Chapter 5
Elementary Stochastic Analysis

Prof. Ali Movaghar

Goals in this chapter

- Characterization of random processes
- Analysis of discrete and continuous time Markov chains
- Long term behavior of random processes
- Birth and death processes
- The methods of stages

Random Process

- A random process can be thought of as a random variable that is a function of time.
  - $X(t)$ is a random process.
  - At time, τ, $X(\tau)$ denotes an ordinary random variable with distribution $F_{X(\tau)}(x)$.
  - $X(t)$ describes the state of a stochastic system as a function of time.

Classification of Random Processes

- State space : the domain of values taken by X(t) can be discrete or continuous.
- Time parameter: which could be discrete or continuous.

Classification of Random … (Con.)

- Variability : refers to the time-related behavior of the random process.
  - A stationary process is a random process does not change its property with time. For a non-stationary
    - $X(t_i)$ and $X(t_{i+1})$ may have different distribution.
    - For i and j, j>i, the dependance between $X(t_i)$ and $X(t_j)$ varies with the time origin t_i

Classification of Random … (Con.)

- Correlation aspect : let $t_1<t_2<\ldots<t_n$. The random variable $X(t_n)$
  - may be independent of $X(t_1), \ldots X(t_{n-1})$: Independent process
  - Only depends on $X(t_{n-1})$: Markov Process

Markov Processes

- Let $t_1<t_2<\ldots<t_n$. Then $P(X(t_n)=j)$ denotes the probability of finding the system in state j at time t_n.
- The Markovian property states that :

$$P[X(t_n)=j \mid X(t_{n-1})=i_{n-1}, X(t_{n-2})=i_{n-2},\ldots] = P[X(t_n)=j \mid X(t_{n-1})=i_{n-1}]$$

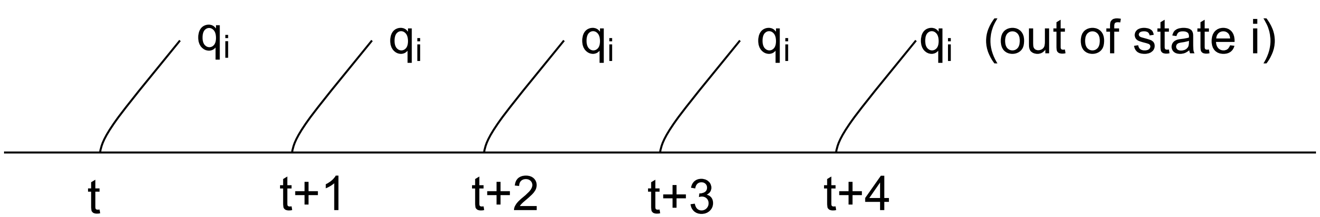
- Note that $t_n, t_{n-1}, t_{n-2}\ldots$ above are arbitrary time instants and j, i_{n-1}, i_{n-2}, … are arbitrary state values.

Markov Processes (Con.)

- The process is memoryless to
  - the states visited in the past
  - the time already spent in the current state

Markov Processes (Con.)

- State residence time of a Markov process can not be arbitrary
  - for a homogeneous Markov process is
    - Geometric for the discrete state space domain.



$$Pr\text{ (resident time = n)} = (1-q_i)^{n-1}q_i$$

- Exponential for the continuous state space domain.

Independent Processes

- Let X_i denote the time between the occurrence of (i-1)th and ith events.
- If the X_i 's are independent, identically distributed random variables, X_i is a renewal process.
- Renewal process is a special case of the independent processes.
  - Example : the arrival process to a queuing system, successive failures of system components.

Independent Processes (Con.)

- Let $F_x(t)$ and $f_X(t)$ denote the distribution and density function of the X_i's.
- Let $S_n = X_1 + \ldots + X_n$ denoting the time until the occurrence of nth event.
  - The distribution of S_n is n-fold convolution of $F_X(t)$ with itself ➔ $F_{(n)}(t)$
- The process S_n is also known as the random walk process.

Independent Processes (Con.)

- Let N(t) denote the number of renewals in time (0,t].
- This process is called renewal counting process; it is stationary, discrete state random process. It is easy to see that
  - $P[\,N(t) \geq n] = P(S_n \leq t) = F_{(n)}(t)$

Poisson Process

- Let renewal process with X_i's are exponentially distributed
  - $f_X(x) = \lambda e^{-\lambda t} \Rightarrow P(N(t)=n)=(\lambda t^n/n!)e^{-\lambda t}$
- Thus N(t) has the Poisson distribution. The resulting renewal counting process is also called a Poisson process.

Poisson Process (Con.)

- A Poisson process has the following properties:
  - Relation to exponentional distribution : Interevent time is exponentially distributed
  - Uniformity: Probability of more than one arrival in a small interval is negligible
  - Memorylessness: Past behavior is totally irrelevent
  - Mixture: Sum of k independent Poisson streams with rates $\lambda_1,\ldots, \lambda_k$ is also Poisson with rate $\lambda = \lambda_1+\ldots+ \lambda_k$.

Poisson Process (Con.)

- Probability split: A k-way probabilistic probabilistic split of a Poisson stream with probabilities $q_1,\ldots,q_k$ creates k independent Poisson substream with rates $q_1 \lambda,\ldots,q_k \lambda$.
- Limiting probability: The average random process of k independent renewal counting processes, $A_1(t),\ldots,A_k(t)$ each having arbitrary interevent-time distribution with finite mean (m_i) is

$$X(t)=[A_1(t)+\ldots+A_k(t)]/k$$

which has Poisson distribution with rate $k/\Sigma(1/m_i)$ as $k \rightarrow \infty$

Analysis of Markov Chains

- State probability $\pi_j(t)=P(X(t)=j)$ denotes the probability that there are j customers in the system at time t.
- Transition probability $P_{ij}(u,t) = P[X(t) = j \mid X(u) = i]$ denotes the probability that the system is in state j at time t, given it was in state i at time u.

Analysis of Markov Chains

- $\pi_j(t) = \Sigma\ \pi_i(u)\ P_{ij}(u,t)$, (*)
- We can rewrite (*) formula in a matrix form.
  - $\Pi(t) = [\pi_0(t), \pi_1(t), \ldots]$ as a row vector
  - $H(u,t)$ as the square matrix of $P_{ij}(u,t)$'s
  - $\Pi(t) = \Pi(u)\ H(u,t)$, (**) gives the state probability vector for both discrete and continuous parameter case.

Analysis of Markov Chains

- In following we show how to find state probability vector for both transient and steady-state behavior of discrete and continuous parameter cases.

Discrete Parameter Case

- Solving Π(t) = Π(u) H(u,t) when time is discrete:
  - Define Q(n) as H(n,n+1) and let u=n and t=n+1:

    Π(n+1) = Π(n) Q(n)
  - A representative element of Q(n) is $P_{ij}(n,n+1)$ denoted as $q_{ij}(n)$ and called one-step transition
  - $q_{ij}(n) \in [0,1]$ and, for each row, $\sum_{j=1}^{n} q_{ij}(n) = 1$
  - Homogenous chains are stationary and so $q_{ij}(n)$'s are independent of the time parameter n: Π(n+1) = Π(n)Q (***)
  - Π(0) is known and so Π(n) can be computed
  - A general expression for Π(n) (transient probability), z-transform is used

Discrete Parameter Case (Con.)

- Let the z-transform of $\Pi(n)$ denoted as $\Phi(z)$.
  - Like $\Pi(n)$, $\Phi(z)$ is a vector $[\phi_0(z), \phi_1(z), \ldots]$ where $\phi_i(z)$ is the z-transform of the probability of being in state i

$$\phi_i(z) = \sum_{k=0}^{\infty} \pi_i(k) z^k$$

  - Multiplying both side of equation (***) by z^{n+1}, summing over all n, we get $\Phi(z) - \Pi(0) = \Phi(z)Qz$, which simplifies to give

$$\Phi(z) = \Pi(0)[I - Qz]^{-1}$$

  - $\Pi(n)$ can be retrieved by inverting $\Phi(z)$. $\Pi(n)$ is the probability state vector for transient behavior of discrete Markov chain.

Discrete Parameter Case (Con.)

- to find $\Phi(z)$, det(I-Qz) should be equal to 0 (remember that the equation $xQ=\lambda x$ has a nontrivial solution x if and only if $(Q-\lambda I)$ is singular or $\det(Q-\lambda I)=0$)
- The equation det(I-Qz)=0 is *characteristic equation* of the Markov chain and its root are *characteristic roots*.
- to invert $\phi_i(z)$, find its partial-fraction expansion and use appendix E to find its inverse. (Study example B.2)

Discrete Parameter Case (Con.)

- Generally the partial-fraction expansion of $\phi_i(z)$ for any state i, will have denominator of the form $(z-r)^k$, where r is characteristic root.
- The inversion of $\phi_i(z)$ to get $\pi_i(n)$ is a sum of terms, each of which contains terms like r^{-n}.
- If system is stable, such terms when $n \to \infty$ cannot blow up. So non-unity roots must be larger than 1.
- Since at least one of the $\pi_i(n)$ must be nonzero when $n \to \infty$, at least one of the roots should be unity.

Discrete Parameter Case (Con.)

- The limiting (or steady-state) behavior of the system denoted as Π is defined by $\lim_{n\to\infty} \Pi(n)$
- If limiting distribution (Π) is independent of the initial conditions, we can obtain it more easily
  - Let e denote a column of all 1's.
  - $\Pi = \Pi Q$ and $\Pi . e = 1$
- A Markov chain can be represented by a directed graph, known as transition diagram. The nodes represent the states and the arcs represent the transition. The arcs are labeled by transition probabilities.

Discrete Parameter Case (Con.)

- **Example:** consider a discrete parameter Markov chain with the following single-step transition probability matrix

$$\begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{bmatrix}$$

- Draw the state transition diagram for this chain. Assuming that the system is in state 1 initially, compute the state probability vector $\Pi(n)$ for $n=1,2,\ldots,\infty$. Also compute the characteristic roots of Q and characterize the limiting behavior.

Discrete Parameter Case (Con.)

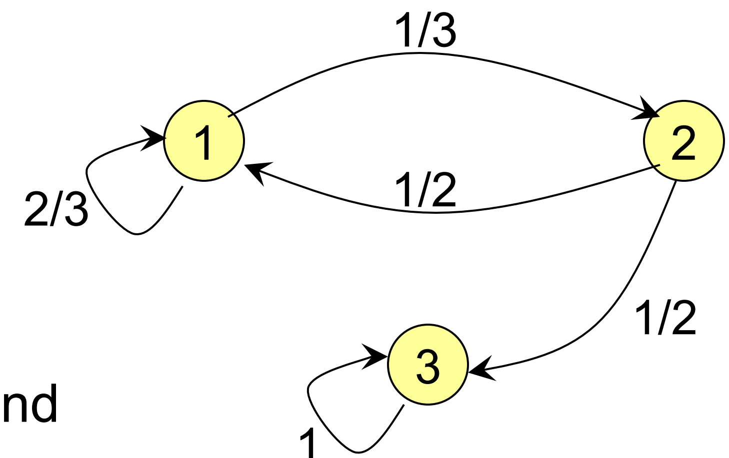
- **Solution:** the first few values of Π(n) can be obtained by simple matrix multiplication. For general case :

$$I - Qz = \begin{bmatrix} 1-2z/3 & -z/3 & 0 \\ -z/2 & 1 & -z/2 \\ 0 & 0 & 1-z \end{bmatrix}$$

$$(1-z)(1-2z/3-z^2/6)=0,\ \ z_0=1,\ \ z_1=-2+\sqrt{10},\ z_2=-2-\sqrt{10}$$

  - As expected one root is unity and the others have magnitude larger than 1. So the Markov is stable.

Discrete Parameter Case (Con.)



- we can see that $n \rightarrow \infty$, the system settle in state 3 with probability 1. therefore the unity root appear only in $\phi_3(z)$.
- The other two roots appear in all ϕ_i's and will lead to a damped oscillatory effect which eventually die down to zero.

Since $\Pi(0)$ =[1 0 0], the first row of $[I-Qz]^{-1}$ is itself $[\phi_1(z), \phi_2(z), \phi_3(z)]$

- $\phi_1(z)=-6/(z^2+4z-6)$ then $\pi_1(n)=0.9487[z_1^{-n-1}-z_2^{-n-1}]$
- $\pi_2(n)$ and $\pi_3(n)$ can be computed similarly. To find out more about z-transform study B.3.1

Continuous Parameter Case

- We start again from $\Pi(t) = \Pi(u)\, H(u,t)$.
  - Let $u = t - \Delta t$.
  - Then $\Pi(t) - \Pi(t-\Delta t) = \Pi(t-\Delta t)\,[H(t-\Delta t,t) - I]$.
  - Divide by Δt and taking limit as $\Delta t \to 0$, we get the following basic equation called (forward) *chapman-kolmogorov* equation:

$$\frac{\partial \Pi(t)}{\partial t} = \Pi(t)Q(t) \quad \text{where } Q(t) = \lim_{\Delta t \to 0} \frac{H(t-\Delta t, t) - I}{\Delta t} \; (*)$$

  - Let $q_{ij}(t)$ denote the (i,j) element of $Q(t)$.
  - Let δ_{ij} denote the Kronecker delta function; $\delta_{ii}=1$ otherwise 0

$$q_{ij}(t) = \lim_{\Delta t \to 0} \frac{P_{ij}(t-\Delta t, t) - \delta_{ij}}{\Delta t}$$

Continuous Parameter Case (Con.)

- As $\Delta t \to 0$, then
$$1 - P_{ii}(t - \Delta t, t) = -q_{ii}(t)\Delta t$$
$$P_{ij}(t - \Delta t, t) = q_{ij}(t)\Delta t \quad \text{for } i \neq j$$
- Thus $q_{ij}(t)$ for $i \neq j$ can be interpreted as the rate at which the system goes from state i to j, and $-q_{ii}(t)$ as the rate at which the system departs from state i at time t.
- Because of this interpretation, Q(t) is called transition rate matrix such that :
$$\sum_{j \neq i} P_{ij}(t - \Delta t, t) + P_{ii}(t - \Delta t, t) = 1$$
- Applying the limit as $\Delta t \to 0$ to above results $\sum_{j=0}^{\infty} q_{ij}(t) = 0$
  - All elements in a row of Q(t) must sum to 0
  - The off-diagonal elements of Q(t) must be nonnegative while those along diagonal must be negative.

Continuous Parameter Case (Con.)

- Solving differential equation (*) gives

$$\frac{\partial \Pi(t)}{\partial t} = \Pi(t)Q(t) \quad \Rightarrow \Pi(t) = \Pi(0)\exp\left[\int_0^t Q(u)du\right]$$

- Again we focus on homogenous chains; the transition rates are independent of time:

$$\frac{\partial \Pi(t)}{\partial t} = \Pi(t)Q \;\Rightarrow \Pi(t) = \Pi(0)\exp(Qt)$$

  - To solve above equation, we use Laplace transforms
  - Let ψ(s) denote the Laplace transform of Π(t)
  - Using the differential property of Laplace transform, above equation yields : $\Psi(s) = \Pi(0)[sI - Q]^{-1}$

Continuous Parameter Case (Con.)

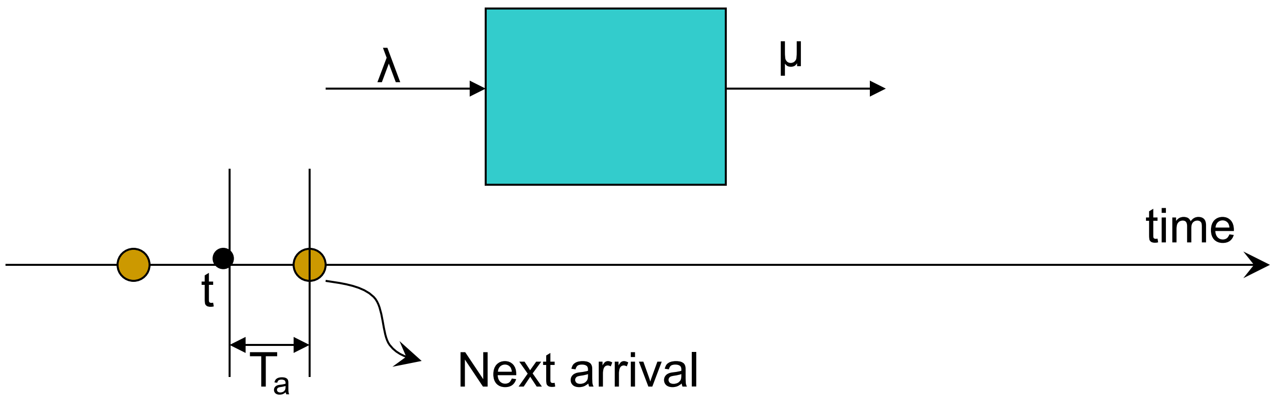
- The limiting (or steady-state) behavior of the system denoted as Π is defined by $\lim_{t\to\infty} \Pi(t)$
- If limiting distribution (Π) exists and is independent of the initial conditions, the derivation $\frac{\partial \Pi(t)}{\partial t}$ must be zero, and we get
  - $\Pi Q=0$ and $\Pi.e=1$
  - It is somehow similar to the equations of discrete parameter case

Converting continuous to discrete parameter

- We can go from a discrete parameter system to a continuous one and vice versa
  - Let Q is a transition rate matrix; its diagonal elements must be negative and largest in magnitude in each row.
  - Let δ be a some positive number larger than all diagonal elements of Q
  - $Q' = \delta^{-1}Q + I$
  - Obviously Q' is a transition probability matrix.

Example

- Obtain steady–state queue length distribution for an open M/M/1 system



- $P_{k,k+1}(t,t+\Delta t)=\Pr(T_a \leq \Delta t)=1-e^{-\lambda\Delta t}= \lambda\Delta t+o(\Delta t)$
- $q_{k,k+1}=\lambda$ for any k and similarly $q_{k,k-1}=\mu$
- Thus λ and μ are the forwards and backwards transition rates.

Example (Con.)

$$\begin{bmatrix}\pi_0 & \pi_1 & \pi_2 & \dots\end{bmatrix}\begin{bmatrix}-\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \dots \\ 0 & \mu & -(\lambda+\mu) & \lambda & \dots \\ 0 & 0 & \mu & -(\lambda+\mu) & \dots \\ . & . & . & . & \dots\end{bmatrix} = \begin{bmatrix}0\end{bmatrix}$$

- This infinite system has a very simple solution.
  - $-\lambda\pi_0+\mu\pi_1=0 \rightarrow \pi_1=\rho\pi_0$ where $\rho=\lambda/\mu$
  - $-\lambda\pi_0-(\lambda+\mu)\pi_1+\mu\pi_2=0 \rightarrow \pi_2=\rho^2\pi_0$
  - ....
  - $\pi_n=\rho^n\pi_0$
  - $\pi_0+\dots+\pi_n=1 \rightarrow \pi_0=(1-\rho)$

Long-Term Behavior of Markov Chains

- Consider a discrete Markov chain Z.
  - Let its limiting distribution be $\Pi = \lim_{t\to\infty} \Pi(t)$
  - Let Π^* be the solution of system equation $\Pi = \Pi Q$ and $\Pi.e=1$. Π^* is called stationary distribution.
  - Run Z with different (and arbitrary) initial distribution. Three possibilities exits:
    - It always settles with a same Π^* ; we have a unique limiting distribution which is equal to stationary distribution.
    - It never settles down. No limiting distribution exists.
    - It always settles, but long-term distribution depends on the initial state. The limiting distribution exists, but is non-unique.

State Classification and Ergodicity

- We introduce a number of concepts that identify the conditions under which a Markov chain will have unique limiting distribution.
  - Let $f_{ii}^{(n)}$ denote the probability that the system, after making a transition while in state i, goes back to state i for the first time in exactly n transitions.

$$\underbrace{i \to \dots \to j \to \dots \to k \to \dots \to i}_{n \text{ steps}}$$

    - If state i has a self-loop, $f_{ii}^{(1)}>0$, otherwise $f_{ii}^{(1)}=0$
    - $f_{ii}^{(0)}$ is always zero.

State Classification … (Con.)

- Let f_{ii} denote the probability that the system ever returns to state i;

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)}$$

  - If $f_{ii}=1$, every time the system leaves state i, it must return to this state with probability 1.
    - During infinitely long period of time, the system must return to state i infinitely often.
    - State i is called recurrent state.
  - If $f_{ii}<1$, each time the system leaves state i, there is a finite probability that it does not come back to state i.
    - Over an infinitely long observation period, the system can visit state i only finitely often.
    - State i is called transient state.

State Classification … (Con.)

- Recurrent states are further classified depending on whether the eventual return occur in a finite amount of time.
- Let θ_{ii} denote the expected time until it reenters state i.

$$\theta_{ii} = \sum_{n=1}^{\infty} n f_{ii}^{(n)}$$

  - If $\theta_{ii} = \infty$, we refer state i as null recurrent.
  - Otherwise state i is called positive recurrent.

State Classification … (Con.)

- If $f_{ii}^{(n)}>0$ only when n equals some integer multiple of a number k>1, we call state i periodic.
- Otherwise (i.e. if $f_{ii}^{(n)}>0$ and $f_{ii}^{(n+1)}>0$ for some n) , state i is called aperiodic.
- A Markov chain is called irreducible, if every state is reachable from every other state (strongly connected graph).

State Classification … (Con.)

- Lemma 1. All states of an irreducible Markov chain are of the same type (i.e. transient, null recurrent, periodic, or positive recurrent and aperiodic).
  - Furthermore, in the periodic case, all states have a same period.
  - We can name an irreducible chain according to its state type.

State Classification … (Con.)

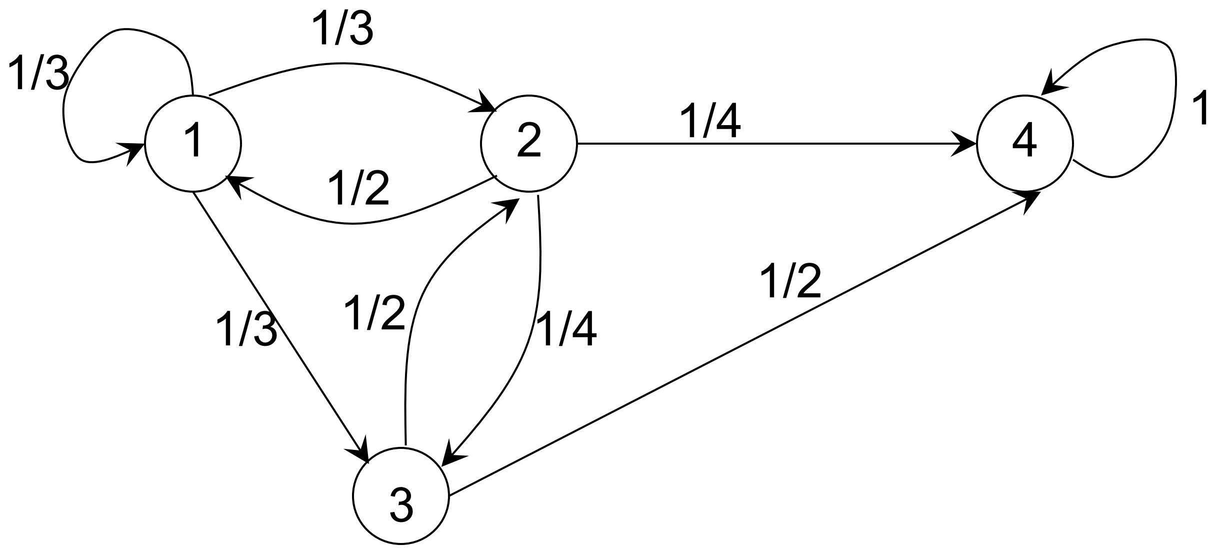
- A Markov chain is called ergodic if it is irreducible, positive recurrent and aperiodic.
  - Aperiodicity is relevant only for discrete time chains.

State Classification … (Con.)

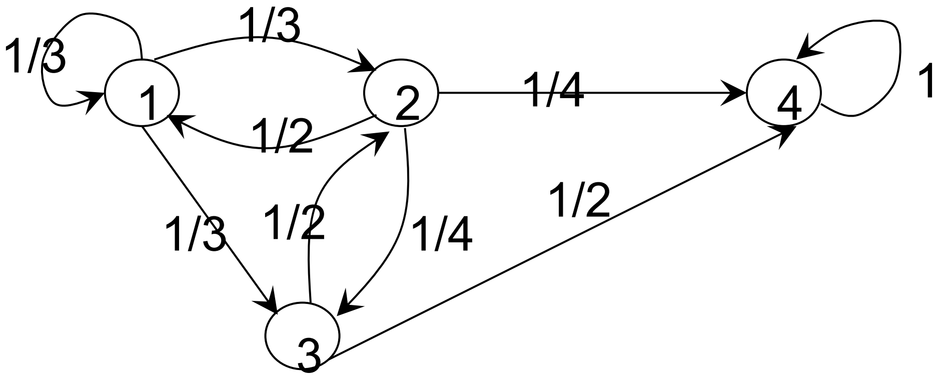
- Lemma 2. An ergodic Markov chain has a unique limiting distribution, independent of the initial state.
  - It is given by $\Pi = \Pi Q$ and $\Pi . e = 1$ for discrete-time case and by $0 = \Pi Q$ and $\Pi . e = 1$ for continuous-time case.

Example 1.

- Classify all states of the discrete-time Markov chain whose state diagram is shown below.



Example 1. (Con.)



- $f_{11}^{(1)}=1/3$, $f_{11}^{(2)}=(1/3)(1/2)=1/6$, $f_{11}^{(3)}=(1/3)(1/2)(1/2)=1/12$

$$f_{11}^{(n)} = \frac{1}{3}.\left(\frac{1}{2}.\frac{1}{4}\right)^{\frac{n-3}{2}}.\frac{1}{2}.\frac{1}{2} \quad \text{for } n > 3 \text{ and odd}$$

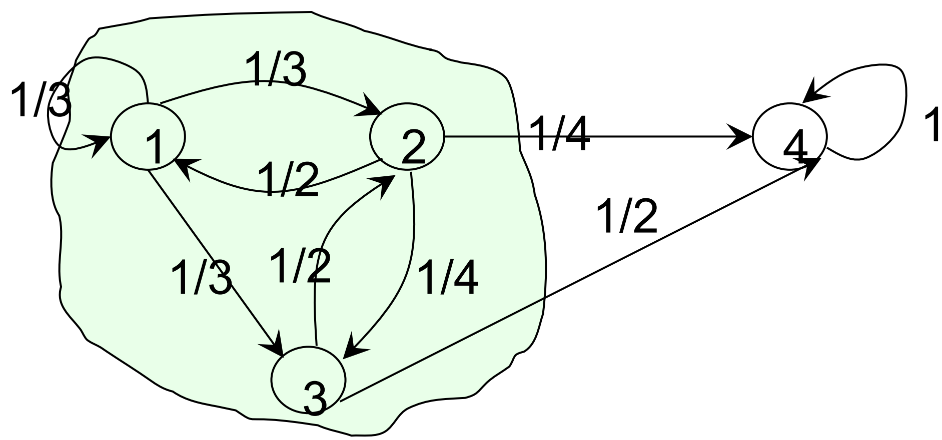
$$f_{11}^{(n)} = \frac{1}{3}.\left(\frac{1}{4}.\frac{1}{2}\right)^{\frac{n-2}{2}}.\frac{1}{2} \quad \text{for } n > 2 \text{ and even}$$

$$f_{11} = \sum_{n=1}^{\infty} f_{11}^{(n)} = \frac{1}{3} + \frac{1}{6}\sum_{m=0}^{\infty}\left(\frac{1}{8}\right)^{m} + \frac{1}{12}\sum_{m=0}^{\infty}\left(\frac{1}{8}\right)^{m} = \frac{13}{21}$$

- State (1) is transient

Example 1. (Con.)

- State (4) is obviously recurrent since $f_{44}=f_{44}^{(1)}=1$
- Without computing f_{22} and f_{33} , we can claim that states (2) and (3) are transient



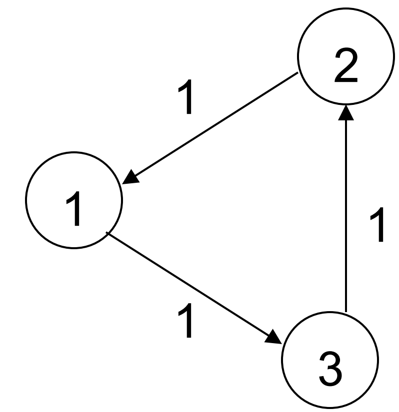
Since the sub-graph consisting node 1, 2, and 3 is strongly connected.

Example 2.

- Consider a discrete-time Markov chain with Q matrix shown below. Classify its states and the long-term behavior. Next consider (Q-I) as the transition rate matrix for a continuous parameter.

$$Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Example 2. (Con.)



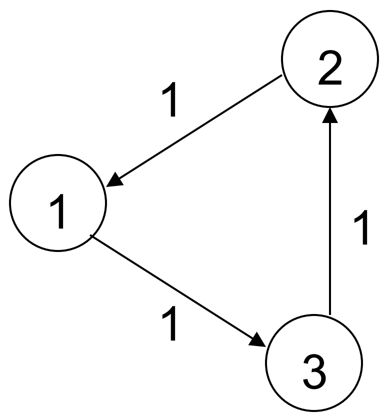
- For discrete case :
  - Chain is irreducible
  - Because of the finite number of states, it is positive recurrent
  - Since the chain cycles through its three states sequentially, it is periodic with period 3, $f_{ii}=f_{ii}^{(3)}=1$
  - Suppose that the chain is state 1 initially, $\Pi(0)=[1\ 0\ 0]$. It is easy to verify using the relation $\Pi(n)=\Pi(0)Q^{(n)}$ that

$$\Pi(n) = \begin{cases} [1\ 0\ 0\ ] \text{ for } n=0, 3, 6, \ldots \\ [0\ 1\ 0\ ] \text{ for } n=1, 4, 7, \ldots \\ [0\ 0\ 1\ ] \text{ for } n=2, 5, 8, \ldots \end{cases}$$

  Therefore no limiting or steady-state distribution exists.
  - If $\Pi(0)=[1/3\ 1/3\ 1/3]$, we get $\Pi(n)=\Pi(0)$.
  - System is nonergodic chain and the limiting distribution depends on initial state.

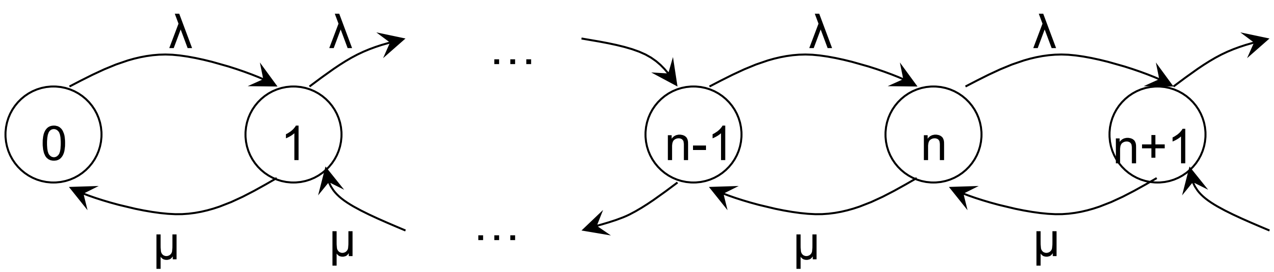
Example 2. (Con.)

- For Continuous case:
  - The chain is no longer periodic .
  - The limiting distribution can be easily obtained as [1/3 1/3 1/3]



Example 3.

- Characterize the Markov chain for the simple M/M/1 queuing system.



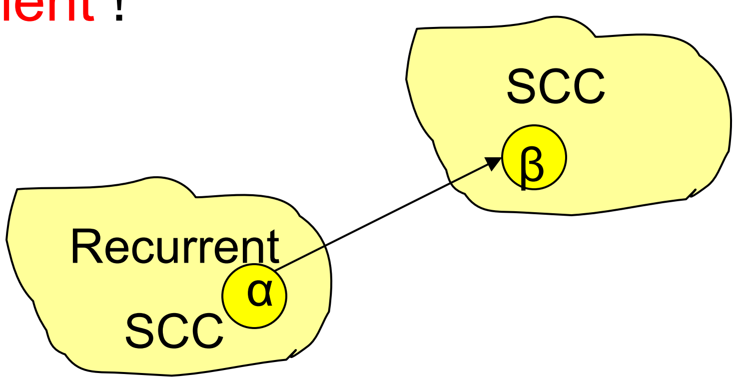
- The chain irreducible.
  - If $\lambda>\mu$, we find that $f_{ii}<1$ and hence chain is transient.
  - If $\lambda\leq\mu$, we find that $f_{ii}=1$ and hence chain is recurrent.
    - If $\lambda<\mu$, chain is positive recurrent.
    - If $\lambda=\mu$, chain is null recurrent.

Analysis of Reducible Chains

- The limiting behavior of reducible chain necessarily depends on the initial distribution.
  - Because not every state is reachable from every initial state.
- We can decompose a reducible chain into maximal strongly connected components (SCC)
  - All states in a SCC are of the same type.
  - In long run, system could only be in one of the recurrent SCC.

Analysis of Reducible Chains (Con.)

- A recurrent SCC cannot have any transition to any state outside that SCC
  - Proof by contradiction
    - There is transition from α to some β
    - As SCC is maximal, there is no path from β to α
    - So α is transient !



Analysis of Reducible Chains (Con.)

- A recurrent SCC cannot have any transition to any state outside that SCC

⇩

- There is no path between various recurrent SCC's. ⇩
- The limiting distribution of chain can be obtained from those of SCC's :
  - If all recurrent SCC have a unique limiting distribution, then so does chain (depending on initial state)

Analysis of Reducible Chains (Con.)

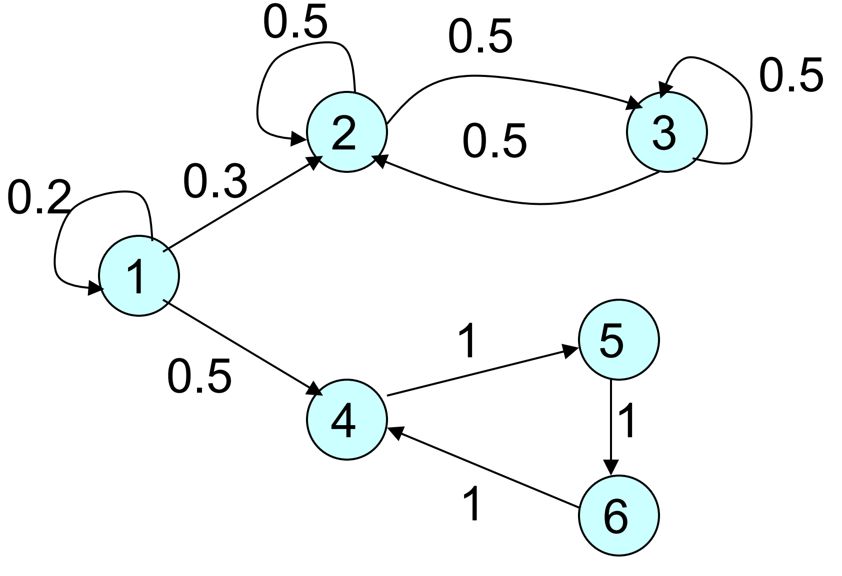
- Define all SCC's in the chain
- Replace each SCC by a single state
  - There is only a transition from a transient state i to a recurrent SCC j.
  - The transition rates can be easily determined

$$q^*_{ij} = \sum_{\forall k \in SCC(j)} q_{ik}$$

- Solve new chain by Z (or Φ) transform method
  - $P_1, \ldots, P_k$ denote limiting state probability for $SCC_1 \ldots SCC_k$
  - $\Pi_i=[\pi_{i1}, \pi_{i2},\ldots]$ denote the stationary distribution for ith SCC.
  - The limiting probability for state k of SCC i is $P_i\, \pi_{ik}$.

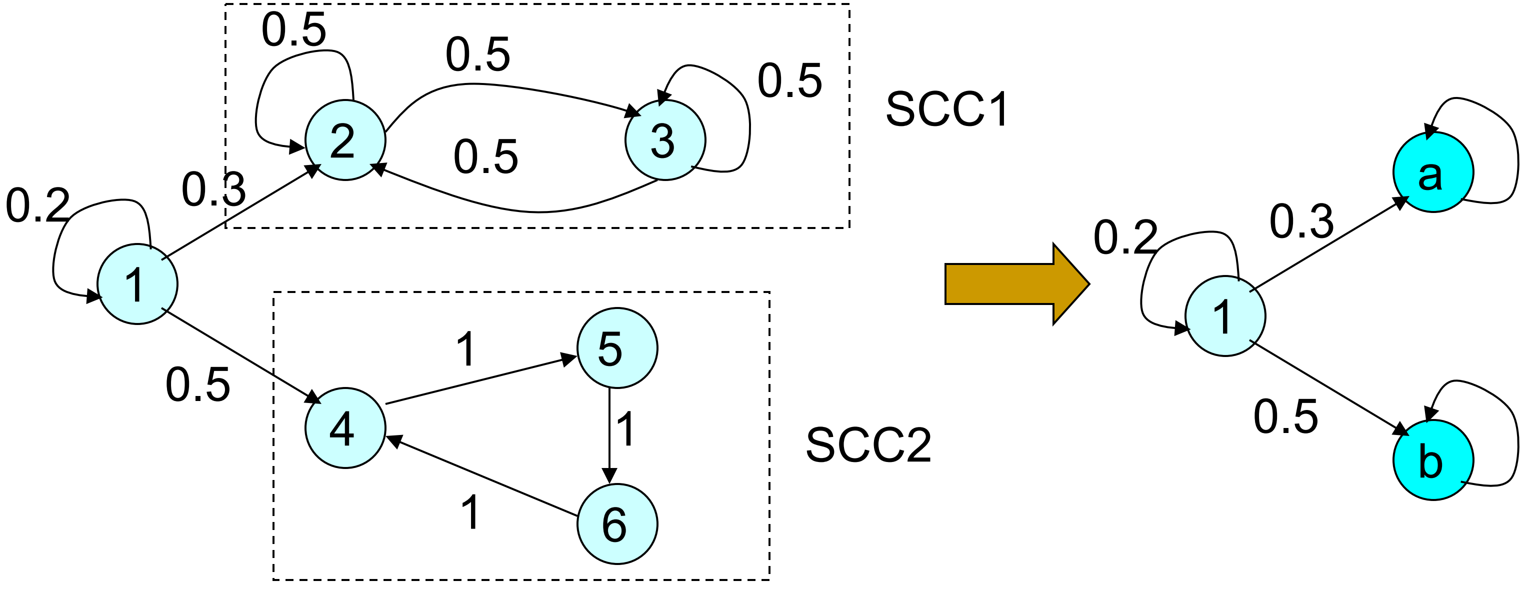
Example

- Characterize the limiting distribution for discrete-time Markov chain shown below. Assume that system is initially in state 1.



Example (Con.)

- Solution.



  - SCC1 is unconditionally ergodic whereas SCC2 is ergodic if the time is not discrete.

Example (Con.)

- The limiting distribution of SCC1 is [1/2 ½] and of SCC2 is [1/3 1/3 1/3]
- Since Π(0)=[1 0 0], we only need to compute the first row of the matrix $[I-Qz]^{-1}$:

$$\begin{bmatrix} \frac{1}{1-0.2z} & \frac{0.3z}{(1-z)(1-0.2z)} & \frac{0.5z}{(1-z)(1-0.2z)} \end{bmatrix}$$

- Π= [0 0.375 0.625] as indicated in Figure P_a =0.3/0.8=0.375