MATH 285: Stochastic Processes

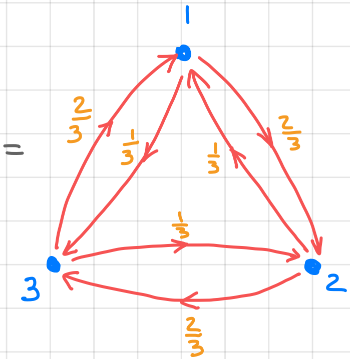
math-old.ucsd.edu/~ynemish/teaching/285

Today: MCMC

- Homework 3 is due on Friday, February 4, 11:59 PM

Example

Consider random walk on $G =$



Transition matrix

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

P is doubly stochastic, so

$$\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

Detailed balance equation:

$$p(j,i) = \frac{\pi(j)}{\pi(i)} p(j,i) = p(i,j) \Rightarrow \text{not reversible}$$

If $\pi = \left(\frac{1}{|S|}, \dots, \frac{1}{|S|}\right)$, (X_n) is reversible only if $P = P^t$

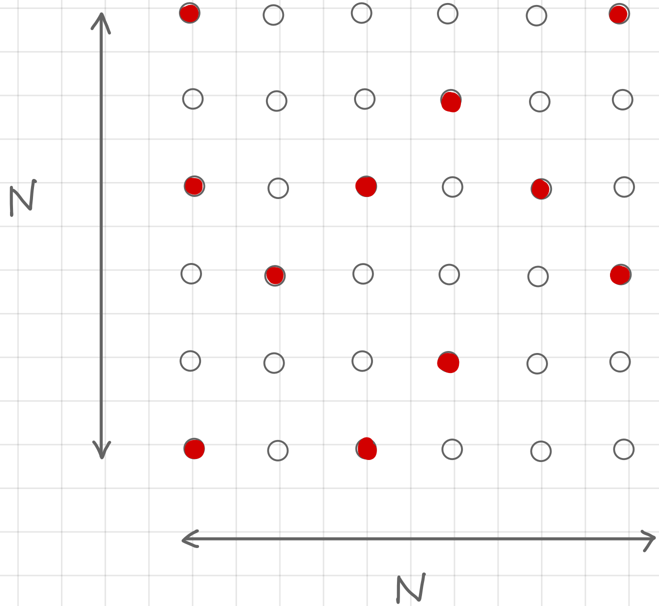
Example: Hard Core Configuration

Hard Core Configuration on $\{1,2,\ldots,N\}^2$ is a function

$$c : \{1,\ldots,N\}^2 \to \{0,1\}$$

such that

$$c(i,j) = 1 \Rightarrow c(i\pm 1, j\pm 1) = 0$$



Denote by HCC_N the set of all hard core configurations on $\{1,\ldots,N\}^2$. Suppose we want to choose a uniform distribution on HCC_N, $\mathbb{P}[Z = c] = \frac{1}{|HCC_N|}$ $\forall\, c \in HCC_N$

Problem: How to compute $|HCC_N|$?

Example: Hard Core Configuration

Computing $|HCC_N|$ for large N is difficult.

Instead we construct a MC on HCC_N whose stationary distribution is the uniform distribution on HCC_N.

Construction: for any two configurations $c \neq c'$

$$p(c, c') = \begin{cases} \frac{1}{N^2} & \text{if } c \text{ and } c' \text{ differ at exactly one point} \\ 0, & \text{otherwise} \end{cases}$$

$$p(c, c) = 1 - \sum_{c' \neq c} p(c, c')$$

Implementation: at each step choose $(i,j) \in \{1, \dots, N\}^2$ uniformly at random and change the value at (i,j) if possible. E.g., $X_n = c$, choose (i,j).

- If $c(i,j) = 1$, then $X_{n+1} = c'$ with $c'(i,j) = 0$, $c'(k,\ell) = c(k,\ell)$

Example: Hard Core Configuration

- If $c(i,j)=0$, and $c(i\pm1, j\pm1)=0$, then $X_{n+1}=c'$ with $c'(i,j)=1$ and $c'(k,l)=c(k,l)$
- If $c(i,j)=0$ and one of $c(i\pm1, j\pm1)\neq 0$, then $X_{n+1}=c$

(i) Then for any $c, c' \in HCC_N$ $\mathbb{P}[X_{n+1}=c' \mid X_n=c] = p(c,c')$

(ii) (X_n) is irreducible $(\forall\, c, c' \in HCC_N \;\; p_{n_0}(c,0)>0,\ p_{n_1}(0,c')>0)$

(iii) $p(c,c')=p(c',c)$ (c and c' differ in only one coordinate)

(iv) Uniform distribution on HCC_N is the stationary distribution

$$\pi(c) = 1/|HCC|_N \;\Rightarrow\; \pi(c)\,p(c,c') = \pi(c')\,p(c',c)$$

$$\Rightarrow \pi \text{ is stationary}$$

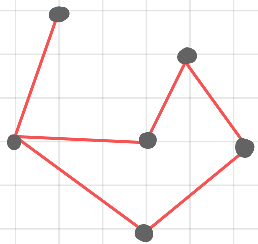
Now if we start the process from any $c \in HCC_N$, then for sufficiently large n $\mathbb{P}[X_n=c] \approx \dfrac{1}{|HCC_N|}$

Example: Graph coloring

Let $G=(V,E)$ be a finite graph. A q-coloring of G (with $q\in\mathbb{N}$) is a function $f: V\to\{1,2,\dots,q\}$ s.t.

$$u\sim v \Rightarrow f(u)\neq f(v)$$

(different colors of neighboring vertices)



Q: How to choose a q-coloring uniformly at random?

Construct a MC: if f and g are two q-colorings, $f\neq g$, set

$$p(f,g)=\begin{cases}\frac{1}{q|V|} & \text{if } f \text{ and } g \text{ differ at exactly one vertex}\\ 0 & \text{otherwise}\end{cases}$$

$$p(f,f)=1-\sum_{g\neq f}p(f,g)$$

(X_n) with transition probabilities $p(f,g)$ is an irreducible MC with stationary distribution $\pi(f)=1/\#\{q\text{-coloring of } G\}$

Metropolis - Hastings Algorithm

Q: How to sample any (strictly positive) distribution π ?

Two-step MC: (1) propose moves (2) accept/reject move

Construction of the Markov Chain

Let S be a finite set, $\pi > 0$ a distribution on S.

(1) Construct an irreducible MC on S with symmetric transition probabilities $q(i,j) = q(j,i)$, $q(i,i) = 1 - \sum_{j \neq i} q(i,j)$

(2) If π (the desired distribution) is not uniform, construct a new MC with transition probabilities

$$p(i,j) = q(i,j) \min\left\{\frac{\pi(j)}{\pi(i)}, 1\right\}$$

- $\pi(i)\, p(i,j) = q(i,j) \min(\pi(j), \pi(i)) = q(j,i) \min\{\pi(i), \pi(j)\}$

  so π is stationary for $p(i,j)$ $\qquad = \pi(j)\, p(j,i)$

Metropolis - Hastings Algorithm

Suppose we know how to simulate a MC with transition probabilities $q(i,j)$. The we can simulate a MC with transition probabilities $p(i,j)$ using the two-step algorithm:

(i) Propose the move:

If $X_n = i$, then for $j \neq i$ choose j with probability $q(i,j)$

(ii) Accept or reject the move:

Accept the move with probability $\min\{\frac{\pi(j)}{\pi(i)}, 1\}$

We get that $\mathbb{P}[X_{n+1} = j \mid X_n = i] = q(i,j)\min\{\frac{\pi(j)}{\pi(i)}, 1\} = p(i,j)$

If we now run (X_n) sufficiently long, then $\mathbb{P}[X_n = j] \approx \pi(j)$

Q: How long should we run (X_n)?

Convergence rate

Suppose that (X_n) is irreducible and aperiodic MC on S with $|S| = N$, and suppose that P is symmetric, $P = P^t$.

Then by the spectral theorem $P = U D U^t$

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix}, \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$$

Perron-Frobenius theorem $\Rightarrow$

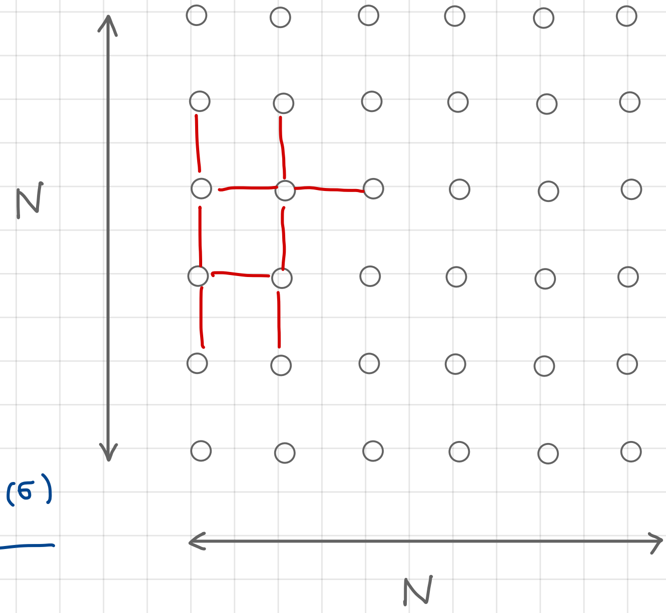
$$\underbrace{\lim_{n\to\infty} P^n}_{= P^\infty} = U \lim_{n\to\infty} D^n U^t = U \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} U^t$$

Then $\| P^n - P^\infty \| = \max_{j \geq 2} |\lambda_j|^n$. Mixing time: n s.t. $\| P^n - P^\infty \|$ is small

E.g. for $\begin{bmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{bmatrix}$ $\lambda_1 = 1$, $\lambda_2 = 1 - 2\varepsilon$ $\to (1-2\varepsilon)^n$ slow mixing

for $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ $\lambda_1 = 1$, $\lambda_2 = 0$ $\to \max_{j \geq 2} |\lambda_j|^n = 0$ fast mixing

Example: Ising model



- $\Lambda_N = \{1, \dots, N\}^2$
- Spin configuration:

  $$\sigma : \Lambda_N \to \{-1, 1\}$$

- Energy: $H(\sigma) = -\sum_{\substack{i \sim j \\ i,j \in \Lambda}} \sigma(i)\sigma(j)$
- Gibbs measure: $P_\beta(\sigma) = \dfrac{e^{-\beta H(\sigma)}}{Z_\beta}$

  where $Z_\beta = \sum_\sigma e^{-\beta H(\sigma)}$ is the partition function (difficult)

- Take $\pi(\sigma) = P_\beta(\sigma)$. Then $\pi(\sigma)/\pi(\sigma') = \exp(-\beta(H(\sigma) - H(\sigma')))$
- For $\sigma \neq \sigma'$ take $q(\sigma, \sigma') = \begin{cases} \frac{1}{N^2} & \text{if } \|\sigma - \sigma'\| = 2 \\ 0 & \text{otherwise} \end{cases}$
- Run MC (X_n) with $p(\sigma, \sigma') = q(\sigma, \sigma') \min\{1, \frac{\pi(\sigma')}{\pi(\sigma)}\}$