty and evolve over time.
  - Weather, stock market,
- A stochastic process is simply a random process through time.
  - A good way to think about it, it is the opposite of a deterministic process.
- In a deterministic process, given the initial conditions and the parameters of the system, we can define the exact "position" of the system at any time. In a stochastic process, we don't know where the process will be, even if we know the initial conditions and parameters.
- Stochastic processes have played a significant role in various engineering disciplines like power systems, robotics, automotive technology, signal processing, manufacturing systems, semiconductor manufacturing, communication networks, wireless networks etc

https://www.quora.com/What-is-a-stochastic-process-What-are-some-real-life-examples

Stochastic Process

- Stochastic processes (and Markov chains) are probability models for such systems.
- A ~ or Random process is a indexed collection of random variables $\{X_t\}$, where t runs through a given set T.
  - X_t = state of system (some measurable characteristic) at t
- ~ is a family of random variable $\{ X_t \mid t \in T \}$, defined on a given probability space, indexed by t that varies over index set T.
- A discrete-time stochastic process is a sequence of random variables

 $X_0, X_1, X_2, \ldots$ typically denoted by $\{ X_t \}$.

Stochastic Process

- Values assumed by X_t are called **states**, set of all possible values of states constitute **state space (*S*)**
- If state space is
  - Discrete – called discrete-state process or **chain**
  - Continuous - called continuous-state process
- If index set T is
  - Discrete – called discrete-time process or **sequence**
  - Continuous - called continuous-time process

Check these Book

- Intro to Probability_2nd ed_ DP Bertsekas
- Prob. Stat. and Stoch. Process - P Olofsson

Markov Chain

- Markov Property: a stochastic process is said to have ~ if probability distribution of future state depends only on present state and not on how the process arrived in that state.
- Formally-The state of the system at time $t+1$ depends only on the state of the system at time t

$$\boxed{P\left[X_{t+1} = x_{t+1} \mid X_t = x_t, X_{t-1} = x_{t-1}, \ldots X_0 = x_t\right] = P\left[X_{t+1} = x_{t+1} \mid X_t = x_t\right]}$$

- A stochastic process $\{X_t\}$ having Markov property is called **Markov Process**
  - **Markov chain** -If the state space of a Markov process is discrete
  - If index set is also discrete – **Discrete-time Markov chain**
    - What we are interested in

Markov Chain

		State Space	
		Discrete	Continuous
Time	Discrete	Discrete-Time Markov Chain	Discrete-Time Markov Process
	Continuous	Continuous-Time Markov Chain	Continuous-Time Markov Process

Stationary Markov Chain

- **Stationary Assumption:** For all states *i, j* and for all *t*, $P(X_{t+1} = j \mid X_t = i)$ is independent of time (t)

  $$P(X_{t+1} = j \mid X_t = i) = P(X_1 = j \mid X_0 = i) = p_{ij}$$

  $$;\quad i,j = 0,1, \dots, s; \quad t = 0,1, \dots,T$$

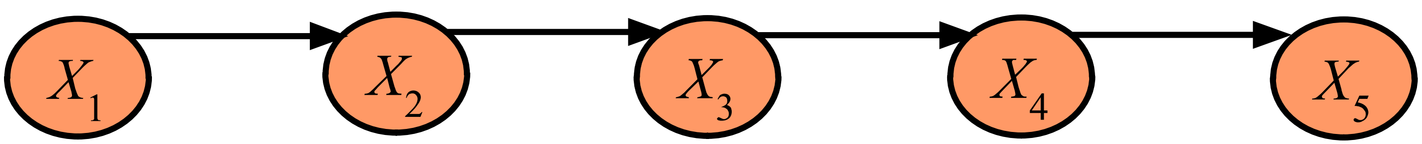
- This means that if system is in state i, the probability that the system will transition to state j is p_{ij} no matter what the value of t is
- p_{ij} = probability that the system will be in state j at time t+1 given that it is in state *i* at time t
  - Called one step **transition probability**
- Stationary Markov Chain – having stationary transition probabilities

Markov Chain

- We generally represent transition probabilities of Markov Chain by a $s \times s$ **Transition Probability Matrix P**

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{22} & \cdots & p_{2s} \\ \vdots & \vdots & & \vdots \\ p_{s1} & p_{s2} & \cdots & p_{ss} \end{bmatrix} \qquad p_{ij} \geq 0 \quad \text{and} \quad \sum_{j=1}^{s} p_{ij} = 1$$

- It can also be represented by stochastic Finite state Machine

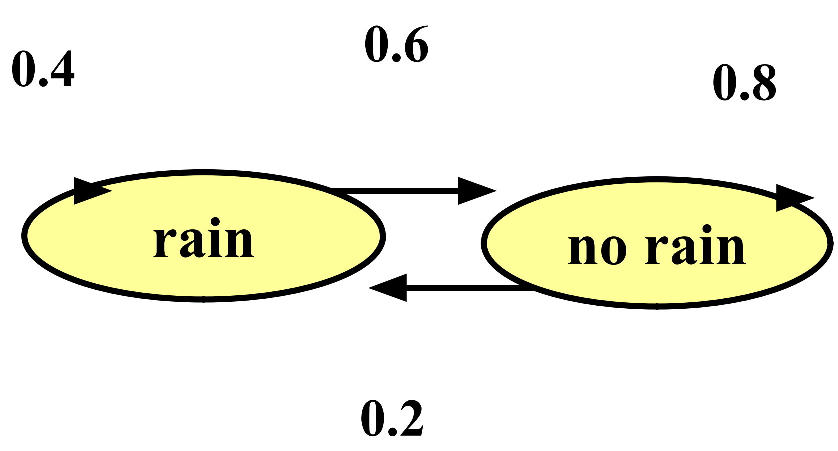


Markov Chain
Simple Example

Weather:

- raining today → 40% rain tomorrow
  → 60% no rain tomorrow
- not raining today → 20% rain tomorrow
  → 80% no rain tomorrow

Stochastic FSM:



Transition Prob. Matrix:

State 1: Rain
State 2: No Rain

$$P = \begin{pmatrix} 0.4 & 0.6 \\ 0.2 & 0.8 \end{pmatrix}$$

Transforming a process to a Markov chain

- Whether or not it rains today depends on previous weather conditions through last two days
  - If it rained for past two days, it will rain tomorrow with prob. 0.7
  - If it rained today but not yesterday, it will rain tomorrow with prob. 0.5
  - If it rained yesterday but not today, it will rain tomorrow with prob. 0.4
  - If it has not rained for past two days, it will rain tomorrow with prob. 0.2
- Let the state at time n – depend only on a single day
  - Not Markov chain
- Convert – n saying that it depend on both day

Transforming a process to a Markov chain

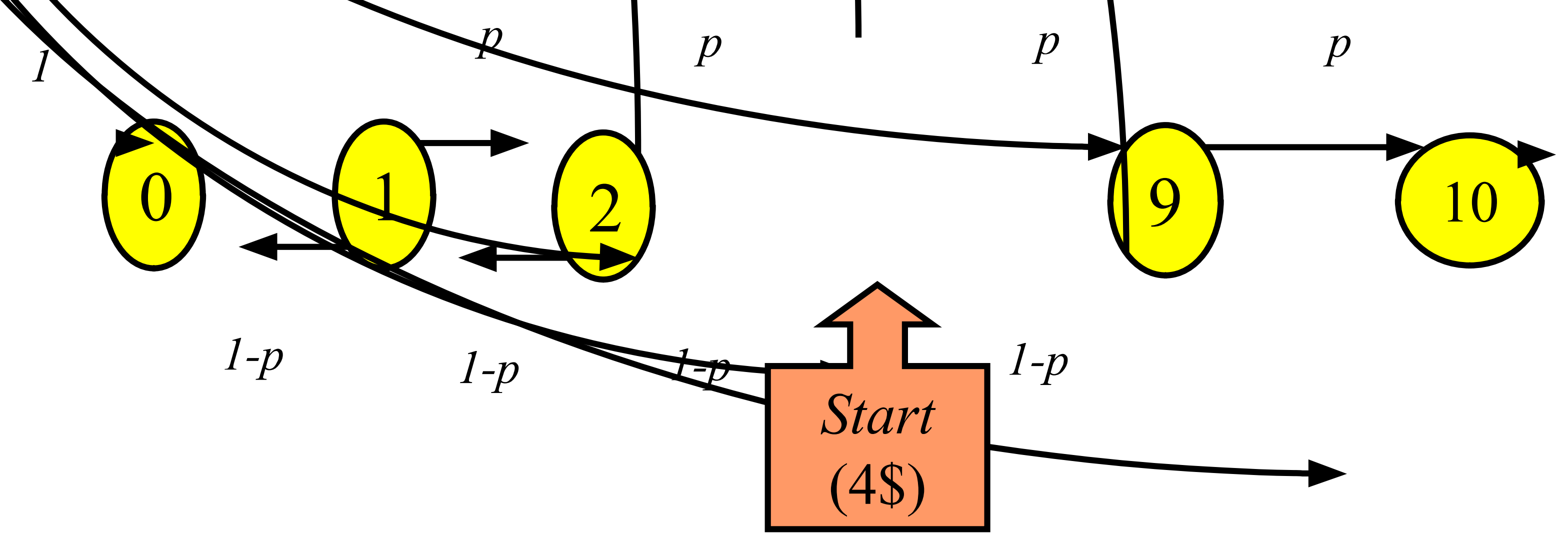
- State 0 - If it rained today and yesterday (RR)
- State 1 - If it rained today but not yesterday (NR)
- State 2 - If it rained yesterday but not today (RN)
- State 3 - If it did not rained either today or yesterday (NN)

$$
\begin{array}{r c}
 & \text{Today, Tomorrow} \\
\underline{\text{Yesterday, Today}} & \begin{array}{cccc} \underline{\text{RR}} & \underline{\text{NR}} & \underline{\text{RN}} & \underline{\text{NN}} \end{array} \\
\begin{array}{r} \text{RR} \\ \text{NR} \\ \text{RN} \\ \text{NN} \end{array} & P = \begin{bmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{bmatrix}
\end{array}
$$

Markov Chain
Gambler's Example

- Gambler starts with \$4
- At each play we have one of the following:
  - Gambler wins \$1 with probability p
  - Gambler looses \$1 with probability $1-p$
- Game ends when gambler goes broke, or gains a fortune of \$10

(Both 0 and 10 are absorbing states)



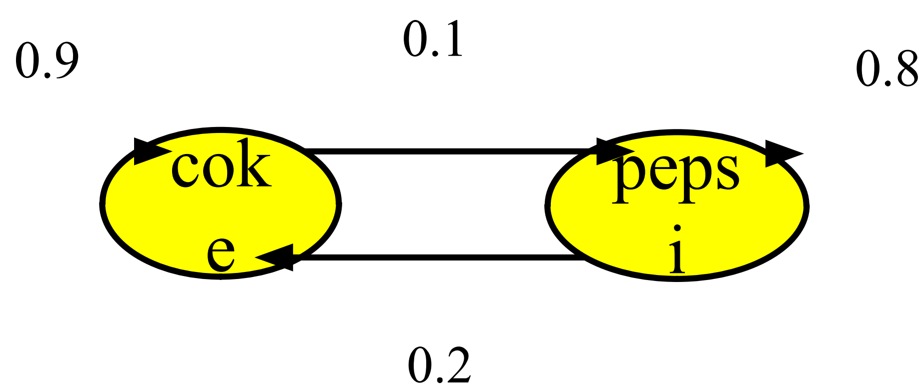
Markov Chain
Coke vs. Pepsi Example

- Given that a person's last cola purchase was Coke, there is a 90% chance that his next cola purchase will also be Coke.
- If a person's last cola purchase was Pepsi, there is an 80% chance that his next cola purchase will also be Pepsi.

transition matrix:

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$$

State 1: Coke
State 2: Pepsi



Characteristics of a Markov chain

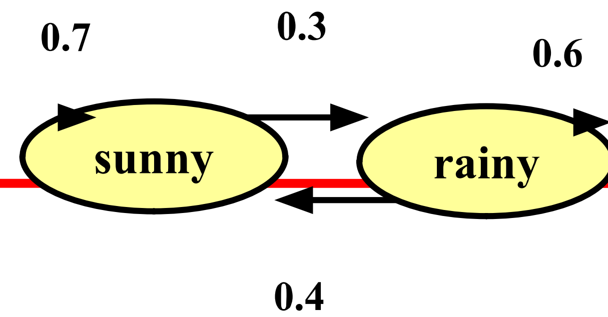
- What do we need to know to describe a Markov chain?
- Next state depends on previous state only, therefore, it is sufficient to know
  - the distribution of its initial state X_0
    - initial distribution P_0 – pmf of X_0
  - the mechanism of transitions from one state to another.
    - one-step transition probabilities p_{ij}.

Characteristics of a Markov chain

- Based on this data, we would like to compute:
  - n-step transition probabilities $\boldsymbol{p}_{ij}^{(n)}$;
  - Q_n the distribution of states at time n, which is our forecast for X_n;
  - The limit of $\boldsymbol{p}_{ij}^{(n)}$; and Q_n as $n \to \infty$, which is our long-term forecast.

(nearly constant)

n-step Transition Probabilities



- We may be interested in
  - It rains on Monday. Make forecasts for Wednesday, and Thursday. (For weather forecast example with above FSM)
- Mathematically
  - $p_{21}^{(2)}$ = P { Wednesday is sunny | Monday is rainy }
- More generally
  - If a Markov chain in state *i* at time m, what is the probability that n periods later the Markov chain will be in state *j*
  - ie. $P(X_{m+n} = j \mid X_m = i) = ?$

n-step Transition Probabilities

- Since we are dealing with stationary Markov chain, we can write
  - $P(X_{m+n} = j \mid X_m = i) = P(X_n = j \mid X_0 = i) = p_{ij}^{(n)}$
    - $p_{ij}^{(n)}$ = n-step probability of transition from state *i* to state *j*
- *If* $P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{22} & \cdots & p_{2s} \\ \vdots & \vdots & & \vdots \\ p_{s1} & p_{s2} & \cdots & p_{ss} \end{bmatrix}$ $\quad p_{ij} \geq 0$ and $\sum_{j=1}^{s} p_{ij} = 1$
- Cleary $p_{ij}^{(1)} = p_{ij}$
- Matrix $\mathbf{P^n}$ represents n-step transition probabilities from any state *i* to state *j*
  - How will you find P^n ?

n-step Transition Probabilities

- $p_{ij}^{(2)}$= probability that the system will be in state *j* two periods from now, considering it is now in state *i*

$$
\begin{aligned}
p_{ij}^{(2)} &= P(X_2 = j | X_0 = i) \\
&= \sum_{k \in S} P(X_2 = j | X_1 = k, X_0 = i) P(X_1 = k | X_0 = i) \quad \text{law of total probability} \\
&= \sum_{k \in S} P(X_2 = j | X_1 = k) P(X_1 = k | X_0 = i) \quad \text{(by Markov property)} \\
&= \sum_{k \in S} p_{kj} p_{ik}.
\end{aligned}
$$

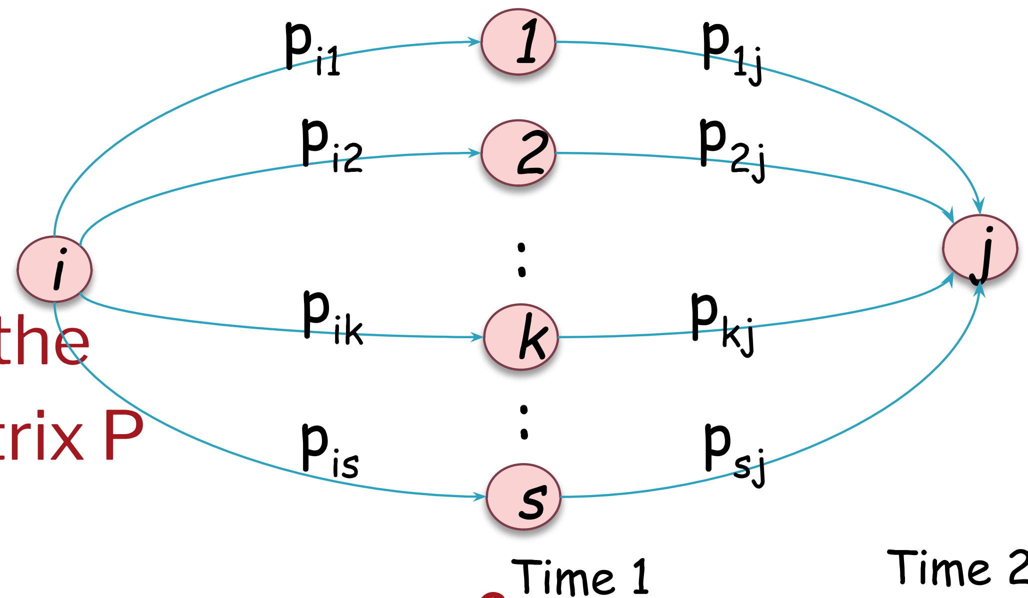
two-step transition matrix as follows:

$$
P^{(2)} = \begin{bmatrix} p_{11}^{(2)} & p_{12}^{(2)} & \dots & p_{1r}^{(2)} \\ p_{21}^{(2)} & p_{22}^{(2)} & \dots & p_{2r}^{(2)} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p_{r1}^{(2)} & p_{r2}^{(2)} & \dots & p_{rr}^{(2)} \end{bmatrix}
$$

n-step Transition Probabilities

- To compute $p_{ij}^{(2)}$, we must go from state *i* to some state *k*, then form state *k* to state *j*

$$p_{ij}^{(2)} = \sum_{k=1}^{s} p_{ik}\, p_{kj} \quad \text{for all states } i,\, j$$



  - Clearly right side of the eq^n is the scalar product of i^{th} row of matrix P with j^{th} column of matrix P
  - Hence, $p_{ij}^{(2)}$ is the ij^{th} element of matrix P.P = $\mathbf{P^2}$
  - i.e. Matrix $\mathbf{P^2}$ represents 2-step transition probabilities for all states *i, j*

$$P^{(2)} = P^2.$$

$$P^2 = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \cdot \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}$$

n-step Transition Probabilities

3-step transitions: We can find: $P(X_3 = j \mid X_0 = i)$ similarly, but conditioning on the state at time 2:

$$p_{ij}^{(3)} = P(X_3 = j \mid X_0 = i)$$

$$= \sum_{k=1}^{s} P(X_3 = j \mid X_2 = k) P(X_2 = k \mid X_0 = i)$$

$$= \sum_{k=1}^{s} p_{kj} p_{ik}^{(2)} = \sum_{k=1}^{s} p_{ik}^{(2)} p_{kj} = i^{th} \text{row of } P^2 . j^{th} \text{col of } P$$

$$= (P^3)_{ij}$$

n-step Transition Probabilities

- Extending the previous reasoning we can find that
- $\mathbf{P}^{(n)}$ = n-step transition probabilities for all states *i, j*

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

or

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

$$= \mathbf{P}^{n-m}.\mathbf{P}^{m} = \mathbf{P}^{m}.\mathbf{P}^{n-m} = \mathbf{P}^{n}$$

The Chapman-Kolmogorov equation

$$p_{ij}^{(m+n)} = P(X_{m+n} = j | X_0 = i) = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}.$$

- Of course for n = 0, $p_{ij}^{(0)} = P(X_0 = j | X_0 = i)$, so we write

$$p_{ij}^{(0)} = \begin{cases} 1 & if \quad j = i \\ 0 & if \quad j \neq i \end{cases}$$

Markov chain

Coke vs. Pepsi Example (cont)

Given that a person is currently a Pepsi purchaser, what is the probability that he will purchase Coke two purchases from now?

Pr[Pepsi→?→Coke] =

Pr[Pepsi→Coke→Coke] + Pr[Pepsi→ Pepsi →Coke] =

$$0.2 * 0.9 + 0.8 * 0.2 = 0.34$$

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

Pepsi → ?

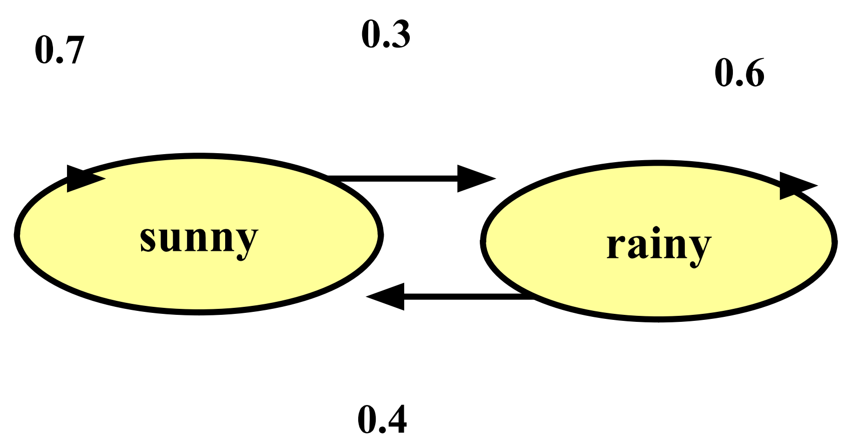
? → Coke

Markov chain

Weather example (cont)

$p_{21}^{(2)}$ = P { Wednesday is sunny | Monday is rainy }

$$= p_{21}p_{11} + p_{22}p_{21}$$
$$= (0.4)(0.7) + (0.6)(0.4)$$
$$= 0.52.$$



$$\text{state } 1 = \text{“sunny”}$$
$$\text{state } 2 = \text{“rainy.}$$

Markov chain

Coke vs. Pepsi Example (cont)

- Given that a person is currently a Coke purchaser, what is the probability that he will purchase Pepsi **three** purchases from now?

$$P^3 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

Unconditional State Probabilities

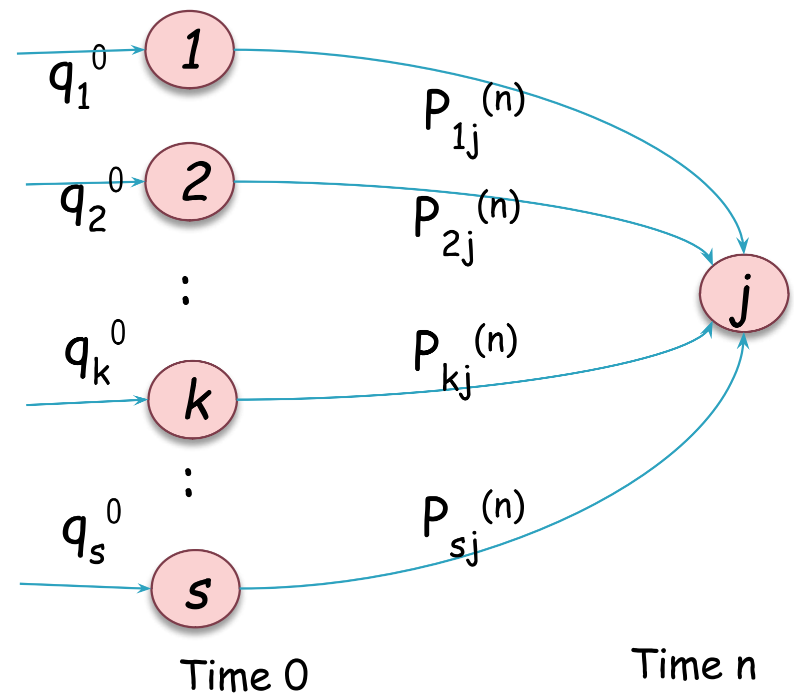
- One or n-step transition probabilities are *conditional* prob.
  - For ex. $P(X_n = j \mid X_0 = i) = p_{ij}^{(n)}$
- Sometimes, we may not know the state of the Markov chain at time 0, but we are interested to determine the prob. that the system is in state j at time n
  - That is $P(X_n = j) = ?$
- Weather example
  - Suppose now that it does not rain yet, but meteorologists predict an 80% chance of rain on Monday. How does this affect our forecasts?

Unconditional State Probabilities

- For, $P(X_n = j) = ?$
  - it is necessary to specify prob. distribution of initial state
    - i.e. $P(X_0 = i)$ for all states *i*,
    - let it be vector $\mathbf{Q_0}$, where $q_i^0 = P(X_0 = i)$ for all states ***i***
- *Then,*

$$P(X_n = j) = \sum_{i=1}^{s} q_i^{\,0} p_{ij}(n)$$

$$= \mathbf{Q_0}(j^{th} \text{ column of } \mathbf{P^n})$$



Markov chain

Coke vs. Pepsi Example (cont)

- Assume each person makes one cola purchase per week
- Suppose 60% of all people now drink Coke, and 40% drink Pepsi
- What fraction of people will be drinking Coke three weeks from now?

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \qquad P^3 = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix}$$

Q_i - the distribution in week i

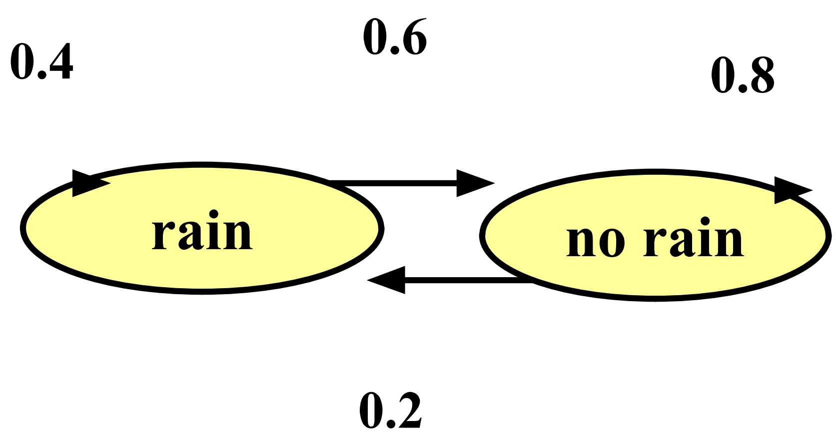
$Q_0=(0.6,0.4)$ - initial distribution

$$\mathrm{P}[X_3=\text{Coke}] = Q_0 * (1^{\text{st}} \text{ column of } P^3) = [0.6 \quad 0.4] * [0.781 \quad 0.438]^T$$
$$= 0.6 * 0.781 + 0.4 * 0.438 = 0.6438$$

$$Q_3 = Q_0 * P^3 = (0.6438, 0.3562)$$

Classification of States in Markov Chain

- We now know probabilities associated with states
- We can classify the states of the system
  - Whether you can get from one state to another
  - Whether you can return to a state
- To help in classifying states, we use a state diagram from the weather example:



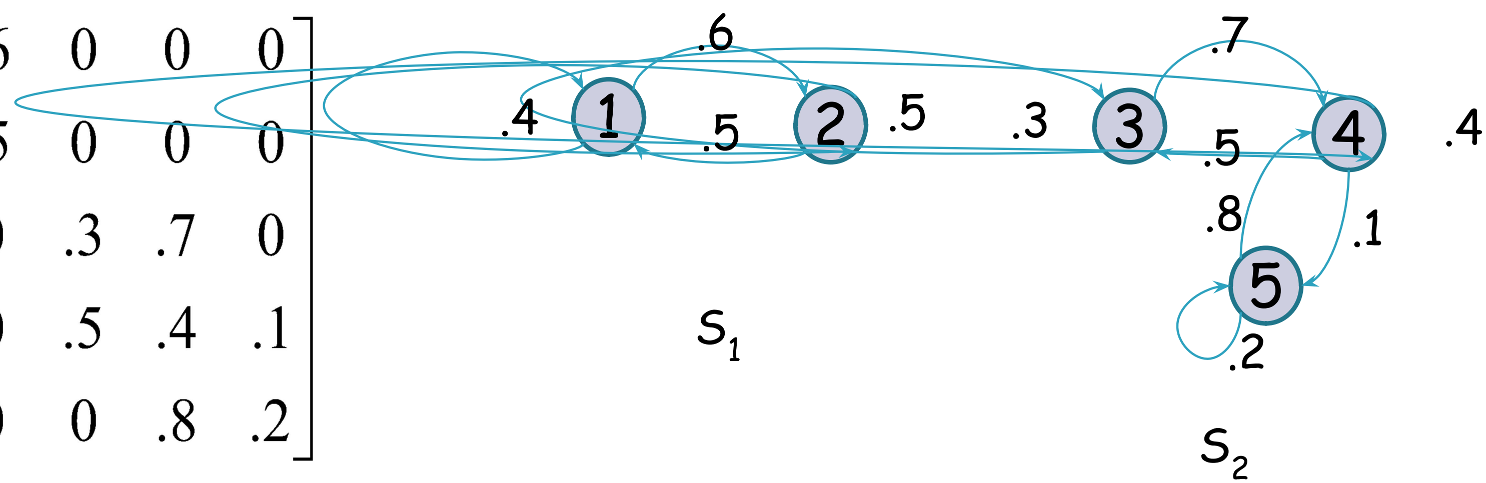
Classification of States - Definitions

- **Path** - a sequence of transitions from state i to state j exists and has positive probability, i.e., $p_{ij}^{(n)}>0$ for some n.
- State j is **Reachable** from state i if there is a path from i to j
- Two states, i and j, **Communicate (**$i \leftrightarrow j$**)** if j is reachable from i, and i is reachable from j.
  - It is easy to check that this is an equivalence relation:
    1. $i \leftrightarrow i$; since $p_{ii}^{(0)}=1$
    2. $i \leftrightarrow j$ implies $j \leftrightarrow i$; and
    3. $i \leftrightarrow j$ and $j \leftrightarrow k$ together imply $i \leftrightarrow k$.

Classification of States – Definitions (cont)

- A set of states S in a Markov Chain is a **closed set** if
  - All the states of S communicate with each other, and
  - No state outside of S is reachable.

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

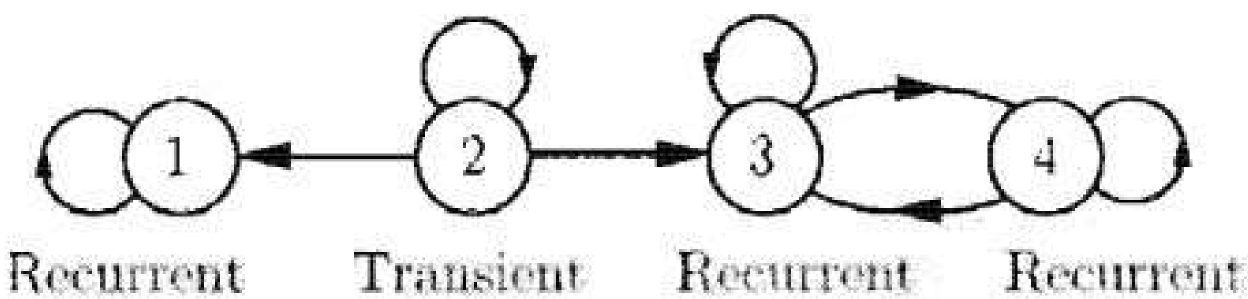


Classification of States – Definitions (cont)

- **Irreducible Markov Chain** - if there is only one Closed set
  - Eg. Weather, Coke vs Pepsi
- A state i is an **Absorbing state** if the process never will leave the state
  - i.e. the state returns to itself with certainty in one transition
    - $p_{ii} = 1$ (closed set with 1 member)
  - Example of Absorbing State - The Gambler's Ruin
    - At each play we have the following:
      - Gambler wins $1 with probability p, or loses $1 with probability 1-p
    - Game ends when gambler goes broke, or gains a fortune of $N
      - Then both $0 and $N are absorbing states

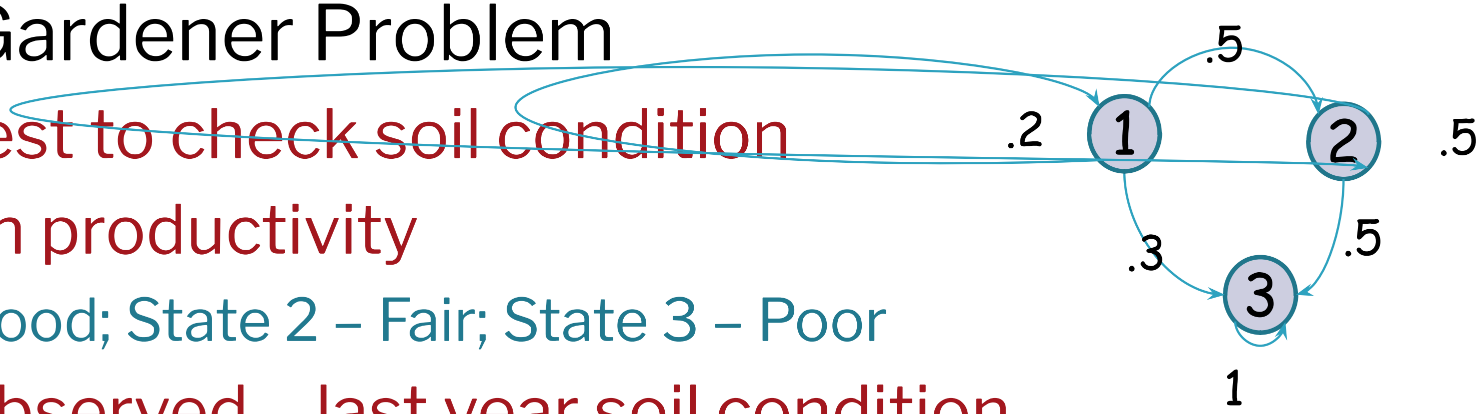
Classification of States – Definitions (cont)

- A state i is a **Transient state** if the process *may never return* the state again.
  - i.e. there exists a state j that is reachable from i, but i is not reachable from j.
  - Mathematically, $\lim_{n\to\infty} p_{ji}^{(n)} = 0, \quad \text{for all j}$
- A state is **Recurrent** if– upon entering the state, the process *definitely will return* the state again.
  - if and only if it is not transient.



Classification of States – Definitions (cont)

- Example – Gardener Problem
  - Chemical test to check soil condition
  - New season productivity
    - State 1 – Good; State 2 – Fair; State 3 – Poor
  - Gardener observed – last year soil condition impacts currents year productivity



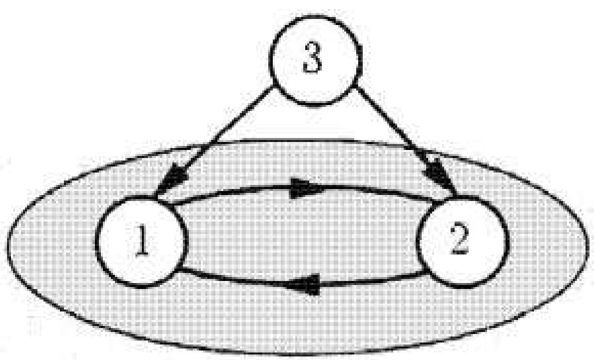
$$P = \begin{bmatrix} .2 & .5 & .3 \\ 0 & .5 & .5 \\ 0 & 0 & 1 \end{bmatrix}$$

- Ex- Gardener Problem
  - State 1, 2 transient
    - Can reach state 3 but never be reached back
  - State 3 absorbing - $p_{33}=1$

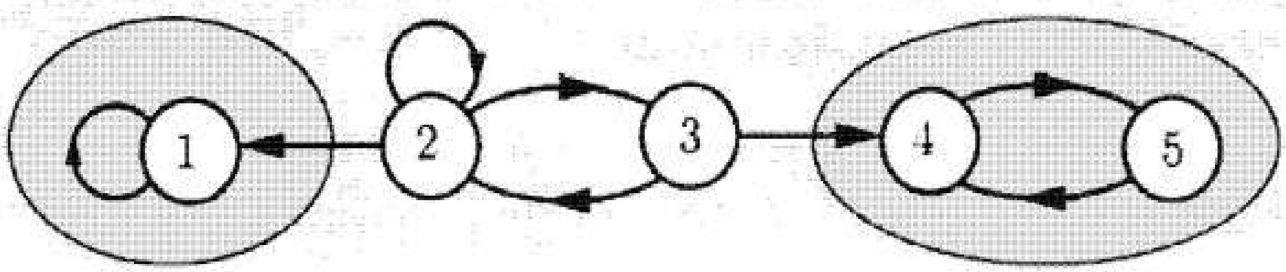
$$P^{100} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Classification of States

If i is a recurrent state, the set of states $A(i)$ that are accessible from i form a **recurrent class** (or simply **class**), meaning that states in $A(i)$ are all accessible from each other, and no state outside $A(i)$ is accessible from them.



Single class of recurrent states (1 and 2) and one transient state (3)



Two classes of recurrent states (class of state 1 and class of states 4 and 5) and two transient states (2 and 3)

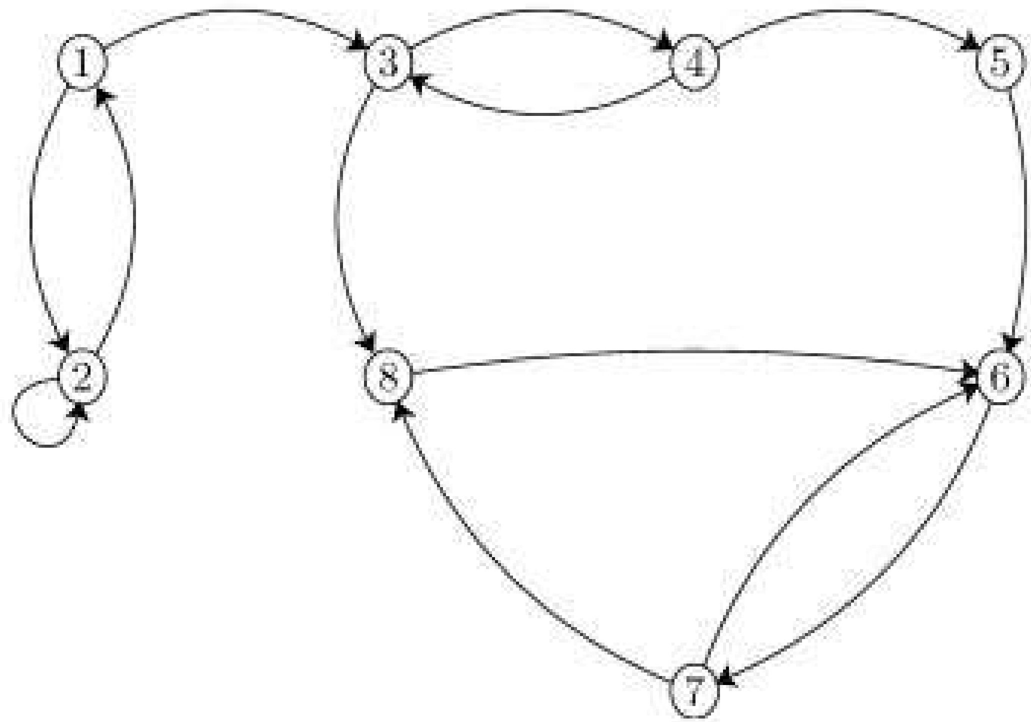
Classification of States

Markov Chain Decomposition

- A Markov chain can be decomposed into one or more recurrent classes, plus possibly some transient states.
- A recurrent state is accessible from all states in its class, but is not accessible from recurrent states in other classes.
- A transient state is not accessible from any recurrent state.
- At least one, possibly more, recurrent states are accessible from a given transient state.

Classification of States

- States i and j are in the same communicating class if i ↔ j: i.e. if each state is accessible from the other.
- Every state is a member of exactly one communicating class



four communicating classes

Class 1 = {state 1, state 2},
Class 2 = {state 3, state 4},
Class 3 = {state 5},
Class 4 = {state 6, state 7, state 8}.

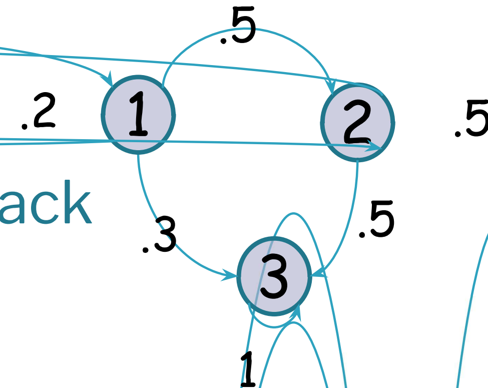
Don't confuse with recurrent class

transient class

recurrent class

Classification of States – Definitions (cont)

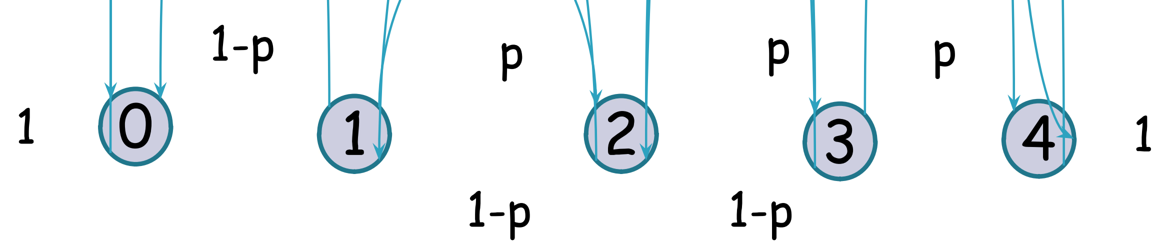
- Ex- Gardener Problem
  - State 1, 2 transient
    - Can reach state 3 but never be reached back
  - State 3 absorbing - $p_{33}=1$



$$P^{100} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

- Ex- Gambler Ruin (simple case – Quits when \$0 or \$4)
  - State 2 – Transient or Recurrent?
    - Ans. Transient

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



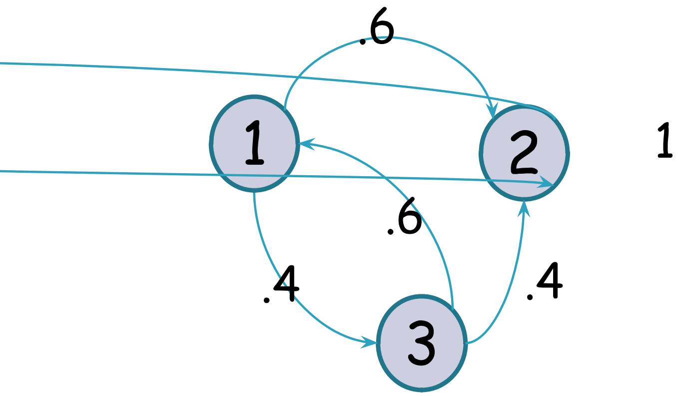
Classification of States – Definitions (cont)

- State *i* is **periodic** with period t > 1 if t is the smallest number such that all paths leading from state i back to state i have a length which is a multiple of **t**
  - i.e a return is possible only in t, 2t, 3t, … steps
  - Mathematically, $p_{ii}^{(n)} = 0$ whenever n is not divisible by t
- A recurrent state that is not periodic is called **aperiodic**

You can show that all states in the same communicating class have the same period. A class is said to be periodic if its states are periodic. Similarly, a class is said to be aperiodic if its states are aperiodic. Finally, a Markov chain is said to be aperiodic if all of its states are aperiodic.

$$\text{If } i \leftrightarrow j, \text{ then } d(i) = d(j).$$

Classification of States – Definitions (cont)



$$P=\begin{bmatrix}0 & .6 & .4\\ 0 & 1 & 0\\ .6 & .4 & 0\end{bmatrix}\qquad P^2=\begin{bmatrix}.24 & .76 & 0\\ 0 & 1 & 0\\ 0 & .76 & .24\end{bmatrix}$$

$$P^3=\begin{bmatrix}0 & .904 & .096\\ 0 & 1 & 0\\ .144 & .856 & 0\end{bmatrix}\quad P^4=\begin{bmatrix}.0576 & .9424 & 0\\ 0 & 1 & 0\\ 0 & .9424 & .0576\end{bmatrix}\quad P^5=\begin{bmatrix}0 & .97696 & .02304\\ 0 & 1 & 0\\ .03456 & .96544 & 0\end{bmatrix}$$

- Continuing with n = 6, 7, … P^n shows that p_{11} and p_{33} are (+)ve for even n and 0 otherwise

  i.e. states 1 and 3 have period 2

Classification of States – Definitions (cont)

- If all states in a Markov Chain are **recurrent, aperiodic**, and **communicate** with one another (a "nice" chain), then the Markov Chain is said to **Ergodic**
- Example –
  - Gambler Ruin
    - Not Ergodic
  - Weather
    - Ergodic
  - Coke vs Pepsi
    - Ergodic
  - Gardener
    - Not Ergodic

How do we check that a Markov chain is aperiodic?

- Remember that two numbers and are said to be co-prime if their greatest common divisor (gcd) is 1
- find two co-prime numbers l and m such that $p_{ii}^{(l)} > 0$ and $p_{ii}^{(\mathrm{m})} > 0$
  - that is, we can go from state to itself in l steps, and also in m steps.
- Then, we can conclude state is aperiodic.
- If we have an irreducible Markov chain, this means that the chain is aperiodic.
  - Since the number 1 is co-prime to every integer, any state with a self transition is aperiodic.

How do we check that a Markov chain is aperiodic?

Consider a finite irreducible Markov chain X_n:

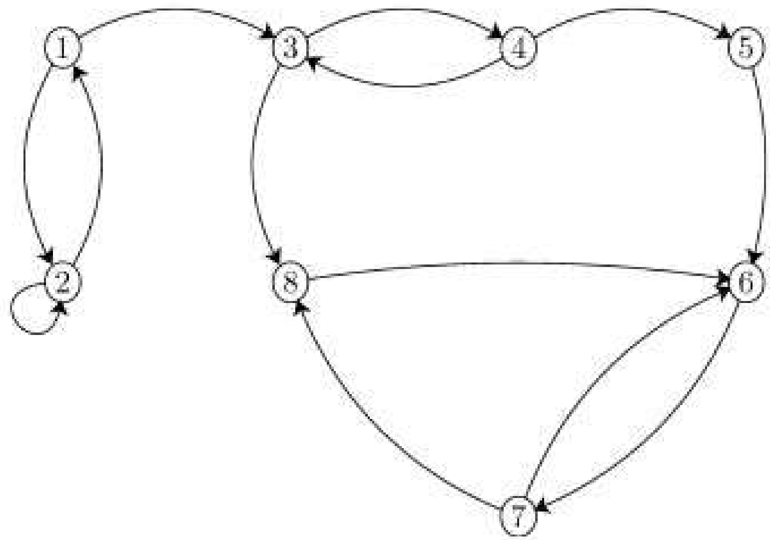
a. If there is a self-transition in the chain ($p_{ii} > 0$ for some i), then the chain is aperiodic.

b. Suppose that you can go from state i to state i in l steps, i.e., $p_{ii}^{(l)} > 0$. Also suppose that $p_{ii}^{(m)} > 0$. If $\gcd(l, m) = 1$, then state i is aperiodic.

c. The chain is aperiodic if and only if there exists a positive integer n such that all elements of the matrix P^n are strictly positive, i.e.,

$$p_{ij}^{(n)} > 0, \quad \text{for all } i, j \in S.$$

How do we check that a Markov chain is aperiodic?



a. Class $1 = \{\text{state } 1, \text{state } 2\}$ is aperiodic since it has a self-transition, $p_{22} > 0$.

b. Class $2 = \{\text{state } 3, \text{state } 4\}$ is periodic with period 2.

c. Class $4 = \{\text{state } 6, \text{state } 7, \text{state } 8\}$ is aperiodic. For example, note that we can go from state 6 to state 6 in two steps $(6 - 7 - 6)$ and in three steps $(6 - 7 - 8 - 6)$. Since $\gcd(2, 3) = 1$, we conclude state 6 and its class are aperiodic.

How do we check that a Markov chain is aperiodic?

Classify the states of the following Markov chains. If a state is periodic, determine its period:

***(a)** $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$

***(b)** $\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & .5 & .5 & 0 & 0 & 0 \\ 0 & .7 & .3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & .4 & .6 \\ 0 & 0 & 0 & 0 & .2 & 8 \end{pmatrix}$

(d) $\begin{pmatrix} .1 & 0 & .9 \\ .7 & .3 & 0 \\ .2 & .7 & .1 \end{pmatrix}$

Long Run Property of Markov Chain

Steady State Probabilities

- n-step transition probabilities for Cola drinkers

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \quad P^2 = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} \quad P^3 = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix} \quad P^5 = \begin{bmatrix} 0.72 & 0.28 \\ 0.56 & 0.44 \end{bmatrix}$$

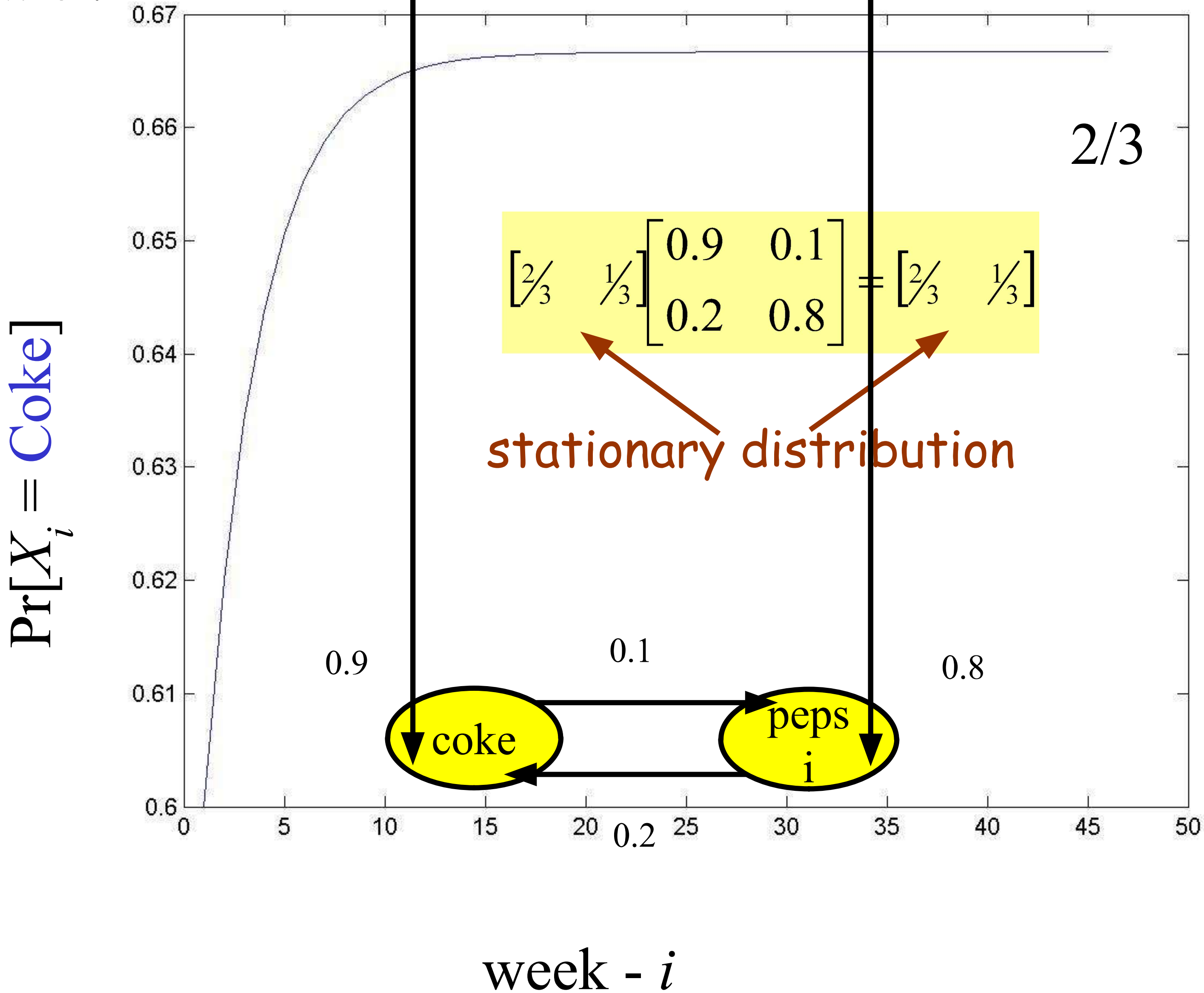
$$P^{10} = \begin{bmatrix} 0.68 & 0.32 \\ 0.65 & 0.35 \end{bmatrix} \quad P^{20} = \begin{bmatrix} 0.67 & 0.33 \\ 0.67 & 0.33 \end{bmatrix} \quad P^{30} = \begin{bmatrix} 0.67 & 0.33 \\ 0.67 & 0.33 \end{bmatrix} \quad P^{40} = \begin{bmatrix} 0.67 & 0.33 \\ 0.67 & 0.33 \end{bmatrix}$$

- After long time a person's next cola purchase probability doesn't depend on
  - Whether (s)he was initially Coke or Pepsi drinker

Markov Chain

Coke vs. Pepsi Example (cont)

Simulation:



$$\begin{bmatrix} 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 \end{bmatrix}$$

Long Run Property of Markov Chain

Steady State Probabilities

- **THEOREM:** Let P be the transition matrix for a s-state **ergodic chain**, then there exists a vector $\boldsymbol{\pi} = [\,\pi_1 \;\; \pi_2 \;\; \pi_3 \;\; .... \;\; \pi_s\,]$ such that

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

- Recall that ijth element of P^n is $p_{ij}^{(n)}$, so theorem tells that

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j, \quad \text{for any initial state } i$$

- π_j are called the steady state probabilities
- Vector $\boldsymbol{\pi}$ is called steady state/ equilibrium distribution
- Steady state
  - doesn't mean that the process settle down to **one state**
  - Means - Prob. of process in state j, after long time, tends to π_j , and independent of initial state distribution

Long Run Property of Markov Chain

Steady State Probabilities

- How can we find **π** ?
- From theorem, for large *n*, and all *i* $p_{ij}^{(n+1)} \cong p_{ij}^{(n)} = \pi_j$ ...... (1)
- Since $p_{ij}^{(n+1)}$ = (ith row of P^n) × (jth column of P)

$$\text{so,} \quad p_{ij}^{(n+1)} = \sum_{k=1}^{s} p_{ik}^{(n)} p_{kj} \quad \text{......... (2)}$$

- If *n* is large, substituting eq^n (1) into (2)

$$\textbf{so,} \quad \pi_j = \sum_{k=1}^{s} \pi_k p_{kj}$$

- In matrix form, **π = π P** .... (3)
- Unfortunately, eq^n (3) has infinite no. of solutions
- To have unique solution, along with eq^n (3), use $\sum_{j=1}^{s} \pi_j = 1$
- That is, solve the system of eq^n

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$$

$$\sum_{j=1}^{s} \pi_j = 1$$

Long Run Property of Markov Chain

Steady State Probabilities

- Coke vs Pepsi Example

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

$$\pi_1 + \pi_2 = 1$$

$$\pi_1 = .9\pi_1 + .2\pi_2$$

$$\pi_2 = .1\pi_1 + .8\pi_2$$

$$\pi_1 + \pi_2 = 1$$

Solving

$$\pi_1 = 2/3$$

$$\pi_2 = 1/3$$

- Suppose, 100 million cola customer, each person purchase 1 cola during any week
- Each selling profits $1
- For $500 million per year, an add firm guarantees to decrease 10% to 5% of Coke customers who switch to Pepsi after a purchase
- Should Coke company hire the add firm?

Long Run Property of Markov Chain

Steady State Probabilities

- Coke vs Pepsi Example
- Total cola purchase in a year (52 week)
  $= 100 * 10^6 * 52 = 5.2$ billion
- Each purchase earns $1 profit for a company
- Current profit in a year for Coke company
  = 2/3 (5.2 billion * $1) [since steady state prob. of buying coke $\pi_1 = 2/3$]
  =$3,466,666,667
- What the add firm offers to Coke compar
  for $500 million per year

$$P = \begin{bmatrix} 0.95 & 0.05 \\ 0.2 & 0.8 \end{bmatrix}$$

Long Run Property of Markov Chain
Steady State Probabilities

- Then what will be the long run/steady state probabilities?
  - Lets find.

$$P = \begin{bmatrix} 0.95 & 0.05 \\ 0.2 & 0.8 \end{bmatrix}$$

$$[\pi_1 \quad \pi_2] = [\pi_1 \quad \pi_2]\begin{bmatrix} .95 & .05 \\ .2 & .8 \end{bmatrix}$$

$$\pi_1 + \pi_2 = 1$$

$$\pi_1 = .95\pi_1 + .2\pi_2$$
$$\pi_2 = .05\pi_1 + .8\pi_2$$
$$\pi_1 + \pi_2 = 1$$

Solving

$$\pi_1 = .8$$
$$\pi_2 = .2$$

- New profit in a year for Coke company will be
  = 0.8 * (5.2 billion * $1) - $500 million
  =$3,660,000,000
- So, they should hire the add firm

Long Run Property of Markov Chain

Steady State Probabilities

- Gardener Problem **(with fertilizer)**

$$P = \begin{bmatrix} .3 & .6 & .1 \\ .1 & .6 & .3 \\ .05 & .4 & .55 \end{bmatrix}$$

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$$

- System of eqn $\sum_{j=1}^{s} \pi_j = 1$ yields following set of equations

$$\pi_1 = .3\pi_1 + .1\pi_2 + .05\pi_3$$

$$\pi_2 = .6\pi_1 + .6\pi_2 + .4\pi_3$$

$$\pi_3 = .1\pi_1 + .3\pi_2 + .55\pi_3$$

$$\pi_1 + \pi_2 + \pi_3 = 1$$

$$.7(1-(\pi_2+\pi_3)) = .1\pi_2 + .05\pi_3 \quad => .8\pi_2 + .75\pi_3 = .7$$

$$.4\pi_2 = .6(1-(\pi_2+\pi_3)) + .4\pi_3 \quad => \pi_2 + .2\pi_3 = .6$$

$$.8(.6-.2\pi_3) + .75\pi_3 = .7$$

$$=> .59\pi_3 = .22 \quad => \pi_3 \approx .3729$$

$$so, \pi_2 = .6 - .2\times.3729 \approx .5254$$

$$\pi_1 = 1-(.3729+.5254) = .1017$$

Therefore, **π** = [.1017 .5254 .3729]

Mean First Passage Times

- For an ergodic 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.
  - Mean/expected hitting time
- assume 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.

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

Since

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

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

Mean First Passage Times

- By solving the linear equations of the equation above, we find all the mean first passage times.

$$m_{ij} - p_{i1}m_{1j} - p_{i2}m_{2j} - \ldots - p_{is}m_{sj} = 1$$

$$\begin{aligned}
m_{1j} - p_{11}m_{1j} - p_{12}m_{2j} - \ldots - p_{1s}m_{sj} &= 1 \\
m_{2j} - p_{21}m_{1j} - p_{22}m_{2j} - \ldots - p_{2s}m_{sj} &= 1 \\
&\vdots \\
m_{sj} - p_{s1}m_{1j} - p_{s2}m_{2j} - \ldots - p_{ss}m_{sj} &= 1
\end{aligned}$$

- In Matrix form

$$\begin{bmatrix} 1-p_{11} & -p_{12} & \cdots & -p_{1s} \\ -p_{21} & 1-p_{22} & \cdots & -p_{2s} \\ \vdots & \vdots & & \vdots \\ -p_{s1} & -p_{s2} & \cdots & 1-p_{ss} \end{bmatrix} \begin{bmatrix} m_{1j} \\ m_{2j} \\ \vdots \\ m_{sj} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{m_{ij}} = (\mathbf{I} - \mathbf{N_j})^{-1}\mathbf{1} \; ; \; j \neq i$$

Mean First Passage Times

- By solving the linear equations of the equation above, we find all the mean first passage times.

- In Matrix form $\mathbf{m}_{\mathbf{ij}} = (\mathbf{I} - \mathbf{N}_{\mathbf{j}})^{-1}\mathbf{1}\ ;\ j \neq i$

$\mathbf{I} = (m - 1)$-identity matrix

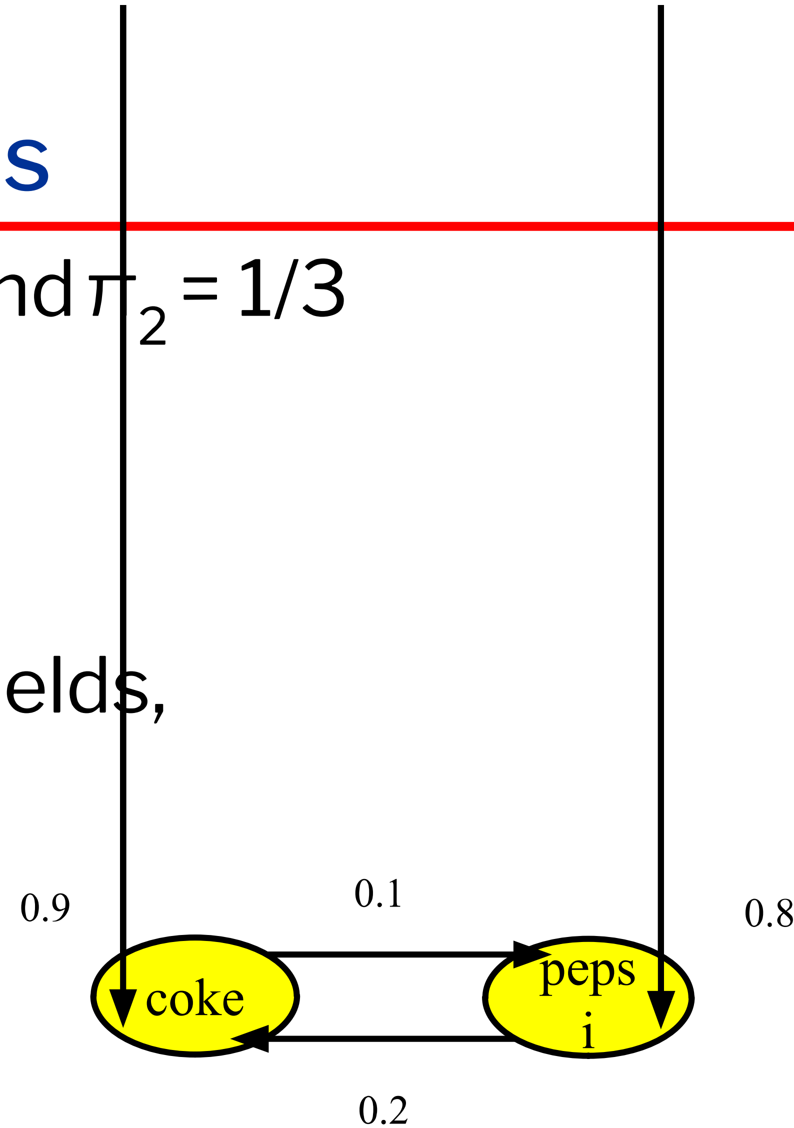
$\mathbf{N}_j$ = transition matrix $\mathbf{P}$ less its jth row and jth column of target state j

$\mathbf{1} = (m - 1)$ column vector with all elements equal to 1

- It can be shown that **mean recurrence time** $m_{ii} = \dfrac{1}{\pi_i}$

Mean First Passage Times

- For the cola example, π_1=2/3 and π_2 = 1/3
  - Hence, m_{11} = 1.5 and m_{22} = 3
  - $m_{12} = 1 + p_{11}m_{12} = 1 + .9m_{12}$
  - $m_{21} = 1 + p_{22}m_{21} = 1 + .8m_{21}$
- Solving these two equations yields,
  - m_{12} = 10 and m_{21} = 5



Mean First Passage Times

- Gardener Problem (with fertilizer)

$$\mathbf{P} = \begin{pmatrix} .30 & .60 & .10 \\ .10 & .60 & .30 \\ .05 & .40 & .55 \end{pmatrix}$$

consider the passage from states 2 and 3 (fair and poor) to state 1 (good). Thus, $j = 1$ and

$$\mathbf{N}_1 = \begin{pmatrix} .60 & .30 \\ .40 & .55 \end{pmatrix}, (\mathbf{I} - \mathbf{N}_1)^{-1} = \begin{pmatrix} .4 & -.3 \\ -.4 & .45 \end{pmatrix}^{-1} = \begin{pmatrix} 7.50 & 5.00 \\ 6.67 & 6.67 \end{pmatrix}$$

$$\begin{pmatrix} m_{21} \\ m_{31} \end{pmatrix} = \begin{pmatrix} 7.50 & 5.00 \\ 6.67 & 6.67 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 12.50 \\ 13.34 \end{pmatrix}$$

on the average, it will take 12.5 seasons to pass from fair to good soil and 13.34 seasons to go from bad to good soil.

Similar calculations can be carried out to obtain μ_{12} and μ_{32} from $(\mathbf{I} - \mathbf{N}_2)$ and μ_{13} and μ_{23} from $(\mathbf{I} - \mathbf{N}_3)$, as shown below.

References

- Operations Research : Applications and Algorithms

  Wayne L. Winston
- Operations Research An Introduction

  Hamdy A. Taha