# Lecture 5: Markov Chains

Ziyu Shao

School of Information Science and Technology
ShanghaiTech University

April 06 & 08, 2020

# Outline

# Outline

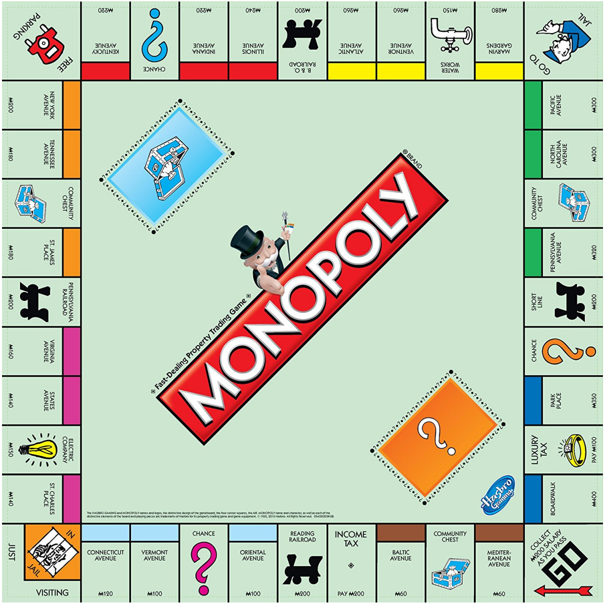
# Definition

State space $\begin{cases} \text{Discrete} \\ \text{Continuous.} \end{cases}$

Index $\begin{cases} \text{Discrete} \\ \text{Continuous} \end{cases}$

- 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




$I = \{0, 1, 2, \dots\}$

$X_0, X_1, X_2, \dots$

# Example: Discrete Time & Continuous State Space

PM 2.5 level record. [every hour]

$\downarrow$

Continuous.

$X_0, X_1, X_2, \ldots$

# Example: Continuous Time & Discrete State Space

$X_t$ : # of emails received up to time $t$.

$X_t \in \{0, 1, 2, \dots\}$.

# Example: Continuous Time & Continuous State Space

Brownian motion.

# 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
- 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 Process

Three basic components of Markov process

- 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 Process

- *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 $\mathcal{T}$
- *Continuous Markov Process*: Continuous $\mathcal{S}$ & Continuous $\mathcal{T}$

# Outline

# Markov Chain

## 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 $n \geq 0$,

$$P\left(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\right) = P\left(X_{n+1} = j | X_n = i\right).$$

$$P(X_4 = j \mid X_0, X_1, X_2, X_3)$$

$$= P(X_4 = j \mid X_3)$$

# Time-homogeneous Markov Chains

$$P(X_{n+1}=j \mid X_n=i) = P(X_1=j \mid X_0=i)$$

$$= \underbrace{P_{ij}}$$

One-step transition matrix.

# Transition Matrix

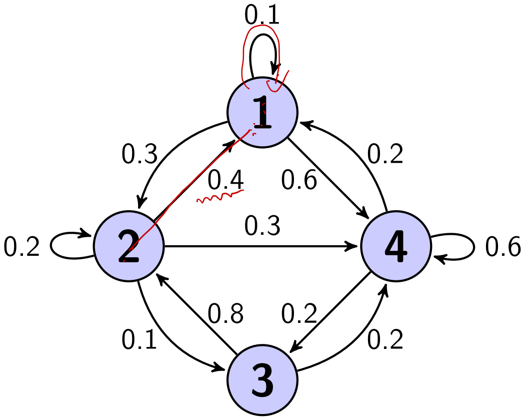
## Definition

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

Q.

# Graphical and Matrix Form of Markov Chain

$\forall i, \ \sum_j q_{ij} = 1$



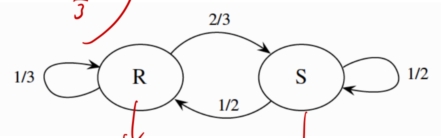

$$Q = \begin{array}{cc} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 0.1 & 0.3 & 0 & 0.6 \\ 0.4 & 0.2 & 0.1 & 0.3 \\ 0 & 0.8 & 0 & 0.2 \\ 0.2 & 0 & 0.2 & 0.6 \end{bmatrix} \end{array}$$

Row sum $=1$ in Transition Matrix.

# Example: Rainy-sunny Markov Chain

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

Tomorrow's weather depends only on today's weather.




Rainy

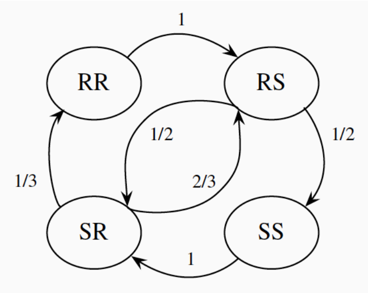
Sunny

# Example: Rainy-sunny Markov Chain

$Y_n = (X_{n-1}, X_n)$

new state.

$\{(RR), (RS), (SR), (SS)\}$



# Stochastic Matrix

transition matrix is a stochastic matrix.

## Definition

A stochastic matrix is a square matrix **P**, which satisfies

1. $P_{ij} \geq 0$ for all i, j.
2. For each row i

$$\sum_j P_{ij} = 1.$$

# Algorithm for Simulating a Markov Chain

Input:

- initial distribution $\alpha$.
- transition matrix $\mathbf{P}$.
- number of steps $n$.

Output:

- $X_0, X_1, \ldots, X_n$.

Algorithm:

- Generate $X_0$ according to $\alpha$ Monte Carlo.
- FOR $i$=1,. . . , $n$
  - Assume that $X_{i-1} = j$
  - Set $\mathbf{p} = j$th row of $\mathbf{P}$
  - Generate $X_i$ according to $\mathbf{p}$
- END FOR

# Outline

# The First Example of Markov Chain in History

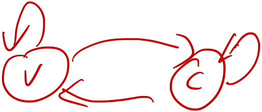
Markov. Alexander Pushkin. Eugene Onegin.

First 20000 letters. Vowel 8638

Consonant 11362.

$$\begin{array}{c|cc} & V & C \\ \hline V & \frac{1104}{8638} & \frac{7534}{8638} \\ C & \frac{7535}{11362} & \frac{3827}{11362} \end{array}$$

$$\begin{array}{c|cc} & V & C \\ \hline V & 0.175 & 0.825 \\ C & 0.526 & 0.474 \end{array}$$




(i.i.d)

Markov showed Law of Large Numbers for dependent r.v.s.

# Example: Chained to the Weather

Minnesota winter.

rain (r) snow (s) clear (c)

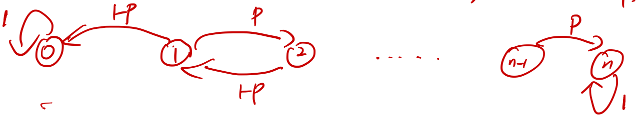
$$
\begin{array}{c c}
 & \begin{array}{ccc} r & s & c \end{array} \\
\begin{array}{c} \underline{r} \\ \underline{s} \\ \underline{c} \end{array} & \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.6 & 0.3 \end{bmatrix}
\end{array}
$$

# Example: Gambler's Ruin

game. win \$1 w.p. $p$

lose \$1 w.p $1-p$

stops either lose all money or win $n$ \$.




A gambler starts with $k$ \$ $(0 < k < n)$

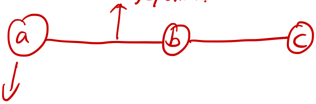
Markov chain state space $S = \{0, 1, \ldots, n\}$, $X_0 = k$.

$$P_{ij} = \begin{cases} p & j = i+1, \quad 0 < i < n \\ 1-p & j = i-1, \quad 0 < i < n \\ 1 & i = j = 0 \text{ or } i = j = n \\ 0 & \text{otherwise} \end{cases}$$

# Example: Random Walk on A Graph

Undirected graph.

edge/link.




node/vertex.

degree of node.

$\deg(a) = 1$

$\deg(b) = 2$

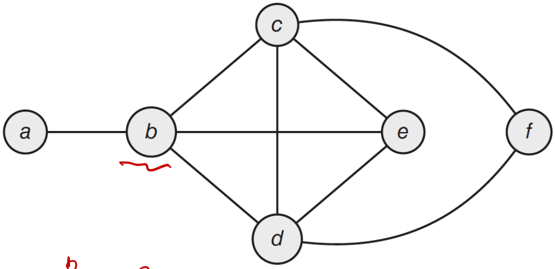
$\deg(c) = 1$

$X_0, X_1, \ldots X_2, \ldots$

random walk.

# Example: Random Walk on A Graph

equal probability for all options




$p_{b,a} = p_{b,c}$

$= p_{b,d}$

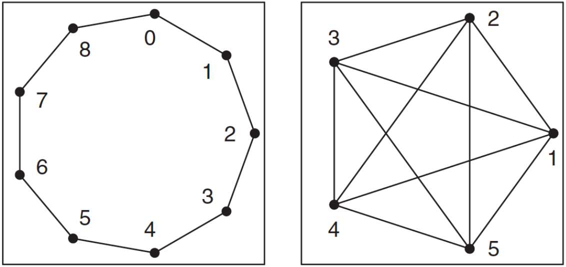
$= p_{b,e}$

$= \frac{1}{4}.$

$p_{b,f} = 0$

# Example: Random Walk on A Graph

$$p_{ij} = \begin{cases} \frac{1}{2} & \text{if } j \equiv i \pm 1 \pmod{k} \\ 0 & \text{otherwise.} \end{cases}$$




(a) cycle graph.

$k$ : # of all nodes.

(b) Complete graph

$k$ : # of all nodes.

$$\deg(v) = k-1, \quad p_{ij} = \begin{cases} \frac{1}{k-1} & \text{if } i \neq j \\ 0 & i = j. \end{cases}$$

# Example: Random Walk on A Graph

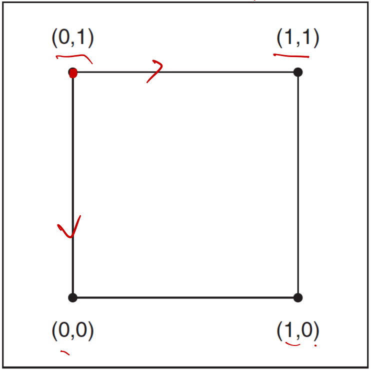
Two vertices are connected iff they differ in exactly one coordinate.

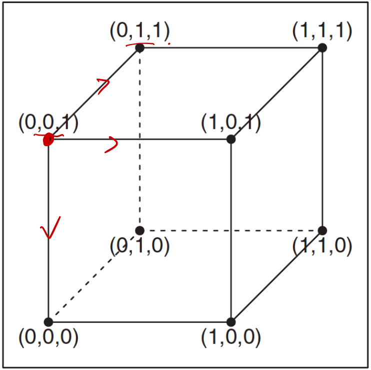
Hamming Distance.

$a = (a_1, a_2, \dots, a_n)$

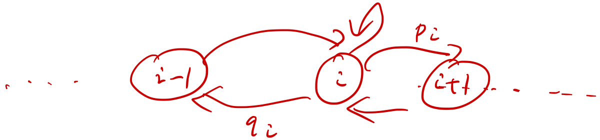
$b = (b_1, b_2, \dots, b_n)$

$d^H(a,b) = \sum_{j=1}^{n} 1_{\{a_j \neq b_j\}}$



# Example: Birth-and-Death Chain




$$P_{ij} = \begin{cases} q_i & \text{if } j = i-1 \\ p_i & \text{if } j = i+1 \\ 1-p_i-q_i & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

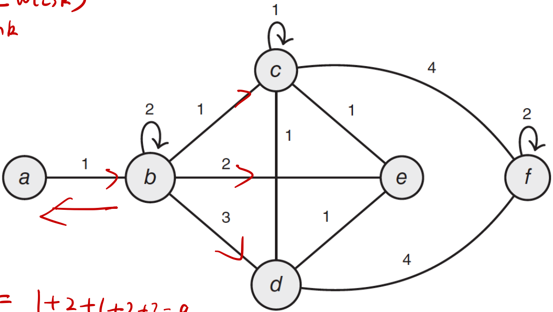
$$0 \le p_i, q_i \le 1$$

$$0 \le p_i + q_i \le 1$$

# Example: Weighted ~~Directed~~ Undirected Graphs

$$P_{i,j} = \begin{cases} \frac{w(i,j)}{w(i)} & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases}$$

$$w(i) = \sum_{k:\, i \sim k} w(i,k)$$




$$w(b) = 1+2+1+2+3 = 9$$

$$P_{b,a} = \frac{w(b,a)}{w(b)} = \frac{1}{9}.$$

# Outline

# $n$-step Transition Probability

$1^\circ$. $P(X_n=j \mid X_0=i) \overset{\text{LOTP}}{=} \sum_k P(X_n=j \mid X_{n-1}=k, X_0=i)\, P(X_{n-1}=k \mid X_0=i)$

$\overset{\text{Markov}}{=} \sum_k P(X_n=j \mid X_{n-1}=k)\, P(X_{n-1}=k \mid X_0=i)$

$\overset{\text{time-homogeneous.}}{=} \sum_k p_{kj} \cdot P(X_{n-1}=k \mid X_0=i)$

## Definition

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 $q_{ij}^{(n)}$:

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

$2^\circ$. $n=2$; $P(X_2=j \mid X_0=i) = \sum_k p_{kj} \cdot P(X_1=k \mid X_0=i) = \sum_k p_{kj} \cdot p_{ik} = (P^2)_{(i,j)}$

$\sum_k p_{ik} \cdot p_{kj}$

$n=3$, $P(X_3=j \mid X_0=i) = (P^3)_{(i,j)}$

# Example: Gambler's Ruin

State space $S = \{0, 1, \dots 8\}$

$X_0 = 3$, $n = 8$;

$$p_{ij} = \begin{cases} 0.6 & j = i+1, \ 0 < i < 8 \\ 0.4 & j = i-1, \ 0 < i < 8 \\ 1 & i = j = 0 \text{ or } i = j = 8 \\ 0 & \text{otherwise} \end{cases}$$

For gambler's ruin, assume that the gambler's initial stake is \$3 and the gambler plays until either gaining \$8 or going bust. At each play the gambler wins \$1, with probability 0.6, or loses \$1, with probability 0.4. Find the gambler's expected fortune after four plays.

① Find transition matrix $P$

② Find $P^4$

③ $X_0 = 3, X_1, X_2, X_3, X_4$

$$E[X_4 \mid X_0 = 3] = \sum_{j=0}^{8} j \cdot P(X_4 = j \mid X_0 = 3) = \sum_{j=0}^{8} j \cdot q_{3j}^{(4)} = \sum_{j=0}^{8} j \cdot P^4_{(3,j)}$$

$$= 3.79$$

# Solution

# Chapman-Kolmogorov Relationship

$m, n \geq 0$, matrix identity

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

$$\Rightarrow \quad P_{ij}^{m+n} = \sum_k P_{ik}^m \cdot P_{kj}^n \quad \forall i, j.$$

$$\Rightarrow P(X_{m+n} = j \mid X_0 = i) = \sum_k P(X_m = k \mid X_0 = i)\, \underline{P(X_n = j \mid X_0 = k)}$$

time-homogeneity

$$= \sum_k P(X_m = k \mid X_0 = i) \cdot \underline{P(X_{m+n} = j \mid X_m = k)}$$

$$i \xrightarrow{m+n} j \iff i \xrightarrow{m} k \xrightarrow{n} j.$$

# Distribution of $X_n$

Let $X_0, X_1, \ldots$ be a Markov chain with transition matrix $\mathbf{P}$ and initial distribution $\boldsymbol{\alpha}$. For all n $\geq 0$, the distribution of $X_n$ is $\boldsymbol{\alpha}\mathbf{P}^n$. That is,

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

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

$$= \sum_i P_{ij}^{(n)} \cdot \alpha_i$$

# Markov Property

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

$$\begin{aligned} P(X_{n+1} = j | X_0 = i_0, \ldots, X_{n-m-1} = i_{n-m-1}, X_{n-m} = i) \\ = P(X_{n+1} = j | X_{n-m} = i) \quad \text{Markov. + math induction.} \\ = P(X_{m+1} = j | X_0 = i) = P_{ij}^{m+1}, \end{aligned}$$

time-homogeneity

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

# Joint Distribution

$P(X_5=i, X_6=j, X_9=k, X_{17}=l) = P(X_5=i)\,P(X_6=j \mid X_5=i) \cdot P(X_9=k \mid X_5=i, X_6=j) \cdot P(X_{17}=l \mid X_5=i, X_6=j, X_9=k)$

$= P(X_5=i)\,P(X_6=j \mid X_5=i) \cdot P(X_9=k \mid X_6=j)\,P(X_{17}=l \mid X_9=k)$

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

Chain rule + Markov property.

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

$= (\alpha P^5)_i \cdot P_{ij} \cdot P_{jk}^{(3)} \cdot P_{kl}^{(8)}$

$= (\alpha P^5)_i\,(P^1)_{ij}\,(P^3)_{jk}\,(P^8)_{kl}$

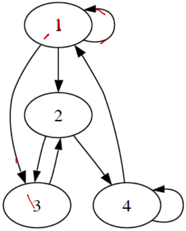
# Outline

# Recurrent and Transient States

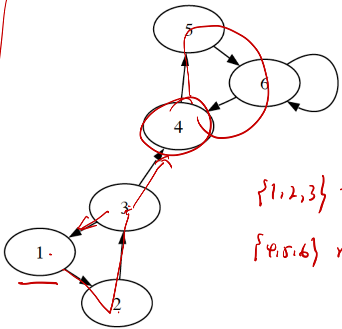
## Definition

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




All states are recurrent.




$\{1,2,3\}$ transient

$\{4,5,6\}$ recurrent

# Recurrence and Transience

Expectation of returny steps.

positive recurrent $\sum_{n=0}^{\infty} n \cdot P_{ii}^{n} < \infty.$

recurrent

Null recurrent $\sum_{n=0}^{\infty} n \cdot P_{ii}^{n} = \infty.$

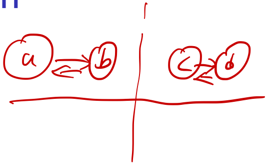
1. State $j$ is recurrent if and only if

$$\sum_{n=0}^{\infty} P_{jj}^{n} = \infty$$

2. State $j$ is transient if and only if

$$\sum_{n=0}^{\infty} P_{jj}^{n} < \infty$$

# Irreducible and Reducible Chain


## Definition

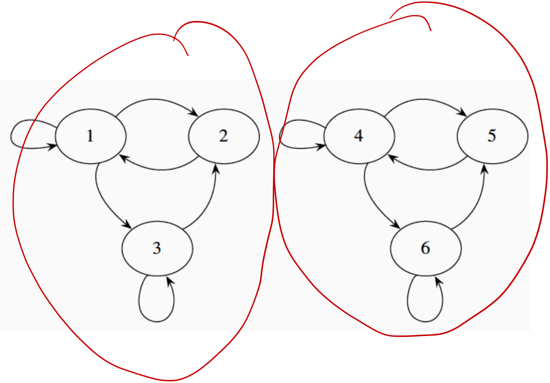
A Markov chain with transition matrix $Q$ is *irreducible* if for any two states $i$ and $j$, it is possible to go from $i$ to $j$ in a finite number of steps (with positive probability). That is, for any states $i$,$j$ there is some positive integer $n$ such that the $(i,j)$ entry of $Q^n$ is positive. 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

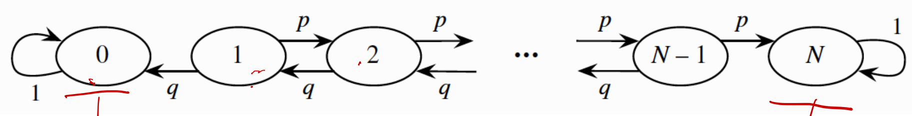


# Gambler's Ruin As A Markov Chain

$X_n$ : wealth of gambler at time $n$

State space : $\{0, 1, \ldots N\}$.

States $\{0, N\}$ : recurrent.




absorbing state

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

absorbing state.

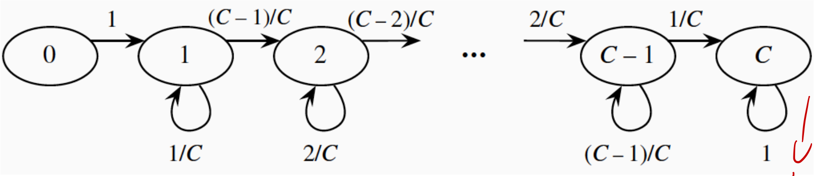
reducible M.C. $\quad (0 \not\to 1)$

# Coupon Collector As A Markov Chain

$X_n$ : # of distinct coupon types at time $n$.

State space : $\{0, 1, \ldots C\}$

$X_0 = 0$.

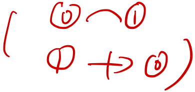



absorbing state.

State $C$ : recurrent

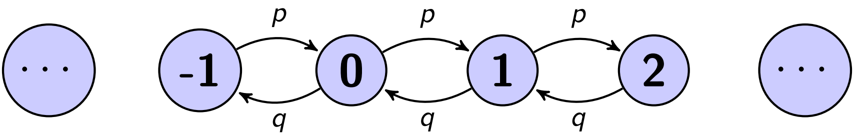
States $\{0, \ldots C-1\}$ : transient

reducible M.C.



# Example: One-Dimensional Random Walk

At each game, the gambler wins \$1 with the probability $p$ and loses \$1 with the probability $q = 1 - p$. Let $X_t$ be the total net winning after $t$ games. It is easy to see $\{X_t\}$ is a Markov chain and its state diagram is:




## Theorem

*If $p = q = 1/2$, this is a* symmetric *random walk and all states in this Markov chain are recurrent. Otherwise $p \neq q$, this is an* asymmetric *random walk and all states in this Markov chain are transient.*

# $N$-Dimension Symmetric Random Walk

## Theorem

*Given a N-dimensional symmetric random walk, all states in the corresponding Markov chain are*

- *Recurrent: when* $N = 1$ *or* $N = 2$,
- *Transient: when* $N \geq 3$.

# Period

## Definition

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

$$d(i) = gcd\{n > 0 : P_{ii}^n > 0\}.$$

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 different from Harvard's book.

## 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

$P_{1,1}^{(1)} > 0$, period(1) $= 1$

$P_{4,4}^{(1)} > 0$, period(4) $= 1$

$P_{2,2}^{(2)} > 0$.

2-3-2

$P_{2,2}^{(3)} > 0$

2-4-1-2

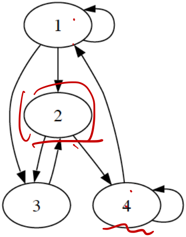
$P_{2,2}^{(4)} > 0$

2-4-1-1-2

period(2) $= \gcd(2,3,4,\dots)$
$= 1$

period(3) $= 1$

irreducible.

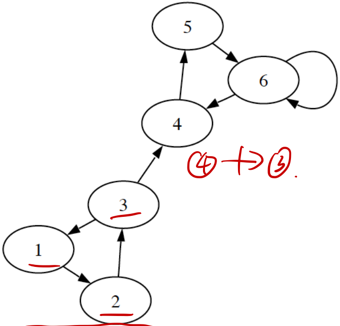



Markov chain

aperiodic

reducible.

reducible.




4 $\not\to$ 3.

# Outline

# Definition

## Definition

A row vector $\mathbf{s} = (s_1, ..., s_M)$ 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

$$\sum_i s_i q_{ij} = s_j.$$

for all $j$, or equivalently,

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

# Example: Two-State Markov Chain

$$Q = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\underline{S Q = S}$$

$$\Rightarrow \underline{S = \left(\frac{3}{7}, \frac{4}{7}\right).}$$

Stationary distribution

# Existence and Uniqueness of Stationary Distribution

**Theorem**

*Any irreducible Markov chain has a unique stationary distribution. In this distribution, every state has positive probability.*

# Convergence to Stationary Distribution

## Theorem

*Let $X_0, X_1, \ldots$ be a Markov chain with stationary distribution* $\mathbf{s}$ *and transition matrix $Q$, such that some power $Q^m$ is positive in all entries. (These assumptions are equivalent to assuming that the chain is irreducible and aperiodic.) 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* $\mathbf{s}$.

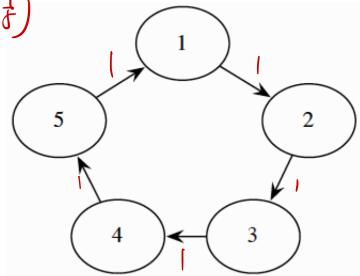
# Periodic Chain

$$Q = \begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 \\ 3 & 0 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 0 & 1 \\ 5 & 1 & 0 & 0 & 0 & 0 \end{array}$$

$sQ = s$

$\Rightarrow s = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$

$Q^n$ 0 and 1.




$\forall i \quad \lim_{n\to\infty} (Q^n)_{i\cdot} \neq s.$

# Fundamental Limit Theorem for Ergodic Markov Chains

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,*

$(P^n)_{ij}$

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

# Outline

# Reversibility

## Definition

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

$$s_i q_{ij} = s_j q_{ji}$$

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.

# Reversible implies Stationary

## Theorem

*Suppose that* $Q = (q_{i,j})$ *is a transition matrix of a Markov chain that is reversible with respect to a nonnegative vector* $\mathbf{s} = (s_1, \ldots, s_M)$ *whose components sum to* 1*. Then* $\mathbf{s}$ *is a stationary distribution of the chain.*

# Check the Detailed Balance Equation

## Theorem

*If for an irreducible Markov chain with transition matrix $Q = (q_{i,j})$, there exists a probability solution $\pi$ 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 $\pi$ is the unique stationary distribution.*

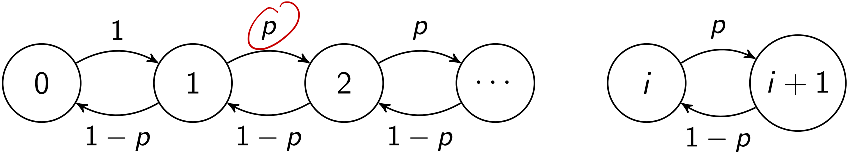
# Double Stochastic Matrix

**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: Simple Random Walk

Consider a negative drift simple random walk, restricted to be non-negative, in which $q_{0,1} = 1$, $q_{i,i+1} = p < 0.5$, $q_{i,i-1} = 1 - p$ for $i \geq 1$. Corresponding state diagram is:



# Example: Simple Random Walk

detailed balance equation

$$\begin{cases} \pi_0 \cdot 1 = (1-p)\pi_1 \\ \pi_i p = \pi_{i+1}(1-p) \quad i \geq 1 \\ \sum_i \pi_i = 1 \end{cases} \Rightarrow \begin{cases} \pi_0 = \frac{1-2p}{2(1-p)} \\ \pi_n = (\frac{1}{2}-p)\left[\frac{p}{1-p}\right]^{n-1} \quad n \geq 1 \end{cases}$$

Since this Markov chain is irreducible.

$\Rightarrow$ positive recurrent

time reversible.

$\pi$ is the stationary distribution

# Outline

# How to Organize the Web?

- First Try: Web Directories
- Yahoo, DMOZ, LookSmart

# Yahoo


YAHOO!
New
Cool
Today's News
More Yahoos
Exploring Mars
The New York Times
Win a $20,000 Trip!
CLICK HERE!
Looking for a Car? Job? Date?
Search options
Yellow Pages - People Search - Maps - Classifieds - News - Stock Quotes - Sports Scores
Arts and Humanities
Architecture, Photography, Literature...
News and Media [Xtra!]
Current Events, Magazines, TV, Newspapers...
Business and Economy [Xtra!]
Companies, Investing, Employment...
Recreation and Sports [Xtra!]
Sports, Games, Travel, Autos, Outdoors...
Computers and Internet [Xtra!]
Internet, WWW, Software, Multimedia...
Reference
Libraries, Dictionaries, Phone Numbers...
Education
Universities, K-12, College Entrance...
Regional
Countries, Regions, U.S. States...
Entertainment [Xtra!]
Cool Links, Movies, Music, Humor...
Science
CS, Biology, Astronomy, Engineering...
Government
Military, Politics [Xtra!], Law, Taxes...
Social Science
Anthropology, Sociology, Economics...
Health [Xtra!]
Medicine, Drugs, Diseases, Fitness...
Society and Culture
People, Environment, Religion...
My Yahoo! - Yahooligans! for Kids - Beatrice's Web Guide - Yahoo! Internet Life
Weekly Picks - Today's Web Events - Chat - Weather Forecasts
Random Yahoo! Link - Yahoo! Shop
National Yahoos Canada - France - Germany - Japan - U.K. & Ireland
Yahoo! Metros Atlanta - Austin - Boston - Chicago - Dallas / Fort Worth - Los Angeles
Get Local Minneapolis / St. Paul - New York - S.F. Bay - Seattle - Washington D.C.
How to Include Your Site - Company Information - Contributors - Yahoo! to Go

# How to Organize the Web?

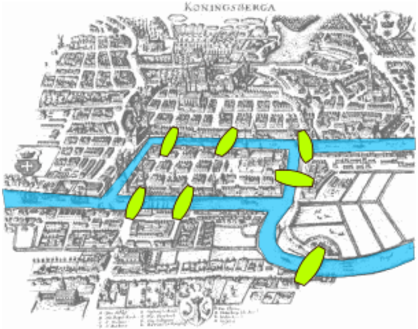
- Second Try: Web Search
- Information Retrieval:
  - find relevant docs in a small and trusted set
  - newspaper articles, patents, etc
- Hardness: web is huge, full of untrusted documents, random things, web spam, etc.

# Challenges for Web Search

- Web contains many sources of information. Who to trust?
  - **Trick**: Trustworthy pages may point to each other!
- What is the best answer to query keywords?
  - Webpages are not equally important (www.nothing.com vs.www.stanford.edu)
  - **Trick**: rank pages containing keywords according to their importances (popularity)
  - Find the page with the highest rank
  - How to rank?

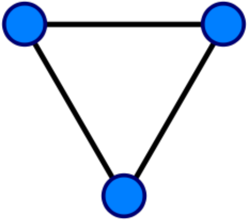
# Modeling Language: Graph Theory

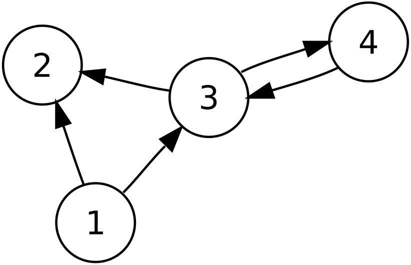
- Origin: 1735 Euler for Seven Bridges of Königsberg



# Key Elements of A Graph

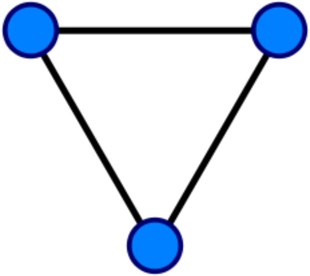
- A graph is an ordered pair $G = (V, E)$
- $V$: a set of vertices or nodes
- $E$: a set of edges or links between nodes
- Edge: undirected/directed





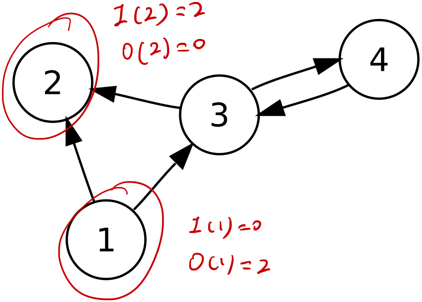
# Undirected Graph

- Degree of vertex $v$: metric for connectivity of vertex $v$.
- $deg(v)$: the number of edges with $v$ as an end vertex



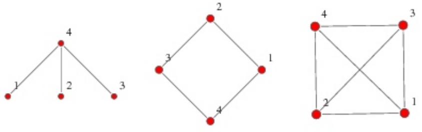
# Directed Graph

- Indegree of vertex $v$: the number of incoming edges ends at $v$.
- Outdegree of vertex $v$: the number of outgoing edges starting from $v$.
- $I(v)$: indegree of $v$
- $O(v)$: outdegree of $v$



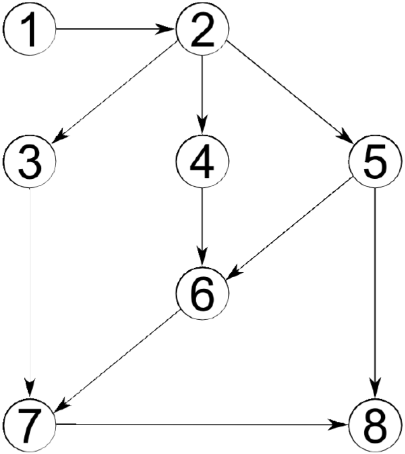
# Adjacency Matrix

- A square $(0, 1)-$matrix to represent a finite graph
- Matrix elements: pairs of vertices are adjacent or not
- Symmetric matrix: for undirected graph



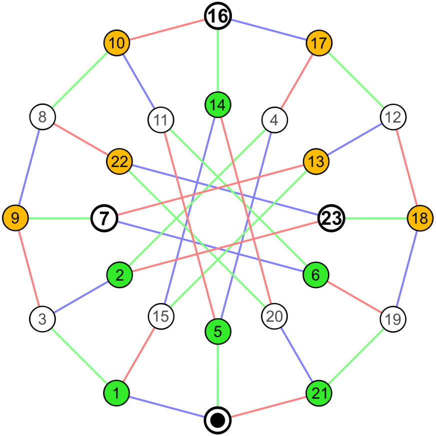
$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

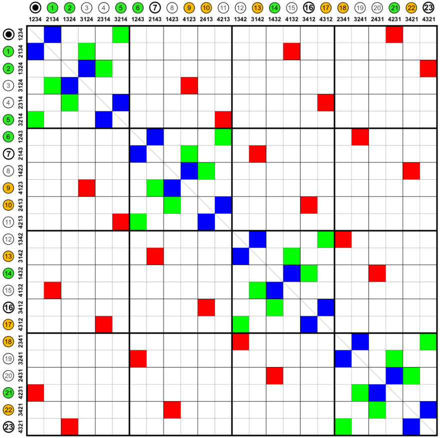
# Adjacency Matrix: Directed Graph



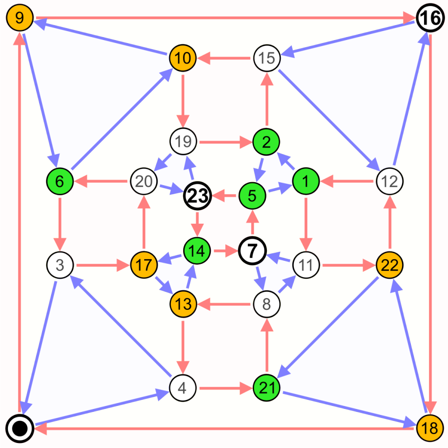

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **1** | | 1 | | | | | | |
| **2** | | | 1 | 1 | 1 | | | |
| **3** | | | | | | | 1 | |
| **4** | | | | | | 1 | | |
| **5** | | | | | | 1 | | 1 |
| **6** | | | | | | | 1 | |
| **7** | | | | | | | | 1 |
| **8** | | | | | | | | |

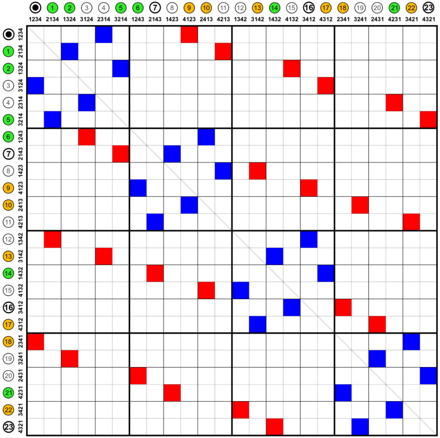
# Adjacency Matrix of Nauru Graph



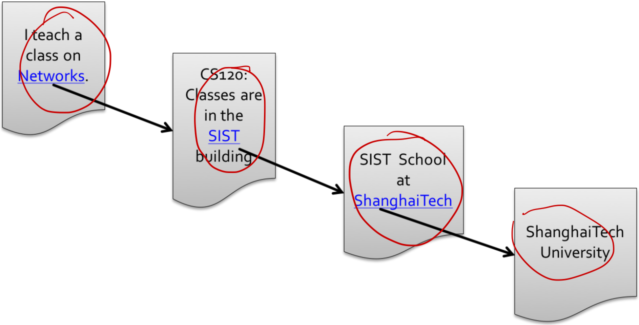
# Adjacency Matrix of Cayley Graph



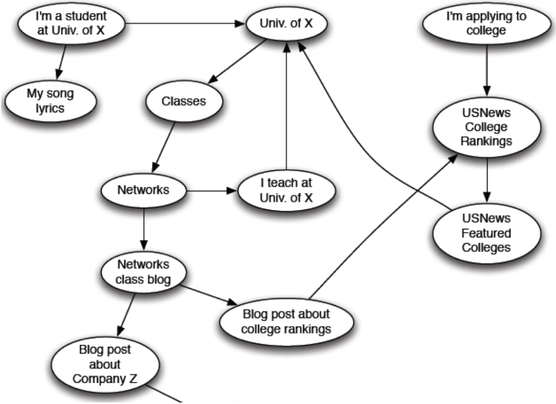



# World Wide Web as A Graph

- Web as a directed graph
- Nodes: webpages
- Edges: hyperlinks

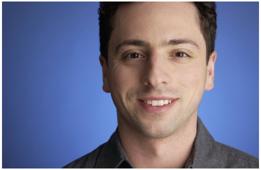


# Web as A Directed Graph



# Milestones in Networking

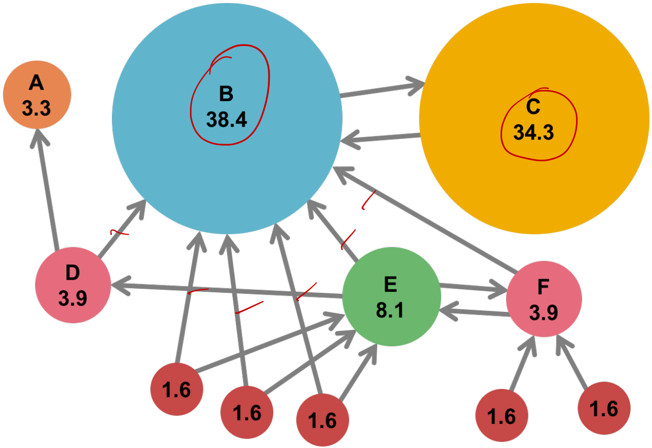
- 1998: Larry Page (1973-) and Sergey Brin (1973-) invented PageRank algorithm and then founded Google.
- 1998: Jon Kleinberg (1971-) invented Hyperlink-Induced Topic Search (HITS) algorithm.



# Links as Votes

- Page is more important if it has more links
- Incoming links or outgoing links?
- Think of incoming links as votes:
  - www.stanford.edu has 23400 incoming links
  - www.nothing.com has 1 incoming link
- Are all in-links are equal?
  - Links from important pages count more
  - Recursive question!

# Example: PageRank Scores

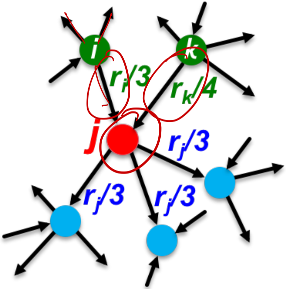


# Simple Recursive Formulation

- Each link's vote is proportional to the importance of its source page.
- If page $j$ with importance $r_j$ has $n$ out-links, each link gets $r_j/n$ votes
- Page j's own importance is the sum of the votes on its in-links

# Example

- $r_j = r_i/3 + r_k/4.$



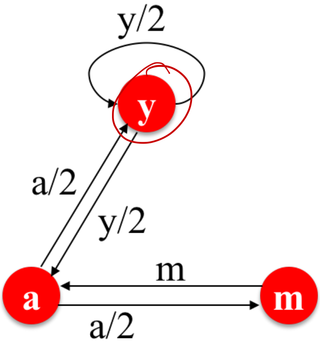
# PageRank: The "Flow" Model

- A "vote" from an important page is worth more
- A page is important if it is pointed to by other important pages
- Define a rank $r_j$ for page $j$:

$$r_j = \sum_{i \to j} \frac{r_i}{O_i}$$

where $O_i$ is the outdegree of $i$.

# Example: Flow Equation




$$
\begin{aligned}
r_y &= r_y/2 + r_a/2 \\
r_a &= r_y/2 + r_m \\
r_m &= r_a/2
\end{aligned}
$$

# Solving the Flow Equations

- Additional constraint forces uniqueness:
  - $r_y + r_a + r_m = 1$.
  - solution: $r_y = 2/5, r_a = 2/5, r_m = 1/5$.
- Gaussian elimination method works for small examples
- We need a better method for large web-size graphs

# PageRank: Matrix Formulation

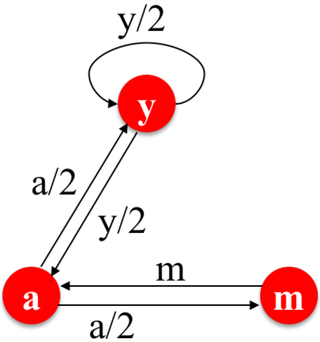
- Adjacency matrix $Q$
  - Each page $i$ has $O_i$ out-links
  - If $i \to j$, then $Q_{i,j} = \frac{1}{O_i}$, else $Q_{i,j} = 0$.
- $Q$ is a stochastic matrix
- Row sum to 1.

# PageRank: Matrix Formulation

- Rank vector $\mathbf{r}$
  - Vector with an entry per page
  - $r_i$ is the importance score of page $i$
  - $\sum_i r_i = 1$
- The flow equations can be written

$$\mathbf{r} = \mathbf{r} \cdot Q$$

# Example




$$r_y = r_y/2 + r_a/2$$
$$r_a = r_y/2 + r_m$$
$$r_m = r_a/2$$

$$Q = \begin{array}{c c} & \begin{array}{ccc} y & a & m \end{array} \\ \begin{array}{c} y \\ a \\ m \end{array} & \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix} \end{array}$$

# Random Walk Interpretation

random walk on a graph. $\begin{cases} V: \text{webpages.} \\ E: \text{hyperlink} \end{cases}$

$$a_{ij} = \frac{1}{O(i)}. \quad \text{if } \underline{i \to j} \in E$$

# The Stationary Distribution

- $\mathbf{r} \cdot Q = \mathbf{r}$.

# Power Iteration Method

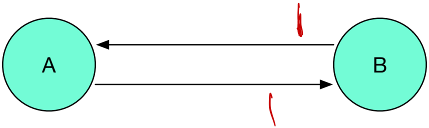
- Given a web graph with ~~n~~ N nodes, where the nodes are pages and edges are hyperlinks.
- Power iteration: a simple iterative scheme
  - Suppose there are $N$ web pages
  - Initialize: $\mathbf{r}(0) = [\frac{1}{N}, \ldots, \frac{1}{N}]$.
  - Iterate: $\mathbf{r}(t+1) = \mathbf{r}(t) \cdot Q$.
  - Stop when $|\mathbf{r}(t+1) - \mathbf{r}(t)|_1 < \epsilon$

# The Google Formulation

$$r_j^{(t+1)} = \sum_{i \to j} \frac{r_i^{(t)}}{O_i} \quad \Longrightarrow \quad \mathbf{r}^{(t+1)} = \mathbf{r}^{(t)} \cdot Q$$

$$\begin{cases} \text{Does this converge?} \\ \text{Converge to what we want?} \\ \text{Result reasonable?} \end{cases}$$

# Example 1: Spider Traps




$$r_A^{(t+1)} = r_B^{(t)}$$
$$r_B^{(t+1)} = r_A^{(t)}$$

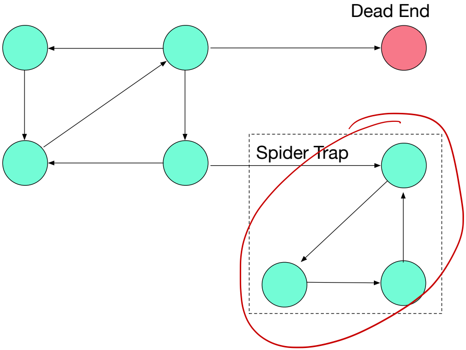
$$\begin{array}{lcccccc} r_A & \rightarrow & 1 & 0 & 1 & 0 & \dots \\ r_B & \rightarrow & 0 & 1 & 0 & 1 & \dots \end{array}$$

# Example 2: Dead End


absorbing state.

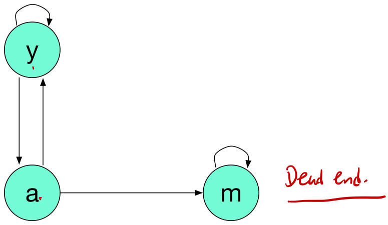
$$
\begin{array}{ccccccc}
r_A & \rightarrow & 1 & 0 & 0 & 0 & \ldots \\
r_B & \rightarrow & 0 & 1 & 0 & 0 & \ldots
\end{array}
$$

# Observations




- Some pages are dead ends (no out-links)
- Spider traps (all out-links are within the group)

# Example




$$\begin{array}{cccc} & y & a & m \\ y & 1/2 & 1/2 & 0 \\ a & 1/2 & 0 & 1/2 \\ m & 0 & 0 & 1 \end{array} \implies \begin{array}{cccccccc} r_y & \rightarrow & 1/3 & 2/6 & 3/12 & \ldots & \underline{0} \\ r_a & \rightarrow & 1/3 & 1/6 & 2/12 & \ldots & \underline{0} \\ r_m & \rightarrow & 1/3 & 3/6 & 7/12 & \ldots & \underline{1} \end{array}$$

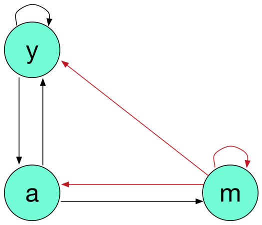
# Google's Solution for Spider Trap

- Teleports!
- At each time, walker at page $i$ has two options

$$\begin{cases} \text{w.p. } \beta & \text{Follow a out-link at random } \frac{1}{O_i} \\ \text{w.p. } 1-\beta & \text{Jump to some random pages} \end{cases} \tag{1}$$

- $\beta = (0.8, 0.9)$

# Google's Solution for Dead Ends




$$\begin{array}{cccc} & y & a & m \\ y & 1/2 & 1/2 & \mathbf{0} \\ a & 1/2 & 0 & \mathbf{1/2} \\ m & 0 & 0 & \mathbf{1} \end{array} \Longrightarrow \begin{array}{cccc} & y & a & m \\ y & 1/2 & 1/2 & 0 \\ a & 1/2 & 0 & 1/2 \\ m & \color{red}{1/3} & \color{red}{1/3} & \color{red}{1/3} \end{array}$$

# Random Teleports

- Pagerank Equation:

$$r_j = \sum_{i \to j} \beta \cdot \frac{r_i}{O_i} + (1-\beta) \cdot \frac{1}{N}$$

- Google Matrix

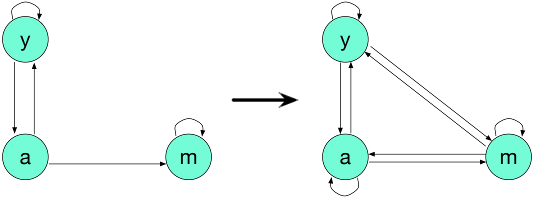
$$G = \beta \cdot Q + (1-\beta) \left[ \frac{1}{N} \right]_{N \times N}$$

$$\mathbf{r} = \mathbf{r} \cdot G$$

# Random Teleports

- Google Matrix

$$Q \to G = \beta \cdot Q + (1-\beta)\left[\frac{1}{N}\right]_{N\times N}$$



# Do Some Calculation

Let $\beta = 0.8$.

$$Q = 0.8 \cdot \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix} + 0.2 \cdot \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

As a result,

$$G = \begin{pmatrix} 7/15 & 7/15 & 1/15 \\ 7/15 & 1/15 & 7/15 \\ 1/15 & 1/15 & 13/15 \end{pmatrix}$$

# Implementation of PageRank in Practice

- BigTable: distributed storage system
- GFS (Google File System): distributed file system
- Mapreduce: distributed computing system (followed by Hadoop & Spark)

# Outline

# Continuous Time Markov Chain (CTMC)

- 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.

# 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

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

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

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.

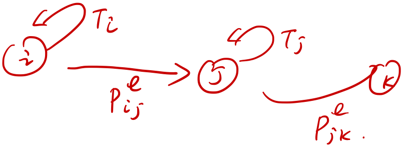
# Embedded Chains

For each state $i$, holding time $T_i \sim \exp(v_i)$, $E(T_i) = \frac{1}{v_i}$

$v_i = 0$ (absorbing state)

$v_i = \infty$ (explosive)

$0 < v_i < \infty$.

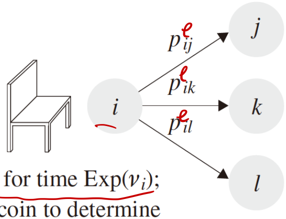



$P^e_{i,j} \triangleq$ discrete M.C. (transition from state $i \to$ state $j$)

$P^e_{i,i} \triangleq 0.$  $P^e = (P^e)_{ij}$ embedded chain

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

# CTMC: Perspective 1




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

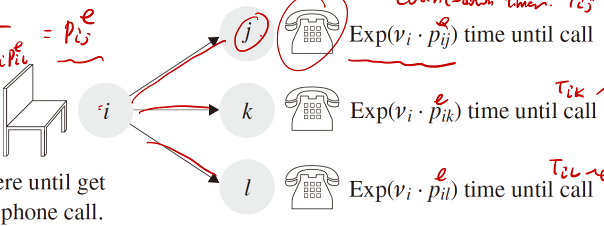
$\sum_{i'} p^e_{ii'} = 1$

# CTMC: Perspective 2

1°. $T_i = \min(T_{ij}, T_{ik}, T_{il}) \sim \exp(v_i(p^e_{ij} + p^e_{ik} + p^e_{il})) = \exp(v_i)$

2°. $P(T_{ij} = \min(T_{ij}, T_{ik}, T_{il})) = \dfrac{v_i p^e_{ij}}{v_i p^e_{ij} + v_i p^e_{ik} + v_i p^e_{il}} = p^e_{ij}$

Count-down timer. $T_{ij} \sim \exp(v_i p^e_{ij})$




$T_{ik} \sim \exp(v_i p^e_{ik})$

$T_{il} \sim \exp(v_i p^e_{il})$

# Generator Matrix

$1^\circ$. $T_{ij} \sim \exp(v_i p^e_{ij})$

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

$\Rightarrow \sum_{j \neq i} q_{ij} = v_i \left( \sum_{j \neq i} p^e_{ij} \right) = v_i$

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

$p^e_{ij} = \frac{q_{ij}}{v_i}$ . $i \neq j$

$2^\circ$. Define a matrix $Q$ $\quad Q_{ij} = q_{ij} \quad i \neq j$

$Q_{ii} = -v_i$

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

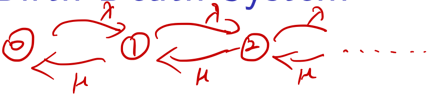
# Instantaneous Rates, Holding Times, Transition Probabilities

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

$$q_{ij} = v_i p_{ij}^e.$$

# Example: Birth-Death System

$\lambda, \mu$. transition rate.




generator matrix

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

$1^0$. Suppose we are in state $i$ $\quad i \geq 1$

$i \to i+1$ or $i \to i-1$

$X_B$ : the time to the next birth $\quad (X_B \sim \exp(\lambda))$

$X_D$ : . . . . . . . . . . . . . death. $\quad (X_D \sim \exp(\mu))$

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

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

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

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

# Example: Birth-Death System

# Stationary Distribution & Generator Matrix

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

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

That is,

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

# Time Reversibility

## Definition

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

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

# Ergodicity of 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 $\pi$ to the detailed balance equation

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

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

# Birth-and-Death Process

$q_{i,i+1} = \lambda_i, \; i \geq 0$

$q_{i,i-1} = \mu_i, \; i \geq 1$

Solve the detailed balance equation

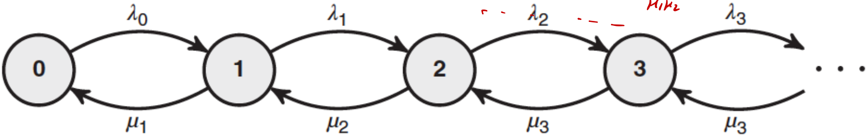
$\pi_i q_{i,i+1} = \pi_{i+1} q_{i+1,i}$

$\Rightarrow \pi_i \lambda_i = \pi_{i+1} \mu_{i+1}$

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

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

$\cdots$




$\sum_{k=0}^{\infty} \pi_k = 1$

# Stationary Distribution for Birth-and-Death Process

For a birth-and-death process with birth rates $\lambda_i$ and death rates $\mu_i$, 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

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

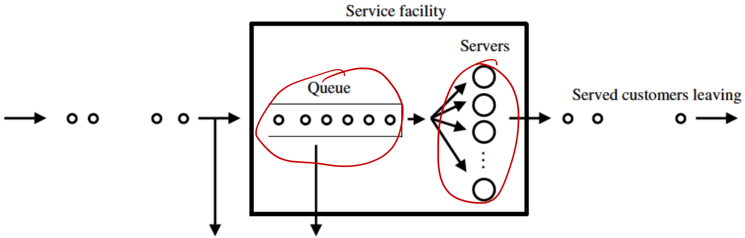
# Common Birth-and-Death Processes

**TABLE 7.1 Types of Birth-and-Death Processes**

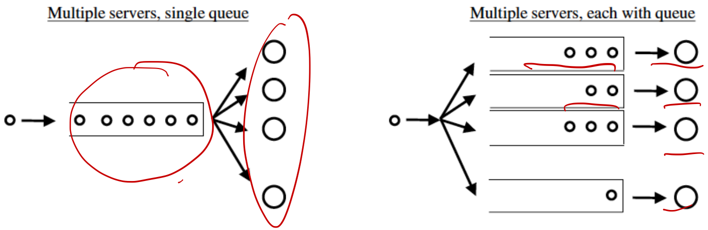
| Type | Birth Rate | Death Rate |
|---|---|---|
| Pure birth | $\lambda_i$ | $\mu_i = 0$ |
| Poisson process | $\lambda_i = \lambda$ | $\mu_i = 0$ |
| Pure death | $\lambda_i = 0$ | $\mu_i$ |
| Linear process | $\lambda_i = i\lambda, i > 0$ | $\mu_i = i\mu$ |
| Yule process | $\lambda_i = \lambda i, i, \lambda > 0$ | $\mu_i = 0$ |
| 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 ($A$) Service-time distribution ($B$) | $M$ | Exponential |
| | $D$ | Deterministic |
| | $E_k$ | Erlang type $k$ ($k = 1, 2, \ldots$) |
| | $H_k$ | 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 |

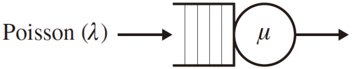
# Little's Law

M/G/1

In a queueing system, let $L$ denote the long-term average number of customers in the system, $\lambda$ 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



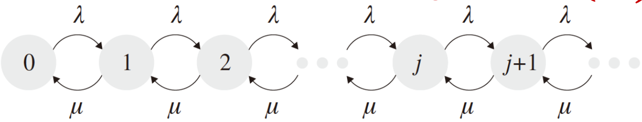
# M/M/1 Queue

Birth-death system

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

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

$\Rightarrow \pi_i = (1-\frac{\lambda}{\mu})(\frac{\lambda}{\mu})^i \quad i \geq 0. \quad (\frac{\lambda}{\mu} < 1)$



Let $\rho = \frac{\lambda}{\mu} < 1 \quad \Rightarrow \pi_i = (1-\rho) \cdot \rho^i$

# M/M/1 Queue

1°. Mean # of customers in the system $E(N) = \sum_{i=0}^{\infty} i\pi_i = \sum_{i=0}^{\infty} i(1-\rho)\rho^i = \frac{\rho}{1-\rho}$

2°. Mean # of customers in the queue $E(N_q) = \sum_{i=1}^{\infty} (i-1)\pi_i = \frac{\rho^2}{1-\rho}$.

3°. Mean time in the system $E(T) = \frac{E(N)}{\lambda} = \frac{1}{\mu-\lambda}$. (Little's law)

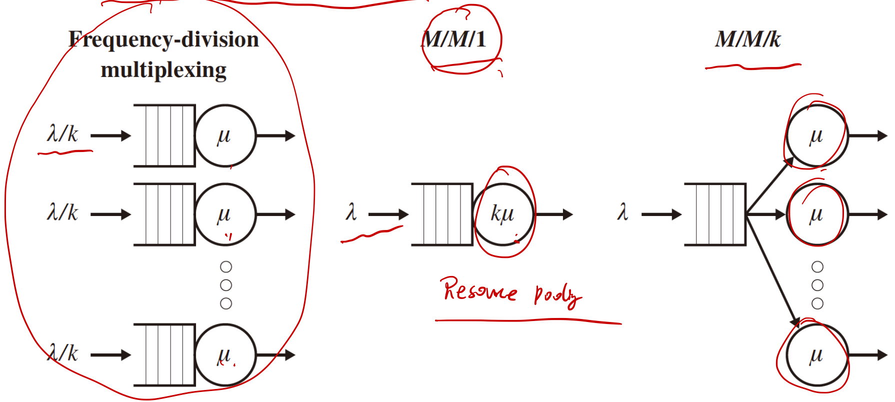
4°. Mean time in the queue. $E(T_q) = E(T) - \frac{1}{\mu}$ (service time)

$$= \frac{\rho}{\mu-\lambda} = \frac{E(N_q)}{\lambda}$$

Little's law

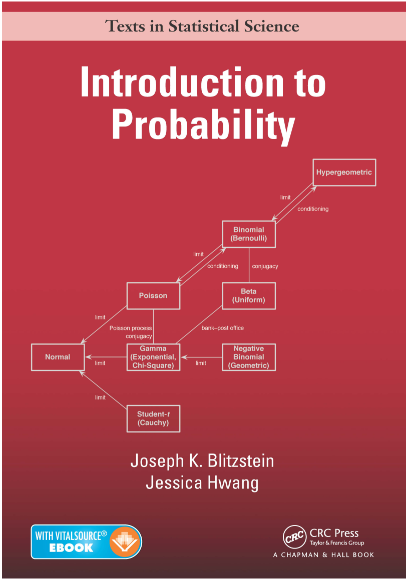
# Application: Three Server Organizations

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



# Outline

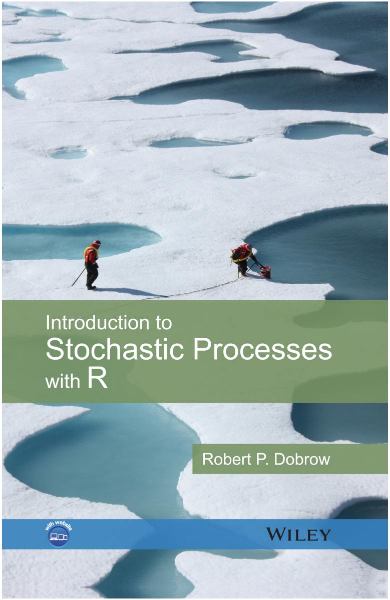
# BH




## BH

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

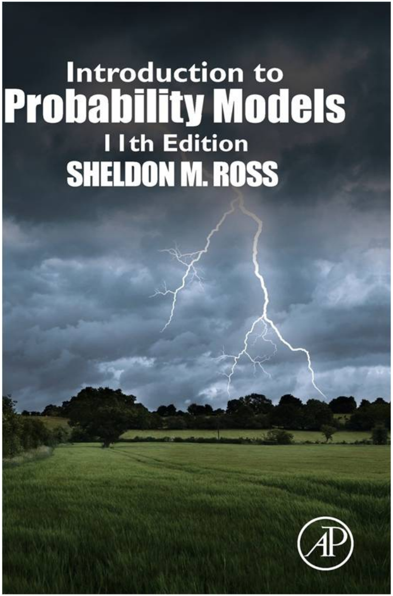
# SRR




## SPR

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

# SMR




## SMR

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