6. Random walk

Sunghyeon Jo

Seoul National University

Project-Agao
Mar 30, 2023

Outline

- Random walk: how to find stable distribution?
  - efficient algorithm via lazy random walk
- Hitting time, Commute time
  - Convert to Laplacian Linear Equation
  - Interpret to Electric Flow

Random Walk

Definition

Given undirected graph $G = (V, E)$.
Random sequence of vertices $v_0, v_1, \cdots$, are called random walk on G, if v_0 is a vertex in G chosen according to some distribution $p_0 \in \mathbb{R}^V$ and $\mathbb{P}[v_{t+1} = v | v_t = u] = w(u, v)/d(u)$ for $(u, v) \in E$

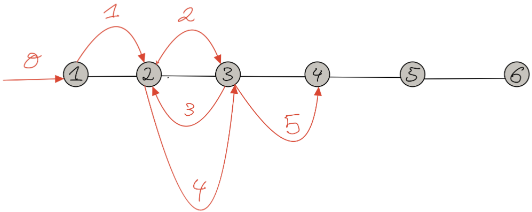

Figure 6.1: A (possibly random) walk where the red edges indicate the edges that the particle moves along. Here the walk visits the vertices $v_0 = 1, v_1 = 2, v_2 = 3, v_3 = 2, v_4 = 3, v_5 = 4$.

Random Walk Matrix

$W = AD^{-1}$ is called Random Walk Matrix.

Let p_t be distribution of v_t. Then, $p_t = W^t p_0$ holds.

Stationary Distribution

Definition

A probability distribution $\pi \in \mathbb{R}^V$ is called a stationary distribution if $W\pi = \pi$, which means π is an eigenvector of W associated with eigenvalue 1.

Stationary Distribution

Theorem

Every graph G has a stationary distribution.

Let $\pi = \frac{d}{1^T d}$ where d is diagonal of the degree matrix D.

Clearly, π is a probability distribution.

Also, $W\pi = AD^{-1}\frac{d}{1^T d} = \frac{A1}{1^T d} = \frac{d}{1^T d} = \pi$ holds.

Stationary Distribution

Question

For every distribution p_0, is p_t always converges?

Answer: No. (Consider a graph with 2 vertices connected by 1 edge)

Question

Can we create a graph with same stationary distribution that always converges (fast)?

Answer: Yes.

Lazy Random Walks

Let $\tilde{W} = \frac{1}{2}(I + W) = \frac{1}{2}(I + AD^{-1})$.

It is a Random Walk Matrix of the graph such that adding $\frac{1}{2}$ probability of staying current vertex.

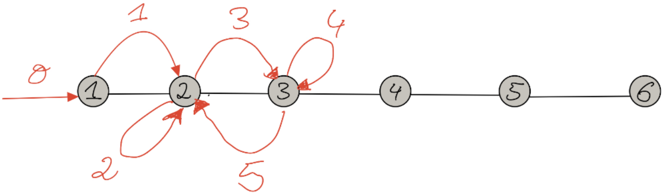

Figure 6.2: A lazy random walk where the red edges indicate the edges that the particle moves along. Here the lazy walk visits the vertices $v_0 = 1, v_1 = 2, v_2 = 2, v_3 = 3, v_4 = 3, v_5 = 2$.

Lazy Random Walks and Normalized Laplacian

Recall that $N = D^{-1/2}LD^{-1/2} = I - D^{-1/2}AD^{-1/2}$,
so $D^{-1/2}AD^{-1/2} = I - N$.

$$\tilde{W} = \frac{1}{2}I + \frac{1}{2}AD^{-1} \tag{1}$$
$$= \frac{1}{2}I + \frac{1}{2}D^{1/2}(I - N)D^{-1/2} \tag{2}$$
$$= I - \frac{1}{2}D^{1/2}ND^{-1/2} \tag{3}$$

we will see relation of eigenvalues and eigenvectors of $\tilde{W}$ and the same of N.

Lazy Random Walks and Normalized Laplacian

Let $\nu_1 \leq \nu_2 \leq \cdots \leq \nu_n$ be eigenvalues of N with their orthogonal eigenvectors $\psi_1, \psi_2, \cdots \psi_n$.

Remark(Spectral Theorem): If A is Hermitian on V, then there exists an orthonormal basis of V consisting of eigenvectors of A. Each eigenvalue is real.

From last lecture, we have $\psi_1 = \frac{d^{1/2}}{(1^T d)^{1/2}}$
Remark: N is PSD since $v^T N v = v^T D^{-1/2} L D^{-1/2} v \geq 0$,
$(d^{1/2})^T N (d^{1/2}) = (d^{1/2})^T D^{-1/2} L D^{-1/2} (d^{1/2}) = 1^T L 1 = 0$.
Therefore, $\psi_1 = \alpha d^{1/2}$ with norm 1 so $\psi_1 = \frac{d^{1/2}}{(1^T d)^{1/2}}$.

Lazy Random Walks and Normalized Laplacian

Theorem

$D^{1/2}\psi_i$ is $\tilde{W}$'s eigenvector associated with eigenvalue $1 - \nu_i/2$.

Exercise

Prove above theorem (straightforward calculation)

Lazy Random Walks and Normalized Laplacian

Theorem

$D^{1/2}\psi_i$ is $\tilde{W}$'s eigenvector associated with eigenvalue $1 - \nu_i/2$.

Exercise

Prove above theorem (straightforward calculation)

Corollary

Every eigenvalue of $\tilde{W}$ is in $[0, 1]$.

Since $0 \leq A + D$, $L \leq L + A + D = 2D$ so $N \leq 2I$. It implies N's eigenvalue is in $[0, 2]$ and the corollary follows.

Convergence of Lazy Random Walks

Idea: want to use decomposition of orthogonal eigenvector, but $\tilde{W}$ is not a symmetric matrix so instead use N's.

$$D^{-1/2}p_0 = \sum_{i=1}^{n} \alpha_i \psi_i \Leftrightarrow p_0 = \sum_{i=1}^{n} \alpha_i D^{1/2} \psi_i$$

Convergence of Lazy Random Walks

Idea: want to use decomposition of orthogonal eigenvector, but $\tilde{W}$ is not a symmetric matrix so instead use N's.

$$D^{-1/2}p_0 = \sum_{i=1}^{n} \alpha_i \psi_i \Leftrightarrow p_0 = \sum_{i=1}^{n} \alpha_i D^{1/2} \psi_i$$

Since $D^{1/2}\psi_i$ is an eigenvector of $\tilde{W}$,

$$\tilde{W} p_0 = \sum_{i=1}^{n} \alpha_i \tilde{W} D^{1/2} \psi_i = \sum_{i=1}^{n} \alpha_i (1 - \nu_i/2) D^{1/2} \psi_i$$

Convergence of Lazy Random Walks

Idea: want to use decomposition of orthogonal eigenvector, but $\tilde{W}$ is not a symmetric matrix so instead use N's.

$$D^{-1/2}p_0 = \sum_{i=1}^{n} \alpha_i \psi_i \Leftrightarrow p_0 = \sum_{i=1}^{n} \alpha_i D^{1/2}\psi_i$$

Since $D^{1/2}\psi_i$ is an eigenvector of $\tilde{W}$,

$$\tilde{W}p_0 = \sum_{i=1}^{n} \alpha_i \tilde{W} D^{1/2}\psi_i = \sum_{i=1}^{n} \alpha_i (1 - \nu_i/2) D^{1/2}\psi_i$$

Distribution after time t is calculated as

$$p_t = \tilde{W}^t p_0 = \sum_{i=1}^{n} \alpha_i (1-\nu_i/2)^t D^{1/2}\psi_i = \alpha_1 D^{1/2}\psi_1 + \sum_{i=2}^{n} \alpha_i (1-\nu_i/2)^t D^{1/2}\psi_i$$

If the graph is connected, $\nu_2 > 0$ so $\lim_{t\to\infty} p_t = \alpha_1 D^{1/2}\psi_1 = \pi$

Convergence of Lazy Random Walks

- Why $\nu_2 > 0$ holds for connected graph?

Convergence of Lazy Random Walks

- Why $\nu_2 > 0$ holds for connected graph?
- Assume $\nu_2 = 0$. Then $0 = N\psi_2 = (D^{-1/2})^T L(D^{-1/2})\psi_2$
- Since each vertex's degree > 0, $L(D^{-1/2}\psi_2) = 0$
- WTS: dimension of L's nullspace is 1

Convergence of Lazy Random Walks

- WTS: dimension of L's nullspace is 1
- Let incident matrix B of size $|E| \times |V|$ s.t. for $e_i = (u_i, v_i)$ with weight w_i then $B_{i,u_i} = \sqrt{w_i}, B_{i,v_i} = -\sqrt{w_i}$. Each row has two nonzero items.
- $B^T B = L$ holds and $N(L) = N(B)$.
- if $Bx = 0$ then $\forall i, x_{u_i} = x_{v_i}$ holds.
- Therefore, dimension of $N(B)$ is at most the number of components of the graph.

Theorem

Number of components is equal to dimension of $N(L)$.

The Rate of Convergence

Theorem

$$\|\boldsymbol{p}_t - \boldsymbol{\pi}\|_\infty \leq e^{-\nu_2 \cdot t/2}\sqrt{n}$$

Therefore, the distribution converges exponentially fast. Since we can calculate $p_{t+1} = \tilde{W} p_t$ in $O(m)$ time, we found an approximation algorithm for stable distribution runs in $O(m\text{polylog}(n))$
Read textbook and prove above theorem.

Hitting time

Definition

Given graph $G = (V, E)$ and a vertex s. Hitting time h is a vector of size V such that h_v is expected time to first arrival at s when starting vertex $v_0 = v$.

Hitting time

Definition

Given graph $G = (V, E)$ and a vertex s. Hitting time h is a vector of size V such that h_v is expected time to first arrival at s when starting vertex $v_0 = v$.

$W = AD^{-1} \Rightarrow h_a = 1 + (W1_a)^T h$ holds for all $a \neq s$.
Therefore, $1 - \alpha \cdot 1_s = (I - W^T)h$.

Hitting time

Assume that G is connected. From $1 - \alpha \cdot 1_s = (I - W^T)h$, $d - \alpha \cdot deg(s) \cdot 1_s = (D - A)h$ holds.

Define $b = d - \alpha \cdot deg(s) \cdot 1_s$ Then $Lh = b$ holds. α is determined uniquely as $||d||_1/deg(s)$.

Electric Networks

Electric network: $Lx = b$ when Demand for vertex u

$$b_u = \sum_{(u,v)\in E} \frac{x_u - x_v}{R_{uv}} = \sum_{(u,v)\in E} w(u,v)(x_u - x_v)$$

Hitting time : $b = d - ||d||_1 \cdot 1_s$ and solve $Lx = h$.

It is same to electrical flow problem with Each vertex v have demand d_v and s have demand $d_s - \sum d_v$ Hitting time is same to voltage difference $h(a) = x(a) - x(s)$

Commuting Time

Commuting time between a and b: $E[H_{a,b} + H_{b,a}]$ where $H_{a,b}$ is hitting time start from a hit to b. It is identical to electric flow problem that routes $||d||_1$ units of flow from b to a.

Effective Resistance

Voltage difference between a in b in an electrical flow routing demand $1_b - 1_a$ also called the effective resistance $R_{eff}(a, b)$

Thank you!