Lecture 2: Markov Models

Ziyu Shao

School of Information Science and Technology
ShanghaiTech University

March 07, 2025

Outline

Outline

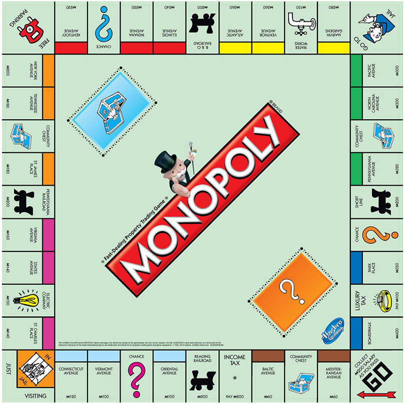
Definition

X_t $t \in [0,1]$

- A stochastic process is a collection of random variables $\{X_t, t \in I\}$. The set I is the *index set* of the process. The random variables are defined on a common state space $\mathcal{S}$.
- I is discrete: discrete-time stochastic processes (sequences of random variables)
- I is continuous: continuous-time stochastic processes (uncountable collections of random variables)

Example: Discrete Time & Discrete State Space

- State Space: $\{1, \ldots, 40\}$
- X_k: the player's board position after k dice rollings.
- Stochastic Process for Monopoly: $X_0, X_1, \ldots$



Example: Discrete Time & Continuous State Space

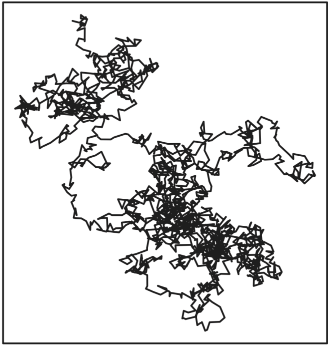
- Air-monitoring with PM2.5 measurements every hour
- State Space: $(0, 2000)$
- X_k: the PM2.5 measurement at the kth hour.
- Stochastic Process for Air-monitoring: $X_0, X_1, \ldots$

Example: Continuous Time & Discrete State Space

- We receive emails at random times day and night.
- State Space: $\{0, 1, 2, \ldots\}$
- $X_t, t \in [0, \infty)$: the number of emails we receive up to time t
- Stochastic Process for Email: $\{X_t\}$

Example: Continuous Time & Continuous State Space

- Two-dimensional Brownian Motion
- State Space: $\mathbb{R}^2$
- $X_t, t \in [0, \infty)$: position of the particle at time t
- Stochastic Process for random motion of particles: $\{X_t\}$



Outline

Model Selection in Stochastic Modeling

- Enough complexity to capture the complexity of the phenomena in question
- Enough structure and simplicity to allow one to compute things of interest

Motivation

- Introduced by Andrey Markov in 1906
- IID sequence of random variables: too restrictive assumption
- Completely dependent among random variables: hard to analysis
- Markov chain: happy medium between complete independence & complete dependence.

Markov Model

Three basic components of Markov model

- A sequence of random variables $\{X_t, t \in \mathcal{T}\}$, where $\mathcal{T}$ is an index set, usually called "**time**".
- All possible sample values of $\{X_t, t \in \mathcal{T}\}$ are called "**states**", which are elements of a sate space $\mathcal{S}$.
- "**Markov property**": given the present value(information) of the process, the future evolution of the process is independent of the past evolution of the process.

Classification of Markov Model

- *Discrete-Time Markov Chain*: Discrete $\mathcal{S}$ & Discrete $\mathcal{T}$
- *Continuous-Time Markov Chain*: Discrete $\mathcal{S}$ & Continuous $\mathcal{T}$
- *Discrete Markov Process*: Continuous $\mathcal{S}$ & Discrete $\mathcal{T}$ (MCMC)
- *Continuous Markov Process*: Continuous $\mathcal{S}$ & Continuous $\mathcal{T}$

Our focus: Discrete(Continuous)-Time Markov Chain with finite state space

Outline

Markov Chain

$P(A|B,C) = P(A|B)$

given B, A and C are independent

Definition

A sequence of random variables $X_0, X_1, X_2, \ldots$ taking values in the *state space* $\{1, 2, \ldots, M\}$ is called a *Markov chain* if for all $n \geq 0$,

$$\underbrace{P(X_{n+1} = j}_{\text{future}} | \underbrace{X_n}_{\text{present}} = i, \underbrace{X_{n-1} = i_{n-1}, \ldots, X_0 = i_0}_{\text{history}}) = P(\underbrace{X_{n+1}}_{\text{future}} = j | \underbrace{X_n}_{\text{present}} = i).$$

$X_{n+1} \perp (X_{n-1}, \ldots, X_0) \,|\, X_n$

Time-homogeneous Markov Chains

Definition

Given a Markov chain $X_0, X_1, X_2, \ldots$ It is called time-homogeneous Markov chain if for all $n \geq 0$,

$$P(X_{n+1} = j | X_n = i) = q_{i,j}.$$

where $q_{i,j}$ is a constant independent of n.

From now on, we focus on time-homogeneous Markov Chains, and we call it Markov chain in brevity.

Transition Matrix

Definition

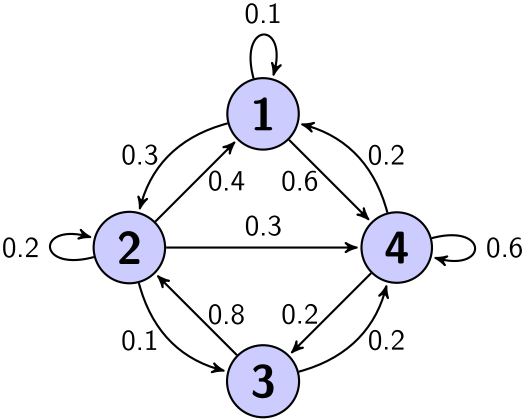
Let $X_0, X_1, X_2, \ldots$ be a Markov chain with state space $\{1, 2, \ldots, M\}$, and let $q_{i,j} = P(X_{n+1} = j | X_n = i)$ be the transition probability from state i to state j. The $M \times M$ matrix $Q = (q_{i,j})$ is called the *transition matrix* of the chain.

Graphical and Matrix Form of Markov Chain

$\forall i;\ \sum_{j=1}^{M} P(X_{n+1}=j \mid X_n=i) = 1$

$\Leftrightarrow \sum_{j=1}^{M} P(X_{n+1}=j, X_n=i) = P(X_n=i)$

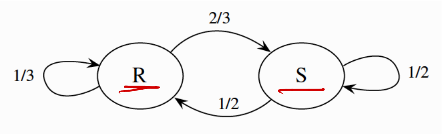
State-transition diagram



$$Q = \begin{bmatrix} 0.1 & 0.3 & 0 & 0.6 \\ 0.4 & 0.2 & 0.1 & 0.3 \\ 0 & 0.8 & 0 & 0.2 \\ 0.2 & 0 & 0.2 & 0.6 \end{bmatrix}$$

Row sum = 1 in Transition Matrix. Stochastic Matrix

Example: Rainy-Sunny Markov Chain

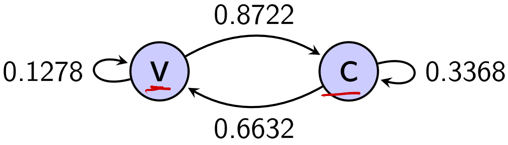


$$\begin{array}{cc} & \begin{array}{cc} R & S \end{array} \\ \begin{array}{c} R \\ S \end{array} & \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{array}$$

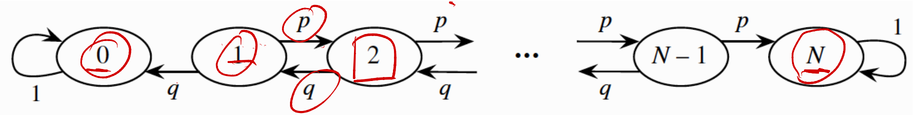
Example: The First Markov Chain in History

- Andrey Andreyevich Markov was interested in investigating the way the vowels and consonants alternate in Russian literature, e.g., “Eugene Onegin” by Pushkin
- He classified 20,000 consecutive characters: 8638 vowels & 11362 consonants

$$
\begin{array}{c} \\ \text{vowel} \\ \text{consonant} \end{array}
\begin{array}{c}
\begin{array}{cc} \quad\text{vowel} & \qquad\qquad \text{consonant} \end{array} \\
\begin{bmatrix} 1104/8638 & 7534/8638 \\ 7535/11362 & 3827/11362 \end{bmatrix}
\end{array}
= \begin{bmatrix} 0.1278 & 0.8722 \\ 0.6632 & 0.3368 \end{bmatrix}
$$



Gambler's Ruin As A Markov Chain



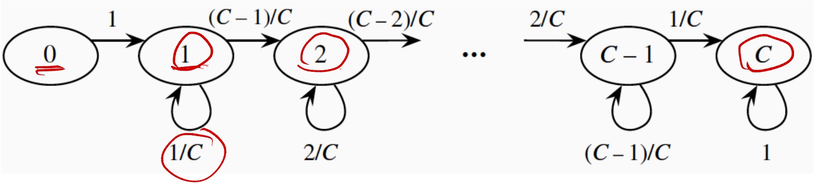
$X_0, X_1, \ldots$

$X_k \in \{0, 1, \ldots N\}$

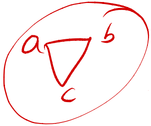
Coupon Collector As A Markov Chain

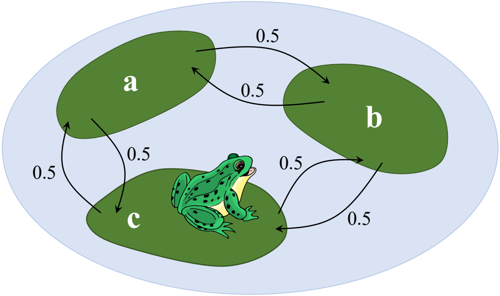
C : # of coupon type.

$X_k \in \{0, 1, \ldots, C\}$



Example: Random Walk on A Graph





$X_0, X_1, X_2, \ldots$

$$
\begin{array}{cc}
 & \begin{array}{ccc} a & b & c \end{array} \\
\begin{array}{c} a \\ b \\ c \end{array} & \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}
\end{array}
$$

Outline

n-step Transition Probability

$q_{a,c} = 0$

$q^{(2)}_{a,c} > 0$

Definition

Let $X_0, X_1, X_2, \ldots$ be a Markov chain with transition matrix Q. The n-step transition probability from i to j is the probability of being at j exactly n steps after being at i. We denote this by $q_{i,j}^{(n)}$:

$$q_{i,j}^{(n)} = P\left(X_n = j | X_0 = i\right).$$

Example: 2-step Transition Probability

$$q_{i,j}^{(2)} = P(X_2 = j \mid X_0 = i) = \sum_k P(X_2 = j, X_1 = k \mid X_0 = i)$$

$$= \sum_k \underbrace{P(X_2 = j \mid X_1 = k, X_0 = i)}_{\text{Markov property}} \cdot P(X_1 = k \mid X_0 = i)$$

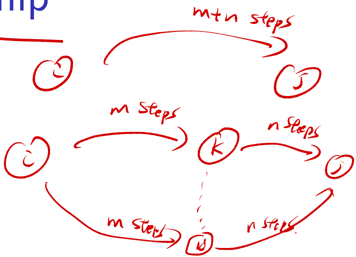
$$= \sum_k P(X_2 = j \mid X_1 = k) \cdot P(X_1 = k \mid X_0 = i)$$

$$q_{i,j}^{(2)} = P(X_2 = j | X_0 = i) = \sum_k q_{i,k} q_{k,j} = \text{(i,j) entry of } Q^2.$$

$$= \sum_k q_{k,j} \cdot q_{i,k} = \sum_k q_{i,k}\, q_{k,j}$$

Chapman-Kolmogorov Relationship

$$Q^{m+n} = Q^m \cdot Q^n$$



$$q_{i,j}^{(m+n)} = P(X_{m+n} = j | X_0 = i) = \sum_k q_{i,k}^{(m)} q_{k,j}^{(n)} = \text{(i,j) entry of } Q^{m+n}.$$

Proof

Distribution of X_n

$P(X_n=j) = \sum_i \underline{P(X_n=j \mid X_0=i)} \cdot P(X_0=i)$

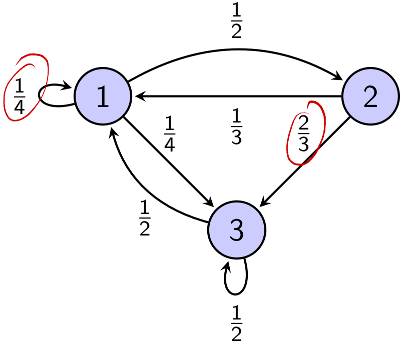
$= \underline{\sum_i q_{i,j}^{(n)} \cdot \alpha_i} \quad [\alpha Q^n]_{(j)}$

Let X_0, X_1,... be a Markov chain with transition matrix Q and initial distribution $\boldsymbol{\alpha}$, where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_M)$, $\alpha_i = P(X_0 = i), i = 1, \ldots, M$. For all n ≥ 0, the distribution of X_n is $\boldsymbol{\alpha} Q^n$. That is, the jth component of $\boldsymbol{\alpha} Q^n$ is $P(X_n = j)$, denoted as:

$$P(X_n = j) = (\alpha Q^n)_j, \ \textit{for all } j.$$

Example

Given a Markov chain $X_0, X_1, X_2, \ldots$ with state space $\mathcal{S} = \{1, 2, 3\}$.



$$Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \end{array}$$

- Find $P(X_3 = 1 | X_2 = 1)$ and $P(X_4 = 3 | X_3 = 2)$.

$q_{1,1} = \frac{1}{4}$. $q_{2,3} = \frac{2}{3}$.

Example

$P(X_0=1, X_1=2, X_2=3)$

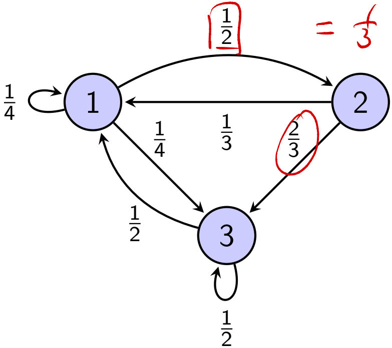
$P(X_2=3 \mid X_1=2)$

$= P(X_0=1) \cdot \underline{P(X_1=2 \mid X_0=1)} \cdot P(X_2=3 \mid X_1=2, X_0=1)$

$= P(X_0=1) \cdot q_{1,2} \cdot q_{2,3}$

Given a Markov chain $X_0, X_1, X_2, \ldots$ with state space $\mathcal{S} = \{1, 2, 3\}$.

$= \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{9}$



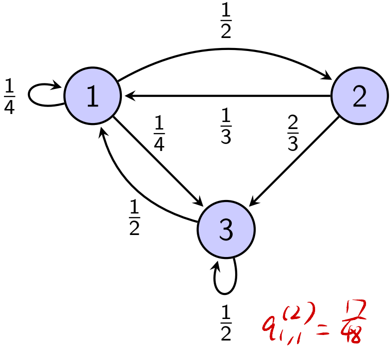
$$\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} & \multicolumn{3}{c}{\begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}} \end{array}$$

- If $\underline{P(X_0 = 1) = \frac{1}{3}}$, find $\underline{P(X_0 = 1, X_1 = 2, X_2 = 3)}$.

Example

$$Q^2 = \begin{bmatrix} \frac{17}{48} & \frac{1}{8} & \frac{25}{48} \\ \frac{5}{12} & \frac{1}{6} & \frac{5}{12} \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \end{bmatrix}$$

Given a Markov chain $X_0, X_1, X_2, \ldots$ with state space $\mathcal{S} = \{1, 2, 3\}$.



$$Q = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \end{array}$$

$q_{1,1}^{(2)} = \frac{17}{48}$ $\qquad$ $q_{1,2}^{(2)} = \frac{1}{8}$

- Find $P(X_2 = 1 | X_0 = 1)$, $P(X_2 = 2 | X_0 = 1)$, and $P(X_2 = 3 | X_0 = 1)$.

$q_{1,3}^{(2)} = \frac{25}{48}$

Example

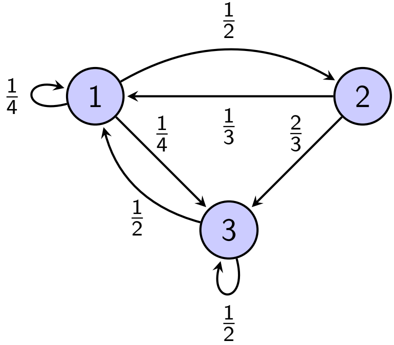
$$E[X_2 | X_0=1] = \sum_{j=1}^{3} j \cdot P(X_2=j \mid X_0=1)$$

$$= \sum_{j=1}^{3} j \cdot q_{1,j}^{(2)}$$

Given a Markov chain $X_0, X_1, X_2, \ldots$ with state space $\mathcal{S} = \{1, 2, 3\}$.

$$= 1 \cdot q_{1,1}^{(2)} + 2 \cdot q_{1,2}^{(2)} + 3 \cdot q_{1,3}^{(2)}$$

$$= \frac{104}{48} = \frac{13}{6}$$



$$\begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \end{array}$$

- Find $E(X_2 | X_0 = 1)$.

Markov Property

Let X_0, X_1,... be a Markov chain. Then, for all m < n, X_n

$$P(X_{n+1} = j | X_0 = i_0, \ldots, X_{n-m-1} = i_{n-m-1}, X_{n-m} = i)$$
$$= P(X_{n+1} = j | X_{n-m} = i)$$
$$= P(X_{m+1} = j | X_0 = i) = P_{ij}^{m+1},$$

for all i, j, i_0, ...,i_{n-m-1}, and n $\geq$ 0.

Joint Distribution

Let X_0, X_1,... be a Markov chain with transition matrix $\boldsymbol{P}$ and initial distribution $\boldsymbol{\alpha}$. For all $0 \leq n_1 < n_2 < \ldots < n_{k-1} < n_k$ and states $i_1, i_2, \ldots, i_{k-1}, i_k$,

$$
\begin{aligned}
P(X_{n_1} = i_1, X_{n_2} = i_2, \ldots, X_{n_{k-1}} = i_{k-1}, X_{n_k} = i_k) \\
= (\alpha P^{n_1})_{i_1} (P^{n_2 - n_1})_{i_1 i_2} \ldots (P^{n_k - n_{k-1}})_{i_{k-1} i_k}.
\end{aligned}
$$

Outline

Recurrent and Transient States

N: # of transitions

$E(N) < \infty$

Positive recurrent

Null recurrent. $E(N) = \infty$

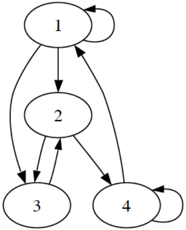
Definition

State i of a Markov chain is **recurrent** if starting from i, the probability is 1 that the chain will eventually return to i. Otherwise, the state is **transient**, which means that if the chain starts from i, there is a positive probability of never returning to i.

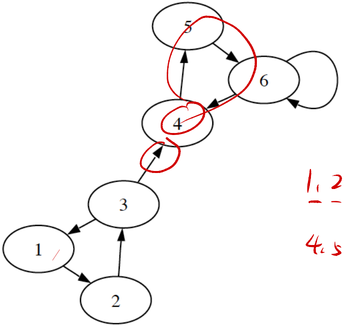
Example

1, 2, 3, 4,

recurrent.



irreducible.



1, 2, 3 transient.

4, 5, 6, recurrent.

reducible.

Irreducible and Reducible Chain

Definition

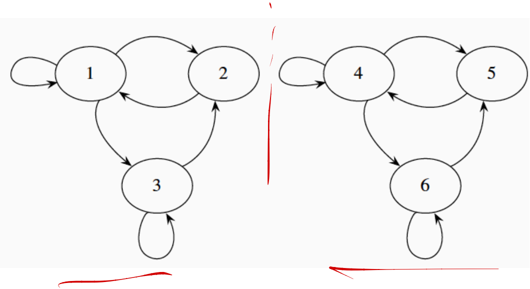
A Markov chain with transition matrix Q is *irreducible* if for any two states i and j, it is possible to go from i to j in a finite number of steps (with positive probability). That is, for any states i,j there is some positive integer n such that the (i,j) entry of Q^n is positive. A Markov chain that is not irreducible is called *reducible*.

Irreducible Implies All States Recurrent

Theorem

In an irreducible Markov chain with a finite state space, all states are recurrent.

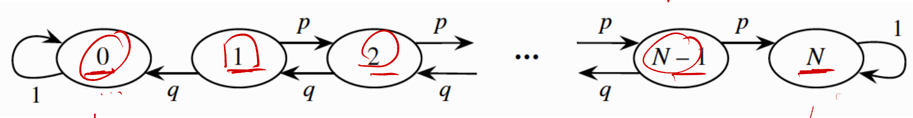
A Reducible Markov Chain with Recurrent States



Gambler's Ruin As A Markov Chain

recurrent states : $\{0, N\}$

transient states : $\{1, 2, \ldots, N-1\}$

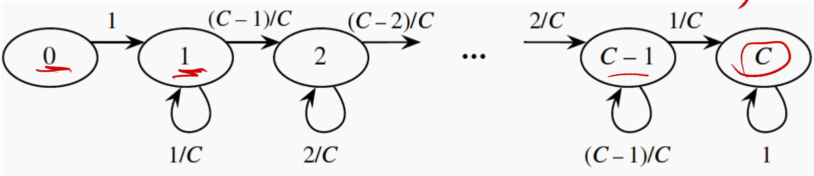


absorbing states.

Coupon Collector As A Markov Chain

recurrent state : $\{C\}$

transient states : $\{0, 1, 2, \ldots C-1\}$



Period

$q_{i,i} > 0$

$d(i) = gcd(1, n_1, n_2)$
$= 1$

Definition

For a Markov chain with transition matrix Q, the *period* of state i, denoted $d(i)$, is the greatest common divisor of the set of possible return times to i. That is,

$$d(i) = gcd\{n > 0 : Q^n_{i,i} > 0\}.$$

If $d(i) = 1$, state i is said to be *aperiodic*. If the set of return times is empty, set $d(i) = +\infty$.

Periodic, Aperiodic Markov Chain

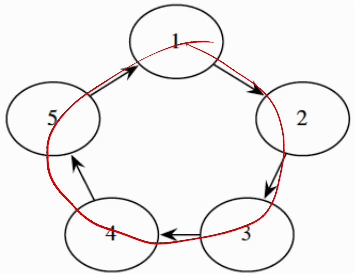
Definition

A Markov chain is called *periodic* if it is irreducible and all states have period greater than 1.

A Markov chain is called *aperiodic* if it is irreducible and all states have period equal to 1.

Example: Periodic Chain

$d(1)=d(2)=d(3)$
$=d(4)$
$=d(5)=5$



Example

$n_1 \rightarrow$ state 1

n_2

2

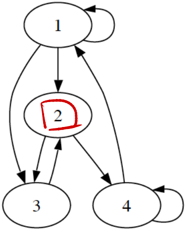
n_1+n_2

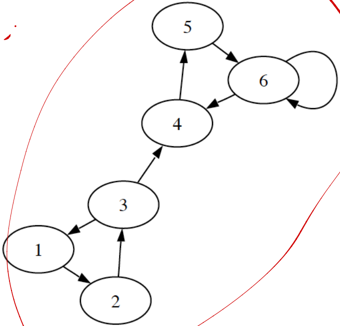
n_1+n_2+1

irreducible;

$d(1)=1$, $d(4)=1$

$d(2)=1$, $d(3)=1$;





aperiodic M.C.

Outline

$\pi_n^{(i)} = P(X_n = i)$; $\pi_{n+1}^{(j)} = P(X_{n+1} = j)$.

$$P(X_{n+1}=j) \overset{LOTP}{=} \sum_i P(X_{n+1}=j \mid X_n = i)\, P(X_n = i)$$

$$= \sum_i q_{i,j}\, P(X_n = i)$$

$$\pi_{n+1}^{(j)} = \sum_i \pi_n^{(i)} \cdot q_{i,j}$$

$$\pi_{n+1} = \pi_n \cdot Q$$

Definition

$\{X_n\}_{n\geq 0}$; n^{th} time state distribution π_n

$\pi_{n+1} = \pi_n Q$

if $\pi_n = s$; $sQ = s$

Definition

$\Rightarrow \pi_{n+1} = sQ = s$ $\Rightarrow \pi_{n+2} = sQ = s$

$\pi_k = s$, $k \geq n$

A row vector $\mathbf{s} = (s_1, \dots, s_M)$ such that $s_i \geq 0$ and $\sum_i s_i = 1$ is a *stationary distribution* for a Markov chain with transition matrix Q if

$$\boxed{\sum_i s_i q_{i,j} = s_j.}$$

for all j, or equivalently,

$$\mathbf{s}Q = \mathbf{s}.$$

Example: Double Stochastic Matrix

$S_j = \sum_i S_i \cdot q_{i,j}$ | $\cancel{\frac{1}{M}} = \sum_i \cancel{\frac{1}{M}} \cdot q_{i,j}$

$\Leftrightarrow 1 = \sum_i q_{i,j}$ ✓

$$
\begin{array}{cc}
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}
\end{array}
$$

Theorem

If each column of the transition matrix Q sums to 1, then the uniform distribution over all states, $(1/M, 1/M, ..., 1/M)$, is a stationary distribution. (A nonnegative matrix such that the row sums and the column sums are all equal to 1 is called a doubly stochastic matrix.)

Example: Two-State Markov Chain

① $0<\alpha<1$, $0<\beta<1$.

irreducible.

aperiodic.

② $Q = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-\alpha & \alpha \\ 1 & \beta & 1-\beta \end{array}$.

③ $\pi_n \xrightarrow[n\to\infty]{\infty} s$.

$SQ = S$. ; $S = (s_0, s_1)$

$(s_0, s_1)\begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} = (s_0, s_1)$

$\Rightarrow \begin{cases} s_0(1-\alpha) + s_1\beta = s_0 \\ s_0\alpha + s_1(1-\beta) = s_1 \\ s_0 + s_1 = 1 \end{cases}$

$\Rightarrow s_0 = \frac{\beta}{\alpha+\beta}$

$s_1 = \frac{\alpha}{\alpha+\beta}$

Stationary distribution $S = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right)$

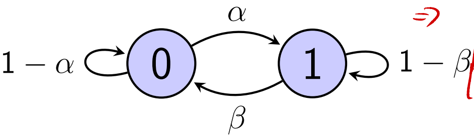
$\pi_{n+1} = \pi_n \cdot Q$

$\Rightarrow \pi_n = \pi_0 Q^n$

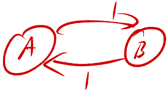
$Q = LDL^{-1}$

$Q^n = LD^nL^{-1}$

$\Rightarrow \pi_n$.



Theorem on Stationary Distribution



Period = 2

$X_0 = A,$

A, B, A, B, A, B

s_n (1, 0) (0, 1), (1, 0) (0, 1)

Theorem

Given a Markov chain with finite state space.

- *If such Markov chain is irreducible, then it has a unique stationary distribution. In this distribution, every state has positive probability.*
- *If such Markov chain is irreducible and aperiodic with stationary distribution* $\mathbf{s}$ *and transition matrix* Q, *then* $P(X_n = i)$ *converges to* s_i *as* $n \to \infty$. *In terms of the transition matrix,* Q^n *converges to a matrix in which each row is* $\mathbf{s}$.

$s_n \to s$

Fundamental Limit Theorem for Ergodic Markov Chains

Finite State space + irreducible + aperiodic

ergodic

A Markov chain is called *ergodic* if it is irreducible, aperiodic, and all states have finite expected return times (positive recurrent).

Theorem

Let $X_0, X_1, \ldots$ be an ergodic Markov chain. There exists a unique, positive, stationary distribution $\boldsymbol{\pi}$, which is the limiting distribution of the chain. That is,

$$\pi_j = \lim_{n\to\infty} P_{ij}^{n}, \text{ for all } i,\ j.$$

$\lim_{n\to\infty} z_n = S$

2

Outline

Reversibility

① $\tau_n = s$; $n \geq k$, $\tau_k = s$ Markov chain → stationary distribution

② time t ⓘ ⟶ ⓙ time $t+1$. $q_{i,j}$

Reverse the time $t+1$ ⓘ ⟶ ⓙ time t $\tilde{q}_{i,j}$

Definition

Let $Q = (q_{i,j})$ be the transition matrix of a Markov chain. Suppose there is $\mathbf{s} = (s_1, \ldots, s_M)$ with $s_i \geq 0$, $\sum_i s_i = 1$, such that

$$s_i q_{i,j} = s_j q_{j,i}$$

$\tilde{q}_{i,j} = P(X_t = j \mid X_{t+1} = i)$

$= P(X_{t+1} = i \mid X_t = j) \dfrac{P(X_t = j)}{P(X_{t+1} = i)}$

for all states i and j. This equation is called the *reversibility* or *detailed balance* condition, and we say that the chain is *reversible* with respect to $\mathbf{s}$ if it holds.

$= q_{j,i} \cdot \dfrac{\pi_t^{(j)}}{\pi_{t+1}^{(i)}} = q_{j,i} \dfrac{s_j}{s_i}$

$q_{i,j} = \tilde{q}_{i,j} \quad \forall i,j$

$= q_{j,i} \dfrac{s_j}{s_i} \Rightarrow s_i q_{i,j} = s_j q_{j,i}$

Check the Detailed Balance Equation

Known π_i. Find Q.

$\pi_i q_{i,j} = \pi_j q_{j,i}$

Theorem

If for an irreducible Markov chain with transition matrix $Q = (q_{i,j})$, there exists a probability solution π to the detailed balance equations

$$\pi_i q_{i,j} = \pi_j q_{j,i}$$

for all pairs of states i, j, then this Markov chain is positive recurrent, time-reversible and the solution π is the unique stationary distribution.

Example: Symmetric Transition Matrix

$\forall i,j$

$\pi_i q_{i,j} = \pi_j \cdot q_{j,i}$

$q_{i,j} = q_{j,i} \neq 0$

$\Rightarrow \pi_i = \pi_j$

$\forall i,j$

Ham(i,j) =1

$$
\begin{array}{cc}
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}
\end{array}
$$

Theorem

If the transition matrix Q for an irreducible Markov chain is symmetric, then the uniform distribution over all states, $(1/M, 1/M, ..., 1/M)$, is the unique stationary distribution.

Example: Random Walk on Undirected Graph

1°. irreducible, aperiodic

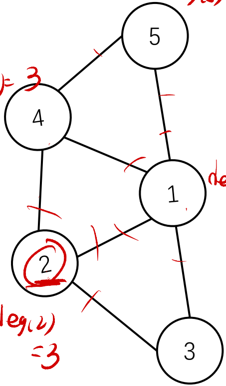
2°. DBE

$\pi_i \cdot q_{i,j} = \pi_j \cdot q_{j,i}$.

$\Rightarrow \pi_i \frac{1}{deg(i)} = \pi_j \cdot \frac{1}{deg(j)}$

$\sum \pi_i = 1$

$\Rightarrow \pi_i = \frac{deg(i)}{\sum_j deg(j)}$



$(i) \rightarrow (j)$

$q_{i,j} = \frac{1}{deg(i)}$

$\frac{\pi_i}{deg(i)} = c, \forall i$.

① ② ③ ④ ⑤

$\vec{deg} = (4, 3, 2, 3, 2)$

$\sum_i deg(i) = 14$.

$\Rightarrow$

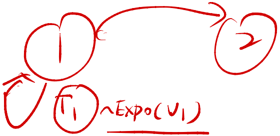
$\pi = \left(\frac{4}{14}, \frac{3}{14}, \frac{2}{14}, \frac{3}{14}, \frac{2}{14}\right)$

$= \left(\frac{2}{7}, \frac{3}{14}, \frac{1}{7}, \frac{3}{14}, \frac{1}{7}\right)$

Example: Random Walk on Undirected Graph

Outline

Continuous Time Markov Chain (CTMC)



DTMC.

- Modeling: Transitions between states & durations of time in each state
- Markov property holds: given the present, past and future are independent.
- Markov property: a form of memorylessness leading to exponential distribution.

DTMC : Q.

CTMC : Q + During time of each state.

Markov Property

Definition

A continuous-time stochastic process $(X_t)_{t\geq 0}$ with discrete state space $\mathcal{S}$ is a *continuous-time Markov chain* if

history info.

$$P(X_{t+s} = j | X_s = i, X_u = x_u, 0 \leq u < s) = P(X_{t+s} = j | X_s = i),$$

for all $s, t \geq 0$, $i, j, x_u \in \mathcal{S}$, and $0 \leq u < s$.

The process is said to be *time-homogeneous* if this probability does not depend on s. That is,

$$P(X_{t+s} = j | X_s = i) = P(X_t = j | X_0 = i), \text{ for } s \geq 0.$$

Holding Times are Exponentially Distributed

① $X_0 = i$, $\forall s \geq 0$,

$$(*1)\ P(T_i > s) = P(X_u = i, 0 < u \leq s \mid X_0 = i)$$

$$= \frac{P(X_u = i, 0 < u \leq s, X_0 = i)}{P(X_0 = i)}$$

Let T_i be the holding time at state i, that is, the length of time that a continuous-time Markov chain started in i stays in i before transitioning to a new state. Then, T_i has an exponential distribution.

$$= \frac{P(X_u = i, 0 \leq u \leq s)}{P(X_0 = i)}$$

$t \geq 0$;

$$(*2)\ P(T_i > s+t) = \frac{P(X_u = i, 0 \leq u \leq s+t)}{P(X_0 = i)} ;$$

object: $\boxed{P(T_i > s+t \mid T_i > s) = P(T_i > t)}$

Proof

② $P(T_i > s+t \mid T_i > s) = \frac{P(T_i > s+t,\ T_i > s)}{P(T_i > s)} = \frac{P(T_i > s+t)}{P(T_i > s)}$

$= \frac{P(X_u = i,\ 0 \le u \le s+t)}{P(X_u = i,\ 0 \le u \le s)} = \frac{P(X_u = i,\ 0 \le u \le s;\ X_u = i,\ s < u \le s+t)}{P(X_u = i,\ 0 \le u \le s)}$

$= P(X_u = i,\ s < u \le s+t \mid X_u = i,\ 0 \le u \le s)$

Markov's property

$= P(X_u = i,\ s < u \le s+t \mid X_s = i)$

Time-homogeneous

$= P(X_u = i,\ 0 < u \le t \mid X_0 = i)$

$= P(T_i > t)$

$\exists\, v_i > 0 \Rightarrow T_i \sim \mathrm{Expo}(v_i)$

③ T_i is memoryless. $T_i \geq 0$ continuous

Embedded Chains

For each state i, holding time $T_i \sim \exp(v_i)$ rate.

$E(T_i) = \frac{1}{v_i}$

$\begin{cases} v_i = 0 & \text{; absorbing state} \\ v_i = \infty & \text{; explosive state.} \end{cases}$

$0 < v_i < \infty$ CTMC.

$$i \xrightarrow[T_i]{P^{e}_{i,j}} j \xrightarrow[T_j]{P^{e}_{j,k}} k$$

$P^e_{i,j}$ = transition prob from state i to state j.

$P^e_{i,i} = 0$

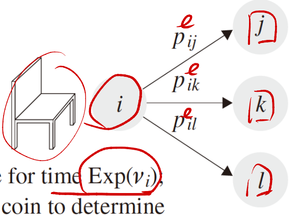
$P^e = \{P^e_{i,j}\}$: transition matrix of embedded chain. DTMC

$$\text{CTMC} = \underbrace{\{v_i\}}_{\text{holding time}} + \underbrace{P^e}_{\text{embedded chain.}}$$

CTMC: Perspective 1

Gilespie Algorithm

τ-leaping



Sit here for time $\mathrm{Exp}(\nu_i)$, then flip coin to determine where to go next.

$$\sum_{i'} p^e_{i,i'} = 1$$

CTMC: Perspective 2

1°. $T_i = \min(T_{i,j}, T_{i,k}, T_{i,l})$

$\sim \exp(v_i \cdot p^e_{i,j} + v_i \cdot p^e_{i,k} + v_i \cdot p^e_{i,l})$

$= v_i (\underbrace{p^e_{i,j} + p^e_{i,k} + p^e_{i,l}}_{1})$

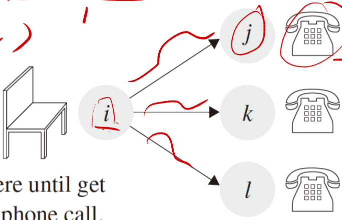
$= v_i$)

X, Y, Z r.v.s. $X \sim Exp(\lambda_1)$, $Y \sim Exp(\lambda_2)$, $Z \sim Exp(\lambda_3)$. independent

$\Rightarrow \min(X, Y, Z) \sim \exp(\lambda_1 + \lambda_2 + \lambda_3)$

$P(X = \min(X,Y,Z)) = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3}$.

Count-down timer



Sit here until get first phone call.

$\text{Exp}(\nu_i \cdot p^e_{ij})$ time until call $\quad T_{i,j} \sim Exp(v_i \cdot p^e_{i,j})$

$\text{Exp}(\nu_i \cdot p^e_{ik})$ time until call $\quad T_{i,k}$

$\text{Exp}(\nu_i \cdot p^e_{il})$ time until call $\quad T_{i,l}$

2°. Transition Prob from i to j.

$$P(T_{i,j} = \min(T_{i,j}, T_{i,k}, T_{i,l})) = \frac{v_i \cdot p^e_{i,j}}{v_i \cdot p^e_{i,j} + v_i \cdot p^e_{i,k} + v_i \cdot p^e_{i,l}} = \frac{\cancel{v_i} p^e_{i,j}}{\cancel{v_i}} = p^e_{i,j}$$

Generator Matrix

1°. $T_{i,j} \sim \exp\left(\underline{v_i p^e_{i,j}}\right)$

$T_{i,j} \sim \exp(q_{i,j})$

$q_{i,j} = \underline{v_i} \cdot \underline{p^e_{i,j}}$ transition rate

v_i : rate at which the process leave state i.

$p^e_{i,j}$: transition prob. from state i to state j.

$$\Rightarrow \sum_{j \neq i} q_{i,j} = \sum_{j \neq i} (v_i p^e_{i,j}) = v_i \underline{\sum_{j \neq i} p^e_{i,j}} = v_i \quad (\underline{p^e_{i,i} = 0})$$

$$\Rightarrow \quad v_i = \sum_{j \neq i} q_{i,j}$$

$$p^e_{i,j} = \frac{q_{i,j}}{v_i} \ (i \neq j) = \frac{q_{i,j}}{\sum_{j \neq i} q_{i,j}}, \ \underline{i \neq j}$$

2°. Define a matrix $\underline{Q}$: $\begin{cases} Q_{i,j} = q_{i,j} & (i \neq j) \\ Q_{i,i} = -v_i \end{cases}$

$$\Rightarrow \sum_j Q_{i,j} = \sum_{j \neq i} Q_{i,j} + Q_{i,i} = \sum_{j \neq i} q_{i,j} - v_i = 0$$

Instantaneous Rates, Holding Times, Transition Probabilities

$P(X_0=j \mid X_0=i)$

① $P_{i,j}(t) = P(X_{t+s}=j \mid X_s=i) = P(X_t=j \mid X_0=i)$

$\lim_{t\to 0} \frac{1-P_{i,i}(t)}{t} = v_i$; $\lim_{t\to 0} \frac{P_{i,j}(t)}{t} = q_{i,j}$, $i\neq j$

② $P(t) = [P_{i,j}(t)]$; $P(t+s) = P(t)\cdot P(s)$; $Q = \lim_{\delta\to 0} \frac{P(\delta)-I}{\delta}$

For a continuous-time Markov chain, let q_{ij}, v_i, and p^e_{ij} be defined as above. For $i \neq j$,

$$q_{ij} = v_i p^e_{ij}.$$

③ $\lim_{\delta\to 0} \frac{P(t+\delta)-P(t)}{\delta} = \lim_{\delta\to 0} \frac{P(t)P(\delta)-P(t)}{\delta}$

$= \lim_{\delta\to 0} \frac{P(t)[P(\delta)-I]}{\delta} = P(t) \cdot \lim_{\delta\to 0} \frac{P(\delta)-I}{\delta}$

$P'(t) = P(t)Q$. Forward Equation

$P(0)=I$;

$\lim_{\delta\to 0} \frac{[P(\delta)-I]\cdot P(t)}{\delta}$

$P'(t) = QP(t)$ Backward Equation

$P(0)=I$.

Finite state space ;

Birth-Death process.

Example: Birth-Death System

$$\begin{tikzcd} 0 \arrow[r, bend left, "\lambda"] & 1 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda"] & 2 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda"] & 3 \arrow[l, bend left, "\mu"] \end{tikzcd}$$

CTMC

1°. Suppose we are in state i, $i \geq 1$

$i \to i+1$ birth.

$i \to i-1$ death.

X_B : the time to the next birth.

X_D : - - - - death.

$X_B \sim \exp(\lambda)$

independent

$X_D \sim \exp(\mu)$.

$\Rightarrow$ holding time at state i : $T_i = \min(X_B, X_D) \sim \exp(\lambda+\mu)$

$\Rightarrow v_i = \lambda+\mu$.

2° $P^e_{i,i+1} = P(X_B = \min(X_B, X_D)) = \frac{\lambda}{\lambda+\mu} = \frac{\lambda}{v_i}$. $\Rightarrow q_{i,i+1} = v_i P_{i,i+1}$

$i \geq 0$ $= \lambda$.

$i \geq 1$, $P^e_{i,i-1} = P(X_D = \min(X_B, X_D)) = \frac{\mu}{\lambda+\mu} = \frac{\mu}{v_i} \Rightarrow q_{i,i-1} = v_i P^e_{i,i-1}$

$= \mu$.

$P^e_{0,1} = 1$, $v_0 = \lambda$ $\Rightarrow q_{0,1} = \lambda$; $q_{i,i+1} = \lambda$, $i \geq 0$

Example: Birth-Death System

generator matrix

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & \cdots & & & \\ \mu & -(\lambda+\mu) & \lambda & 0 & \cdots & & \\ 0 & \ddots & \ddots & \ddots & \lambda & & \\ & & \ddots & \ddots & \ddots & \lambda & \\ & & \mu & \ddots & \ddots & & \lambda \\ & & & & & & -(\lambda+\mu) \\ & & & & \mu & & \ddots \\ & & \cdots & & & & \end{bmatrix}$$

Stationary Distribution & Generator Matrix

π_t:

$\pi_t^{(j)} = P(X_t = j) \overset{LOTP}{=} \sum_i P(X_t = j \mid X_0 = i)\, P(X_0 = i) = \sum_i P_{ij}(t) \cdot \pi_0^{(i)}$.

$\pi_t = \pi_0 \cdot P(t)$ $\quad (P(t) = [P_{ij}(t)])$ $\quad \Rightarrow \pi'(t) = \pi_0 P'(t)$

$= \pi_0 \cdot P(t) \cdot Q$

$= \pi_t \cdot Q$

$\dfrac{d\pi(t)}{dt} = \pi_t \cdot Q$

π_0.

A probability distribution $\boldsymbol{\pi}$ is a stationary distribution of a continuous-time Markov chain with generator $\boldsymbol{Q}$ if and only if

$\pi P(t) = \pi, \ \forall t \geq t_0$

$\pi P'(t) = 0 \Rightarrow \pi Q P(t) = 0, \ \forall t \geq t_0$

$\Rightarrow \pi Q = 0$

$$\boldsymbol{\pi} \boldsymbol{Q} = \boldsymbol{0}.$$

That is,

$$\sum_i \pi_i Q_{ij} = 0, \ \text{for all } j.$$

$\pi_t = \pi, \ \forall t \geq t_0$

$\pi'_t = 0 = \pi Q$.

Time Reversibility

MCMC

Reverse of time process

Diffusion model

X_t : $t \in [0, T)$

$\bar{X}_t$: $X(T-t)$

Definition

A continuous-time Markov chain with generator $\boldsymbol{Q}$ and unique stationary distribution $\boldsymbol{\pi}$ is said to be *time reversible* if

$$\pi_i q_{ij} = \pi_j q_{ji}, \text{ for all } i, j.$$

DTMC: $\pi_i P_{ij} = \pi_j \cdot P_{ji}$

Ergodicity of CTMC

Design CTMC:

Target distribution $\pi \xrightarrow{\text{DBE}} \underset{\text{irreducible}}{q_{i,j}} \longrightarrow \nu_i,\ P^e_{i,j} \longrightarrow$ CTMC

- An irreducible CTMC with a finite state space is positive recurrent.
- For an irreducible CTMC with transition rate matrix $Q = \{q_{ij}\}$, if there exists a probability solution π to the detailed balance equation

  DBE

$$\pi_i q_{i,j} = \pi_j q_{j,i}, \ \forall i, j \in S$$

  then this Markov chain is positive recurrent, time-reversible, and the solution π is the unique stationary distribution.

Birth-and-Death Process

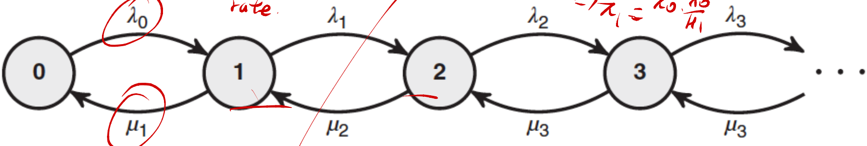
① irreducible

② DBE. $\pi_i \cdot \lambda_i = \pi_{i+1} \cdot \mu_{i+1}$.

$\pi_0 \cdot \lambda_0 = \pi_1 \cdot \mu_1$.

$\Rightarrow \pi_1 = \pi_0 \cdot \frac{\lambda_0}{\mu_1}$

$\pi_0 \cdot \lambda_0$ prob. flow rate.



$\pi_1 \cdot \mu_1$.

$\pi_1 \cdot \lambda_1 = \pi_2 \cdot \mu_2$

$\Rightarrow \pi_2 = \pi_1 \frac{\lambda_1}{\mu_2} = \pi_0 \cdot \frac{\lambda_0}{\mu_1} \cdot \frac{\lambda_1}{\mu_2}$

flow Balance

$\pi_k = \ldots$

③ $\sum_{i \geq 0} \pi_i = 1$

Stationary Distribution for Birth-and-Death Process

For a birth-and-death process with birth rates λ_i and death rates μ_i, for $i = 1, 2, \ldots,$ assume that

$$\sum_{k=0}^{\infty} \prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i} < \infty$$

Then, the unique stationary distribution $\boldsymbol{\pi}$ is

$$\pi_k = \pi_0 \prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i} \text{ for } k = 1, 2, \ldots,$$

where

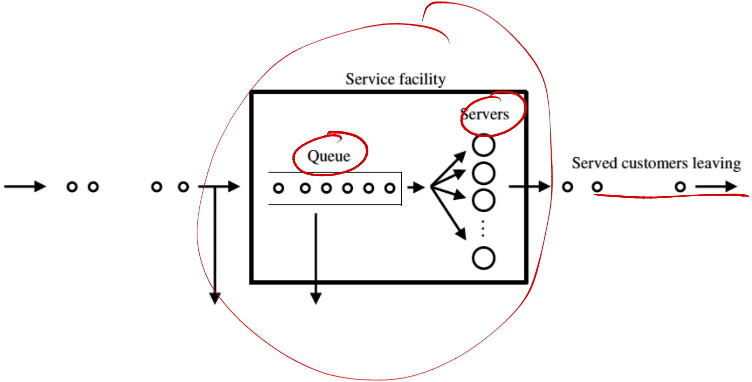
$$\pi_0 = \left(\sum_{k=0}^{\infty} \prod_{i=1}^{k} \frac{\lambda_{i-1}}{\mu_i} \right)^{-1}.$$

Common Birth-and-Death Processes

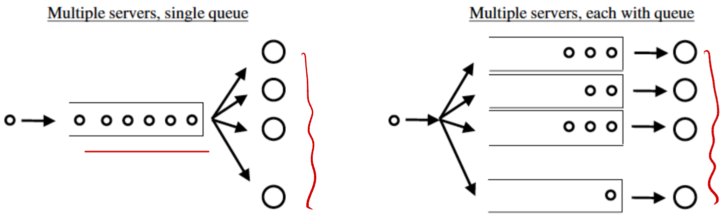
Type	Birth Rate	Death Rate
Pure birth	λ_i	$\mu_i = 0$
Poisson process	$\lambda_i = \lambda$	$\mu_i = 0$
Pure death	$\lambda_i = 0$	μ_i
Linear process	$\lambda_i = i\lambda, i > 0$	$\mu_i = i\mu$
Yule process	$\lambda_i = \lambda i, i, \lambda > 0$	$\mu_i = 0$
Linear with immigration	$\lambda_i = i\lambda + \alpha, i, \alpha > 0$	$\mu_i = i\mu$

Outline

A Typical Queueing System



Multi-Server Queueing System



Kendall Notation: A/B/X/Y/Z

Table 1.1 Queueing notation $A/B/X/Y/Z$

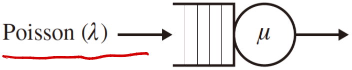
Characteristic	Symbol	Explanation
Interarrival-time distribution (A)	M memoryless	Exponential
	D	Deterministic
	E_k	Erlang type k $(k = 1, 2, \ldots)$
Service-time distribution (B)	H_k	Mixture of k exponentials
	PH	Phase type
	G	General
Parallel servers (X)	$1, 2, \ldots, \infty$	
System capacity (Y)	$1, 2, \ldots, \infty$	
Queue discipline (Z)	FCFS	First come, first served
	LCFS	Last come, first served
	RSS	Random selection for service
	PR	Priority
	GD	General discipline

Little's Law

In a queueing system, let L denote the long-term average number of customers in the system, λ the rate of arrivals, and W the long-term average time that a customer is in the system. Then,

$$L = \lambda W.$$

M/M/1 Queue



M/M/1 Queue

CTMC

$\left[\frac{\lambda}{\mu} < 1 \right]$

Birth-Death System

$q_{i,i+1} = \lambda, \quad i \geq 0$

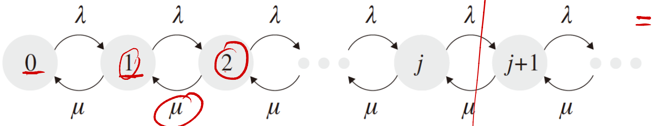
$q_{i,i-1} = \mu, \quad i \geq 1$

$\pi_i \cdot \lambda = \pi_{i+1} \cdot \mu$

$\Rightarrow \pi_i = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^i$

$L = E[N] = \sum_{i=0}^{\infty} i \cdot \pi_i = \frac{\lambda}{\mu - \lambda};$

$L_Q = \sum_{i=1}^{\infty} (i-1) \cdot \pi_i = \frac{\lambda^2}{\mu(\mu-\lambda)};$



$W = E[T] = \sum_{i=0}^{\infty} \frac{i+1}{\mu} \cdot \pi_i = \frac{1}{\mu - \lambda}$

$i = 0; \quad \frac{1}{\mu}$

$i \geq 1; \quad \frac{i}{\mu} + \frac{1}{\mu} = \frac{i+1}{\mu}$

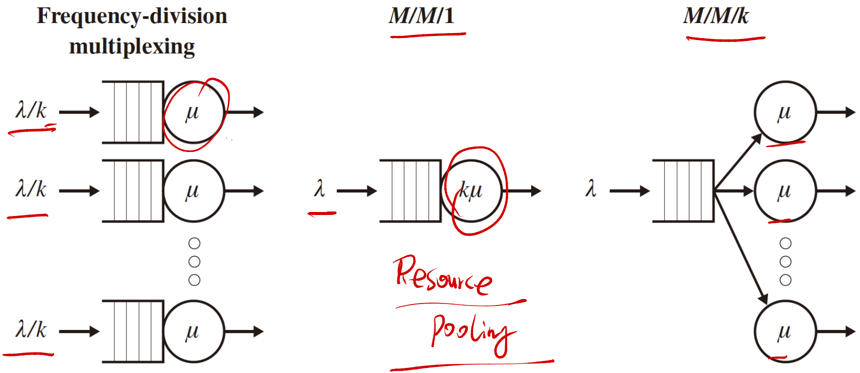
$W_Q = \sum_{i=0}^{\infty} \frac{i}{\mu} \cdot \pi_i = \frac{\lambda}{\mu(\mu-\lambda)}$

$L_Q = \lambda \cdot W_Q$

$L = \lambda W$

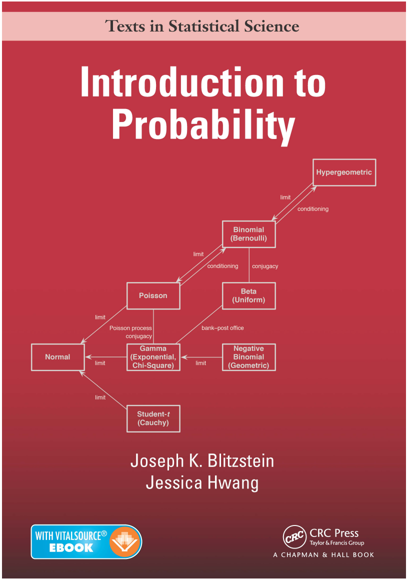
Application: Three Server Organizations

- Three data center systems with the same arriving rate λ and the same total service rate $k\mu$.
- different server organizations: FDM, M/M/1 and M/M/k



Outline

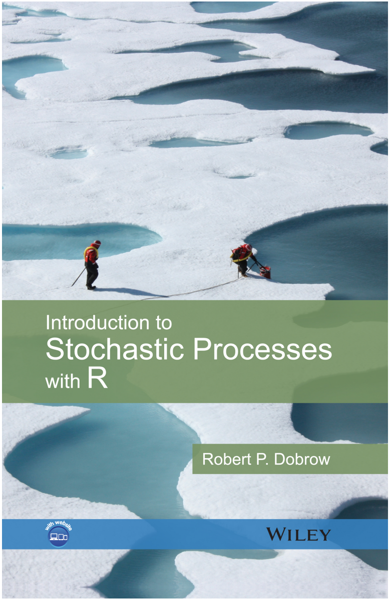
BH



BH

- Introduction to Probability
- Chapman & Hall/CRC, 2014.
- Chapman & Hall/CRC, 2019.
- Chapter 11

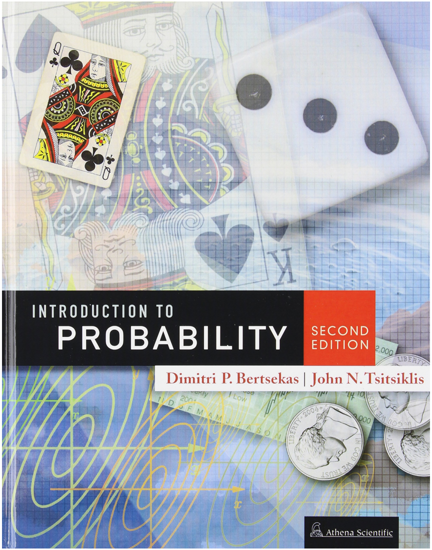
SRR



SPR

- Introduction to Stochastic Processes with R
- John Wiley & Son, 2016.
- Chapters 2 & 3 & 7

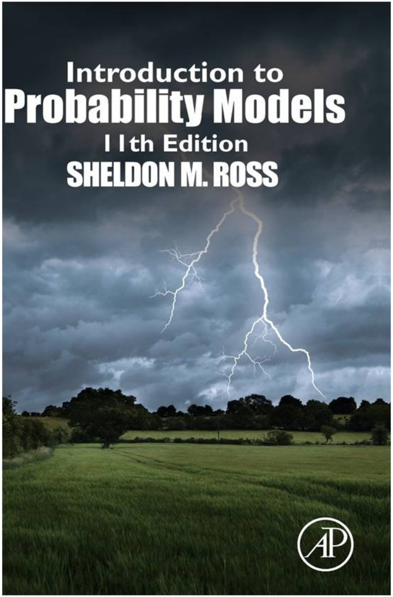
BT



BT

- Introduction to Probability (2nd Edition)
- Athena Scientific, 2008.
- Chapter 7

SMR



SMR

- Introduction to Probability Models
- Academic Press, 11 edition, 2014.
- Chapters 4 & 6 & 8