Advanced Algorithms

南京大学

尹一通
Markov Chain

- stochastic process: $X_0, X_1, X_2, \ldots$
  
  $X_t \in \Omega$
  
  Markov property (memoryless):
  
  $X_{t+1}$ depends only on $X_t$
  
  $\Pr[X_{t+1} = y \mid X_0 = x_0, \ldots, X_{t-1} = x_{t-1}, X_t = x]$
  
  $= \Pr[X_{t+1} = y \mid X_t = x]$

Markov chain: discrete time discrete space stochastic process with Markov property.
Transition Matrix

• **Markov chain**: $X_0, X_1, X_2, \ldots \in \Omega$

  \[
  \Pr[X_{t+1} = y \mid X_0 = x_0, \ldots, X_{t-1} = x_{t-1}, X_t = x] = \Pr[X_{t+1} = y \mid X_t = x] = P_{xy}^{(t)} = P_{xy}
  \]

  *(time-homogenous)*

• **homogeneity**: transition does not depend on time

• **transition matrix** $P$ over $\Omega \times \Omega$

  *(row-)stochastic matrix*: $P1 = 1$

\[
\Pr[X_{t+1} = y] = \sum_{x \in \Omega} \Pr[X_{t+1} = y \mid X_t = x] \Pr[X_t = x]
\]
Transition Matrix

- **Markov chain**: $\mathcal{M}, X(\Omega, \mathcal{P}) \ldots \in \Omega$
  
  Distribution: $p^{(t)}(x) = \Pr[X_t = x]$

- **transition matrix** $P$ over $\Omega \times \Omega$

  $$P(x, y) = \Pr[X_{t+1} = y \mid X_t = x]$$

\[
\begin{align*}
p^{(t+1)} &= p^{(t)} P \\
p^{(0)} &\xrightarrow{P} p^{(1)} &\xrightarrow{P} \ldots &\xrightarrow{P} p^{(t)} &\xrightarrow{P} \ldots
\end{align*}
\]

- **initial distribution**
  - $p^{(0)}$
- **distribution of** $X_t$
  - $p^{(t)}$
\[ P = \begin{bmatrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \]
Convergence

\[
P = \begin{bmatrix}
0 & 1 & 0 \\
1/3 & 0 & 2/3 \\
1/3 & 1/3 & 1/3
\end{bmatrix}
\]

\[
P^{20} \approx \begin{bmatrix}
0.2500 & 0.3750 & 0.3750 \\
0.2500 & 0.3750 & 0.3750 \\
0.2500 & 0.3750 & 0.3750
\end{bmatrix}
\]

\[
\forall \text{initial distribution } p^{(0)}:\quad p^{(20)} = p^{(0)}P^{20} \approx (\frac{1}{4}, \frac{3}{8}, \frac{3}{8})
\]
Stationary Distribution

Markov chain $\mathcal{M} = (\Omega, P)$

- stationary distribution $\pi$:
  \[ \pi P = \pi \quad \text{(fixed point)} \]

- Perron-Frobenius Theorem:
  - stochastic matrix $P$:
    \[ P1 = 1 \]
  - 1 is also a left eigenvalue of $P$ (eigenvalue of $P^T$)
  - the left eigenvector $\pi P = \pi$ is nonnegative
- stationary distribution always exists
The transition matrix $P$ is given by:

$$P = \begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
\frac{1}{3} & \frac{2}{3} & 0 & 0 \\
0 & 0 & \frac{3}{4} & \frac{1}{4} \\
0 & 0 & \frac{1}{4} & \frac{3}{4}
\end{bmatrix}$$

The matrix $P^{20}$ is approximately:

$$P^{20} \approx \begin{bmatrix}
0.4 & 0.6 & 0 & 0 \\
0.4 & 0.6 & 0 & 0 \\
0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0.5 & 0.5
\end{bmatrix}$$
$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$P^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$P^{2k} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  
$P^{2k+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
reducible

periodic
Fundamental Theorem of Markov Chain:

If a finite Markov chain $\mathcal{M} = (\Omega, P)$ is irreducible and aperiodic, then $\forall$ initial distribution $\pi^{(0)}$ (ergodic)

$$\lim_{t \to \infty} \pi^{(0)} P^t = \pi$$

where $\pi$ is a unique stationary distribution satisfying

$$\pi P = \pi$$
Irreducibility

• \( y \) is **accessible** from \( x \):
  \[ \exists t, \ P^t(x, y) > 0 \]

• \( x \) **communicates** with \( y \):
  • \( x \) is accessible from \( y \)
  • \( y \) is accessible from \( x \)

• MC is **irreducible**: all pairs of states communicate
Reducible Chains

stationary distributions: \( \pi = \lambda \pi_A + (1 - \lambda) \pi_B \)

stationary distribution: \( \pi = (0, \pi_B) \)

\[ P = \begin{bmatrix} P_A & 0 \\ 0 & P_B \end{bmatrix} \]
Aperiodicity

- **period** of state $x$:
  
  $$d_x = \gcd\{t \mid P^t(x, x) > 0\}$$

- **aperiodic** chain: all states have period 1

- **period**: the gcd of lengths of cycles

  $$\begin{array}{c}
  x \\
  \square \quad \square \quad \triangle \quad \square \quad \square \quad \triangle \quad \square \quad \square \quad \triangle \\
  \end{array}$$

  $$d_x = 3$$

A chain is *aperiodic* if $\forall x \in \Omega$, $P(x, x) > 0$. (every state has a self-loop)
If a finite Markov chain \( \mathcal{M} = (\Omega, P) \) is irreducible and aperiodic, then \( \forall \) initial distribution \( \pi^{(0)} \)

\[
\lim_{t \to \infty} \pi^{(0)} P^t = \pi
\]

where \( \pi \) is a unique stationary distribution satisfying

\[
\pi P = \pi
\]
Fundamental Theorem of Markov Chain:

If a Markov chain $\mathcal{M} = (\Omega, P)$ is irreducible and ergodic, then $\forall$ initial distribution $\pi^{(0)}$

$$\lim_{t \to \infty} \pi^{(0)} P^t = \pi$$

where $\pi$ is a unique stationary distribution satisfying $\pi P = \pi$

**ergodic**: convergent

- finit chain: aperiodic
- infinit chain: aperiodic + non-null persistent
Random Walk on Graph

undirected graph \( G(V,E) \)

- **uniform random walk**: \( \Omega = V \)

  at each step, the current position is \( u \in V \):
  - pick a neighbor \( v \) of \( u \) uniformly at random;
  - move to vertex \( v \);

- **transition matrix**:
  \[
  P(u, v) = \begin{cases} 
  \frac{1}{\deg(u)} & \text{if } uv \in E \\
  0 & \text{if } uv \notin E
  \end{cases}
  \]

  irreducible \( \iff \) \( G \) is connected

  aperiodic \( \iff \) \( G \) is non-bipartite
Random Walk on Graph

undirected graph $G(V,E)$

• lazy random walk: $\Omega = V$

at each step, the current position is $u \in V$:

• (lazy) for probability $1/2$, do nothing;

• else: pick a neighbor $v$ of $u$ uniformly at random and move to vertex $v$;

• transition matrix:

$$P(u, v) = \begin{cases} 
\frac{1}{2} & \text{if } u = v \\
\frac{1}{2\deg(u)} & \text{if } uv \in E \\
0 & \text{otherwise}
\end{cases}$$

irreducible $\iff G$ is connected

always aperiodic!
Random Walk on Graph

undirected graph $G(V,E)$

- **uniform random walk:**
  $$P(u, v) = \begin{cases} 
  \frac{1}{\deg(u)} & \text{if } uv \in E \\
  0 & \text{if } uv \notin E
  \end{cases}$$

- **lazy random walk:**
  $$P(u, v) = \begin{cases} 
  \frac{1}{2} & \text{if } u = v \\
  \frac{1}{2\deg(u)} & \text{if } uv \in E \\
  0 & \text{otherwise}
  \end{cases}$$

stationary distribution $\pi(u) = \frac{\deg(u)}{2|E|}$

uniform walk:
$$\left(\pi P\right)_v = \sum_{u \in V} \pi(u)P(u, v) = \sum_{u \sim v} \frac{\deg(u)}{2|E|} \frac{1}{\deg(u)} = \frac{\deg(v)}{2|E|} = \pi(v)$$

lazy walk:
$$P' = \frac{1}{2}(I + P) \quad \pi P = \pi \quad \pi P' = \pi$$
Reversibility

**Detailed Balance Equation:**

\[ \pi(x)P(x, y) = \pi(y)P(y, x) \]

**time-reversible** Markov chain:

\[ \exists \pi, \forall, x, y \in \Omega, \quad \pi(x)P(x, y) = \pi(y)P(y, x) \]

**stationary distribution:**

\[ (\pi P)y = \sum_x \pi(x)P(x, y) = \sum_x \pi(y)P(y, x) = \pi(y) \]

**time-reversible:** when start from \( \pi \)

\[
\Pr[X_0 = x_0 \land X_1 = x_1 \land \ldots \land X_n = x_n] = \Pr[X_0 = x_n \land X_1 = x_{n-1} \land \ldots \land X_n = x_0]
\]
Reversibility

**Detailed Balance Equation:**
\[ \pi(x)P(x, y) = \pi(y)P(y, x) \]

**time-reversible Markov chain:**
\[ \exists \pi, \forall x, y \in \Omega, \quad \pi(x)P(x, y) = \pi(y)P(y, x) \]

**stationary distribution:**
\[ (\pi P)y = \sum_x \pi(x)P(x, y) = \sum_x \pi(y)P(y, x) = \pi(y) \]

**time-reversible:** when start from \( \pi \)
\[ (X_0, X_1, \ldots, X_n) \sim (X_n, X_{n-1}, \ldots, X_0) \]

**ergodic flow**
Random Walk on Graph

undirected graph $G(V,E)$

- **uniform random walk:** $P(u, v) = \begin{cases} \frac{1}{\deg(u)} & \text{if } uv \in E \\ 0 & \text{if } uv \notin E \end{cases}$

- **lazy random walk:** $P(u, v) = \begin{cases} \frac{1}{2} & \text{if } u = v \\ \frac{1}{2\deg(u)} & \text{if } uv \in E \\ 0 & \text{otherwise} \end{cases}$

**Detailed Balance Equation:**

$$\pi(x)P(x, y) = \pi(y)P(y, x)$$

$u = v$ or $u \sim v$: detailed balanced equation holds for free

$u \sim v$: DBE holds when $\pi(u) \propto \frac{1}{P(u, v)} \propto \deg(u)$
Random Walk on Graph

undirected graph $G(V,E)$ \quad \text{max-degree } \Delta = \max_v \deg(v)

$$P(u, v) = \begin{cases} 
1 - \frac{\deg(u)}{2\Delta} & \text{if } u = v \\
\frac{1}{2\Delta} & \text{if } uv \in E \\
0 & \text{otherwise}
\end{cases}$$

**Detailed Balance Equation:**

$$\pi(x)P(x, y) = \pi(y)P(y, x)$$

$\pi$ is uniform
Metropolis Algorithm

• **symmetric** transition matrix $Q$ over state space $\Omega$:
  
  \[
  \begin{align*}
  Q^T &= Q \\
  Q1 &= 1
  \end{align*}
  \]

  \[1Q = 1\] uniform stationary

• **Goal**: a Markov chain with stationary distribution $\pi$

Metropolis-Hastings Algorithm:

at each step, the current state is $x \in \Omega$:

• **(proposal)** propose $y \in \Omega$ with probability $Q(x,y)$;

• **(filter)** accept the proposal and move to $y$ with probability $\min\{1, \pi(y)/\pi(x)\}$;
• symmetric transition matrix $Q$ over state space $\Omega$:

**Metropolis-Hastings Algorithm:**

at each step, the current state is $x \in \Omega$:

- *(proposal)* propose $y \in \Omega$ with probability $Q(x, y)$;
- *(filter)* accept the proposal and move to $y$ with probability $\min\{1, \frac{\pi(y)}{\pi(x)}\}$;

For $x \neq y$,

$$P(x, y) = Q(x, y) \min\left\{1, \frac{\pi(x)}{\pi(y)}\right\}$$

For $x = y$,

$$P(x, x) = 1 - \sum_{y \neq x} P(x, y)$$

**Detailed Balance Equation:**

$$\pi(x)P(x, y) = \pi(y)P(y, x)$$
Constraint Satisfaction Problem

• variables: \( X = \{x_1, x_2, \ldots, x_n\} \)
• domain: \( \Omega \), usually \( \Omega = [q] \) for a finite \( q \)
• constraints: \( C = (\psi, S) \) where \( \psi: \Omega^k \rightarrow \{0,1\} \) and scope \( S \subseteq X \) is a subset of \( k \) variables
• CSP instance \( I \): a set of constraints defined on \( X \)
• assignment: \( \sigma \in \Omega^X \) assigns values to variables
• a constraint \( C = (\psi, S) \) is satisfied if \( \psi(\sigma_S) = 1 \)
• CSP solution: an assignment \( \sigma \) is a solution to a CSP instance if it satisfies all constraints.
Input: a CSP instance $I$ on $n$ variables with domain $[q]$;
Sample a uniform random CSP solution.

**Metropolis Algorithm:**
Initially, start with an arbitrary CSP solution; at each step, the current CSP solution is $\sigma=(\sigma_1, \ldots, \sigma_n)$:

- **(proposal)** pick a variable $i \in [n]$ and value $c \in [q]$ uniformly at random;
- **(filter)** accept the proposal and change $\sigma_i$ to $c$ if it does not violate any constraint;

**detailed balanced equation** → **uniform stationary distribution!**
Initially, start with an arbitrary independent set; at each step:

- (proposal) pick a vertex \( v \in V \) and \( b \in \{0,1\} \) uniformly at random;
- (filter) change \( v \)'s state to \( b \) if the it gives an independent set;

\[
\sigma \in \{0,1\}^V
\]
\[
\forall \ uv \in E:
\quad \text{NOT } \sigma_u=\sigma_v=1
\]
Initially, start with an arbitrary proper $q$-coloring; at each step:

- *(proposal)* pick a vertex $v \in V$ and color $c \in [q]$ uniformly at random;
- *(filter)* change $v$’s color to $c$ if the it gives a proper coloring;

**proper $q$-coloring**

$\sigma \in [q]^V$

$\forall \ uv \in E: \ \sigma_u \neq \sigma_v$
Glauber Dynamics

**Input:** a CSP instance $I$
on $n$ variables with domain $[q]$;
Sample a uniform random CSP solution.

Glauber Dynamics:
Initially, start with an arbitrary CSP solution;at each step, the current CSP solution is $\sigma=(\sigma_1, \ldots, \sigma_n)$:

- pick a variable $i \in [n]$ uniformly at random;
- change value of $\sigma_i$ to a uniform value $c$ among all $\sigma_i$’s available values $c$: changing $\sigma_i$ to $c$ won’t violate any constraint;

detailed balanced equation $\rightarrow$ uniform stationary distribution!
Glauber Dynamics:

Initially, start with an arbitrary independent set; at each step:

- pick a uniform vertex \( v \in V \);
- change \( v \)'s state to a uniform random \( b \in \{0,1\} \) if all \( v \)'s neighbors have state 0;
Glauber Dynamics:
Initially, start with an arbitrary proper $q$-coloring; at each step:

- pick a uniform vertex $v \in V$;
- change $v$'s color to a uniform random color $c$ among $v$'s current available colors;