Randomized Algorithms

Lecture 15: Metropolis and Glauber

Raul Garcia-Patron

School of Informatics
University of Edinburgh

MCMC for Independent Sets (Recap)

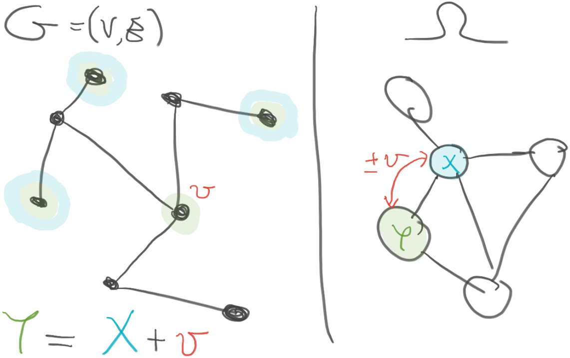
The IS Markov chain generates a uniform random sequence of ISs:

Algorithm GENERATEIS$(n; G = (V, E))$

1. Start with an arbitrary IS X_0
2. **for** $i \leftarrow 0$ **to** “whenever"
3. Choose v uniformly at random from V.
4. **if** $v \in X_i$ **then**
5. $X_{i+1} \leftarrow X_i \setminus \{v\}$
6. **elseif** ($v \notin X_i$ **and** $X_i \cup \{v\}$ is also an IS in G) **then**
7. $X_{i+1} \leftarrow X_i \cup \{v\}$
8. **else** $X_{i+1} \leftarrow X_i$

Discussion of update

1. The difference between X_i and X_{i+1} is in at most 1 vertex.
2. Update need to be a valid IS.
3. Probability of jumping between two IS is either zero or $1/|V|$.



Metropolis equivalent form

- $\{N(X)|x \in \Omega\}$ is the set of ISs that differ from x in a single vertex.
- Assume $X + v = Y$ We have $P_{x,y} = P_{y,x} = 1/|V|$:
  - $P_{x,y}$: we jump from X to Y only if vertex v is selected (probability $1/|V|$) followed by the algorithm deterministically adding v to X (line 7).
  - $P_{y,x}$: we jump from Y to X only if vertex v is selected (probability $1/|V|$) followed by the algorithm deterministically removing v to Y (line 5).

$$P_{x,y} = \begin{cases} 1/|V| & \text{if } x \neq y \text{ and} y \in N(x). \\ 0 & \text{if } x \neq y \text{ and} y \notin N(x). \\ 1 - |N(x)|/|V| & \text{if } x = y \end{cases} \quad (1)$$

Metropolis for uniform sampling

For a finite state space Ω and neighborhood structure $\{N(X)|x \in \Omega\}$, let $N = \max_{x \in \Omega} |N(x)|$. Let M be any number such that $M \geq N$. Consider a Markov chain where

$$P_{x,y} = \begin{cases} 1/M & \text{if } x \neq y \text{ and} y \in N(x). \\ 0 & \text{if } x \neq y \text{ and} y \notin N(x). \\ 1 - |N(x)|/M & \text{if } x = y \end{cases} \quad (2)$$

If this chain is **irreducible** and **aperiodic**, then the stationary distribution is the uniform distribution.

- For $x \neq y$ we have $P_{x,y} = P_{y,x} = 1/M$, implies $\pi_x = \pi_y = 1/|\Omega|$.
- Previous IS example we had $M = V$.

The Metropolis Algorithm

We may want to sample from a nonuniform distribution.

For a finite state space Ω and neighborhood structure $\{N(X)|x \in \Omega\}$, let $N = \max_{x\in\Omega} |N(x)|$. Let M be any number such that $M \geq N$. For all $x \in \Omega$, let $\pi_x > 0$ be the desired probability of state x in the stationary distribution. Consider a Markov chain where

$$P_{x,y} = \begin{cases} 1/M \min(1, \pi_y/\pi_x) & \text{if } x \neq y \text{ and} y \in N(x). \\ 0 & \text{if } x \neq y \text{ and} y \notin N(x). \\ 1 - \sum_{y\neq x} P_{x,y} & \text{if } x = y \quad (\textit{rejection}) \end{cases} \tag{3}$$

MC **finite**, **irreducible** and **aperiodic**: π stationary distribution.

- We do not need to know π_x or π_y, but only their ratio π_x/π_y!

The Metropolis Algorithm: proof by detail balance

$$P_{x,y} = \begin{cases} 1/M \min(1, \pi_y/\pi_x) & \text{if } x \neq y \text{ and} y \in N(x). \\ 0 & \text{if } x \neq y \text{ and} y \notin N(x). \\ 1 - \sum_{y \neq x} P_{x,y} & \text{if } x = y \end{cases} \quad (4)$$

Proof.

We want to prove detail balance. For $x \neq y$:

- If $\pi_x \leq \pi_y : P_{x,y} = 1/M$ and $P_{y,x} = \pi_x/(\pi_y M)$
  $\Rightarrow \pi_x P_{x,y} = \frac{\pi_x}{M} = \pi_y \frac{\pi_x}{\pi_y M} = \pi_y P_{y,x}$.
- If $\pi_x > \pi_y$: very similar proof.

□

Example: Independent sets (also Hardcore model)

We want to sample from independent sets x with probability

$$\pi_x = \lambda^{|x|}/Z, \tag{5}$$

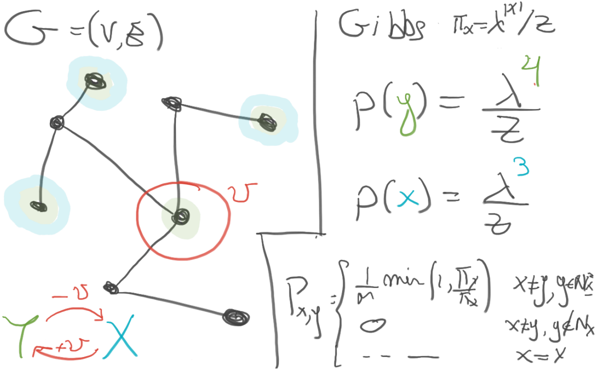
i.e., its Gibbs distribution.

- $|x|$ is the size of the IS x.
- $\lambda > 0$ a constant parameter
- $Z = \sum_x \lambda^{|x|}$ is a normalization constant.
  Hard to compute, but we do not need it to sample!

The value of λ bias the distribution:

- $\lambda = 1$, uniform distribution.
- $\lambda > 1$ larger IS have larger probability.
- $\lambda < 1$ smaller IS have larger probability.

Metropolis for independent sets (Gibbs distribution)



Metropolis for Hardcore model II

Algorithm METROPOLISIS$(n; G = (V, E))$

1. Start with an arbitrary IS X_0
2. **for** $i \leftarrow 0$ **to** "whenever"
3. Choose v uniformly at random from V.
4. **if** $v \in X_i$ **then**
5. $X_{i+1} \leftarrow X_i \setminus \{v\}$ with probability ?
6. **elseif** ($v \notin X_i$ **and** $X_i \cup \{v\}$ is also an IS in G) **then**
7. $X_{i+1} \leftarrow X_i \cup \{v\}$ with probability ?
8. **else** $X_{i+1} \leftarrow X_i$

- Remember: $\pi_x = \lambda^{|x|}/Z$.
-

$$P_{x,y} = \begin{cases} 1/M \min(1, \pi_y/\pi_x) & \text{if } x \neq y \text{ and} y \in N(x). \\ 0 & \text{if } x \neq y \text{ and} y \notin N(x). \\ 1 - \sum_{y \neq x} P_{x,y} & \text{if } x = y \end{cases} \quad (6)$$

Metropolis for Hardcore model III

Algorithm METROPOLISIS$(n; G = (V, E))$

1. Start with an arbitrary IS X_0
2. **for** $i \leftarrow 0$ **to** “whenever"
3. Choose v uniformly at random from V.
4. **if** $v \in X_i$ **then**
5. $X_{i+1} \leftarrow X_i \setminus \{v\}$ with probability $\min(1, 1/\lambda)$
6. **elseif** ($v \notin X_i$ **and** $X_i \cup \{v\}$ is also an IS in G) **then**
7. $X_{i+1} \leftarrow X_i \cup \{v\}$ with probability $\min(1, \lambda)$
8. **else** $X_{i+1} \leftarrow X_i$

- Remark $M = |V|$ again.
- $\frac{\pi_y}{\pi_x} = \frac{\lambda^{|y|}}{\lambda^{|x|}}$
- It is crucial that at any moment we need to compute $Z = \sum_x \lambda^{|x|}$.

Glauber Dynamics

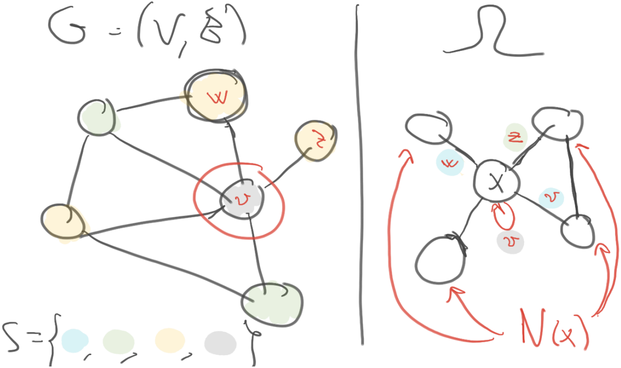
- Let V and S be finite sets and suppose $\Omega \subseteq S^V$.
  Ex: V vertices of a graph and S a set of colors (graph coloring).
- Let π be a probability distribution whose support is Ω.
- The Glauber chain moves from state x as follows:
  1. An element v is chosen uniformly at random from V.
  2. A new state y is chosen s.t.: $y(w) = x(w)\, \forall\, w \neq v$.

Definition

Given $x \in \Omega, v \in V : \Omega(x, v) = \{y \in \Omega : y(w) = x(w)\, \forall\, w \neq v\}$. The chain transition reads:

$$P_{x,y} = \pi(y|\Omega(x, v)) = \begin{cases} \frac{\pi(y)}{\pi(\Omega(x,v))} & \text{if } y \in \Omega(x, v) \\ 0 & \text{if } y \notin \Omega(x, v) \end{cases} \tag{7}$$

Glauber Dynamics for graph coloring



Glauber Dynamics for uniform independent sets

Algorithm GLAUBERIS$(n; G = (V, E))$

1. Start with an arbitrary IS X_0
2. **for** $i \leftarrow 0$ **to** "whenever"
3. Choose v uniformly at random from V.
4. Set $X_{t+1}(w) = X_t(w) \quad \forall w \neq v$
5. **if** $\exists w' \in N(v)$ such that $X_t(w') = 1$ **then**
6. $X_{t+1}(v) = 0$
7. **else** Generate a random bit b
8. **if** $b = 1$ **then** $X_{t+1}(v) = 1$
9. **else** $X_{t+1}(v) = 0$.

- When $y(v) = 1$ leads to a valid IS, also does $y(v) = 0$. Uniform distribution $\Rightarrow p(y(v) = 0) = p(y(v) = 1) = 1/2$.
- If $y(v) = 1$ not valid: then $y(v) = 0$ always is an IS (x is).

Comparing Metropolis and Glauber Dynamics.

- Consider the following Metropolis algorithm for uniform sampling of IS:
  1. Pick one vertex v uniformly at random.
  2. $\forall w \neq v : y(w) = x(w)$
  3. Select $y(v) = 0$ or $y(v) = 1$ with probability 1/2.
  4. If $y \notin \Omega$ reject it.
- In this scenario $M = |V|$ and $\pi_y/\pi_x \in \{0, 1\}$.
- Metropolis is equivalent to previous Glauber:
  - Whenever a neighbors of v has value 1 we can predict with certainty that the new value in v will be $y(v) = 0$ whether there is rejection or not ($x(v) = 0$ due to neighbors).
  - If all neighbors are 0s, there will be no rejection.
- We will see in the tutorial an example where Metropolis and Glauber for same distribution π lead to different Markov chains (graph coloring).