CS 598: Spectral Graph Theory. Lecture 9

Random Walks and Eigenvalues

Alexandra Kolla

Today

- Random walks on graphs review
- Normalized Laplacian, normalized Adjacency Matrix
- Matrix form of random walks, lazy random walk
- The stable distribution
- Convergence and the second eigenvalue
- Examples

Random Walks on Graphs

- G=(V,E,w) weighted undirected graph.
- Random walk on G starts on some vertex and moves to a neighbor with prob. proportional to the weight of the corresponding edge.
- We are interested in the probability distribution over vertices after a certain number of steps.

Random Walks on Graphs

- G=(V,E,w) weighted undirected graph.
- Let vector $p_t \in R^n$ denote the probability distribution at time t. We will also write $p_t \in R^V$, and $p_t(u)$ for the value at vertex u.
- Since it's a probability vector, $p_t(u) \geq 0$ and $\sum_u p_t(u) = 1$ for every t.
- Usually, we start our walk at one vertex, so $p_0(u) = 1$ for some vertex u and o for the rest.

Random Walks on Graphs

- To derive p_t from p_{t+1} note that the probability of being at node u at time t+1 is the sum over all neighbors v of u of the probability that the walk was on v at time t times the probability it moved from v to u in one step:

$$p_{t+1}(u) = \sum_{v:(u,v)\in E} \frac{w(u,v)}{d(v)} p_t(v)$$

Where $d(v)$=$\sum_u w(u,v)$ is the weighted degree of v.

Lazy Random Walks

- We will often consider lazy random walks, which are a variant where we stay put with probability ½ at each time step, and walk to a random neighbor the other half of the time.

$$p_{t+1}(u) = \frac{1}{2} p_t(u) + \frac{1}{2} \sum_{v:(u,v)\in E} \frac{w(u,v)}{d(v)} p_t(v)$$

- Lazy random walks closely related to diffusion processes (at each time step, some substances diffuses out of each vertex)

Normalized Adjacency Matrix

- Need to define normalized versions of Adjacency matrix and Laplacian.
- Normalized Adjacency matrix is what you would expect:

$$M_G = {D_G}^{-1/2} A_G {D_G}^{-1/2}$$

With eigenvalues $1 = \mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$

and first eigenvector $\sqrt{\mathbf{d}}$ (see blackboard)

Normalized Laplacian

- Normalized Laplacian is also what you would expect:

$$N_G = D_G{}^{-1/2} L_G D_G{}^{-1/2} = I - M_G$$
$$= I - D_G{}^{-1/2} A_G D_G{}^{-1/2}$$

With eigenvalues $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_n$

and first eigenvector $\sqrt{\mathbf{d}}$ as well

Matrix Form of Random Walk

- Best way to understand random walks is with linear algebra. Equation

$$p_{t+1}(u) = \frac{1}{2} p_t(u) + \frac{1}{2} \sum_{v:(u,v)\in E} \frac{w(u,v)}{d(v)} p_t(v)$$

Is equivalent to (verify on blackboard)

$$p_{t+1} = \frac{1}{2}(I + AD^{-1})\, p_t$$

The lazy r.w. matrix is:

$$W_G = 1/2(I + A_G {D_G}^{-1})$$

Matrix Form of Random Walk

$$W_G = 1/2(I + A_G {D_G}^{-1})$$

- Is an a-symmetric matrix!! (the only one we will deal with in class). But it is closely related to normalized adjacency and Laplacian :

$$W_G = \frac{1}{2} {D_G}^{\frac{1}{2}}(I + M_G){D_G}^{-\frac{1}{2}} = I - 1/2{D_G}^{\frac{1}{2}} N_G {D_G}^{-\frac{1}{2}}$$

So W is diagonalizable and for every evector v or M

with evalue μ, ${D_G}^{\frac{1}{2}}$v is right-evector of W with evalue ½(1 + μ).

- For asymmetric matrices, evectors not necessarily orthogonal!

Why Lazy Random Walks?

$$W_G = 1/2(I + A_G D_G{}^{-1}) = \frac{1}{2} D_G{}^{\frac{1}{2}}(I + M_G) D_G{}^{-\frac{1}{2}}$$

- All evals of W are between 1 and 0: Perron evalue of M is 1, so M has evalues between 1 and -1.
- We let $1 = \omega_1 \geq \omega_2 \geq \cdots \geq \omega_n \geq 0$
- Where $\omega_i = 1/2(1 + \mu_i)$

The Stable Distribution

$$W_G = 1/2(I + A_G {D_G}^{-1}) = \frac{1}{2} {D_G}^{\frac{1}{2}} (I + M_G) {D_G}^{-\frac{1}{2}}$$

- Regardless of starting distribution, lazy r.w. always converges to stable distribution (I think Sariel called it irreducible, aperiodic MC).
- In stable distribution, every vertex is visited with probability proportional to its (weighted) degree.

$$\boldsymbol{\pi}(\mathrm{i}) = \frac{\boldsymbol{d}(i)}{\sum_j \boldsymbol{d}(j)}$$

The Stable Distribution

$$W_G = 1/2(I + A_G {D_G}^{-1}) = \frac{1}{2} {D_G}^{\frac{1}{2}} (I + M_G) {D_G}^{-\frac{1}{2}}$$

$$\boldsymbol{\pi}(\mathrm{i}) = \frac{\boldsymbol{d}(i)}{\sum_j \boldsymbol{d}(j)}$$

- **π** is right evector of W with evalue 1 (see blackboard).
- Other reason to consider lazy walks, is that they always converge. (e.g. consider bipartite graphs)

The Stable Distribution

$$W_G = 1/2(I + A_G D_G{}^{-1}) = \frac{1}{2} D_G{}^{\frac{1}{2}} (I + M_G) D_G{}^{-\frac{1}{2}}$$

$$\boldsymbol{\pi}(\mathrm{i}) = \frac{\boldsymbol{d}(i)}{\sum_j \boldsymbol{d}(j)}$$

- **π** is right evector of W with evalue 1 (see blackboard).

- Other reason to consider lazy walks, is that they always converge. (e.g. consider bipartite graphs)
- See blackboard for proof that lazy walk converges to **π**

Rate of Convergeance

$$W_G = 1/2(I + A_G {D_G}^{-1}) = \frac{1}{2} {D_G}^{\frac{1}{2}} (I + M_G) {D_G}^{-\frac{1}{2}}$$

$$\boldsymbol{\pi}(\mathrm{i}) = \frac{\boldsymbol{d}(i)}{\sum_j \boldsymbol{d}(j)}$$

- Rate of convergence to the stable distribution is dictated by the second eigenvalue of W.
- Assume that r.w. starts at some vertex a. Let χ_a the characteristic vector of a, which is our starting distribution. For every vertex b, we will bound how far $p_t(\mathrm{b})$ can be from $\boldsymbol{\pi}(\mathrm{b})$.

Rate of Convergeance

- Assume that r.w. starts at some vertex a. Let χ_a the characteristic vector of a, which is our starting distribution. For every vertex b, we will bound how far $p_t(\mathrm{b})$ can be from $\boldsymbol{\pi}(\mathrm{b})$:

- Theorem. For all a,b, if $p_0 = \chi_a$ then

$$|p_t(\mathrm{b}) - \boldsymbol{\pi}(\mathrm{b})| \leq \sqrt{\frac{d(b)}{d(a)}}\,{\omega_2}^t$$

How Many Steps to Converge?

- To have $|p_t(\mathrm{b}) - \boldsymbol{\pi}(\mathrm{b})| \leq \varepsilon$, we need t to be such that $\sqrt{\frac{d(b)}{d(a)}}\omega_2{}^t \leq \varepsilon$.
- Define $\omega_2 = 1 - \gamma$, where γ is the spectral gap between first and second eigenvalue(remember discussion about expansion and large spectral gap).
- Number of steps to convergeance depends on 1/ γ , use $1 - \gamma \leq e^{-\gamma}$ (blackboard).

How Many Steps to Converge?

Path graph, tree graph…