CS 0368-4246: Combinatorial Methods in Algorithms (Spring 2025) 16/06/2025, 2025

Lecture 11: Graph Random Walks

Instructor: *Or Zamir* Scribes: *Nir Shahar*

1 Introduction

We have already seen in previous courses the concept of a random walk on a "straight line" (i.e, the number line), and now we want to generalize the concept to general graphs.

Given an undirected graph G and an initial vertex v_0, we can go to a random neighboring vertex of it in each "step". This will create the series of steps $v_0, v_1, \ldots, v_t$. This concept is also known as a "markov chain", where we have states and move between them randomly.

Problem: what is the distribution of v_t for "large enough" t? Does it even converge? We will see under very basic assumptions, this distribution will indeed converge very quickly to the same distribution - and it won't even matter from which source v_0 vertex we begin! Intuitively reasonable assumptions (for this convergence) might be:

1. Connected - so we won't get "stuck" in one aisle
2. Not bipartite - so we won't "ping-pong" between two states and never converge

Claim 1. *Our goal is to show that these assumptions are enough for $D(v_t)$ to converge (quickly), and it will converge to $\frac{1}{2m}\vec{d}$ (the normalized degree vector).*

Proof. Let D_0 be the distribution of v_0, which is e_{v_0} (1 in the index of v_0 and 0 everywhere else. We could have also worked with any other distribution). Now, consider the t's step. A sends a "1" in a vertex v, to a "1" in all of its neighbors. Instead, we want it to send us into a new valid distribution, which we can achieve by normalizing it with D^{-1}. Now, AD^{-1} is the transition matrix that will take us a step forward in time:

$$D_t = (AD^{-1})D_{t-1}$$

We raise the question: what is the spectrum of AD^{-1}? We noticed last week that this matrix is similar (in the linear algebra sense) to $D^{-\frac{1}{2}}AD^{-1}D^{\frac{1}{2}} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ and we have defined $N_G := I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$. Denote the eigenvalues of AD^{-1} by $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$, and N_G's by $\lambda_1 \leq \cdots \leq \lambda_n$. Notice that if v is an eigenvector of some B with eigenvalue λ, then it is also an eigenvector of $I - B$ with eigenvalue $1 - \lambda$. Thus, $\lambda_i = 1 - \mu_i$. Last week, we have proven that:

- always $\lambda_1 = 0$ (implies $\mu_1 = 1$)
- $\lambda_2 > 0 \iff G$ is connected ($\iff \mu_2 < 1$)

Question:

1. Prove $\mu_n \geq -1$

2. Prove $\mu_n = -1 \iff G$ is bipartite

Denote by $\mu^* := \max(|\mu_2|, |\mu_n|)$, from this homework question, we know $\mu^* \leq 1$ and $\mu^* < 1 \iff G$ is connected and not bipartite. We now claim that if G is connected and not bipartite ($\mu^* < 1$) then D_t converges to $\frac{1}{2m}\vec{d}$ (when $t \to \infty$). First, notice that $\frac{1}{2m}\vec{d}$ is an eigenvector of AD^{-1} with eigenvalue 1 (so it is a stationary distribution - it does not change when stepping to the next t). Noting $D_t = (AD^{-1})^t D_0$, we will write $D_0 = (p_1, \ldots, p_n)$ in the unitary eigenvector basis $\{v_1, \ldots, v_n\}$ for our eigenvalues $\mu_1, \ldots, \mu_n$:

$$D_0 = \sum_{i=1}^{n} \alpha_i v_i$$

Now, what happens after t steps?

$$\begin{array}{rcl} D_t & = & \\ (AD^{-1})^t D_0 & = & \\ (AD^{-1})^t \left(\sum \alpha_i v_i\right) = \sum \alpha_i (AD^{-1})^t v_i & = & \\ \sum \alpha_i \mu_i^t v_i & = & (\mu_1 = 1) \\ \alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i \underbrace{\mu_i^t}_{|\mu_i|<1} v_i & & \end{array}$$

And since $|\mu_i^t| \leq \mu^* < 1$ (from our assumption), then $\mu_i^t \underset{t\to\infty}{\to} 0$ hence the limit must be:

$$D_t \underset{t\to\infty}{\to} \alpha_1 v_1$$

But what is $\alpha_1 v_1$? We know $v_1 = \frac{1}{2m}\vec{d}$ since its the eigenvector for eigenvalue $\mu_1 = 1$, and we must have $\alpha_1 = 1$ because every D_t in the process is a distribution, hence at the limit it must be a distribution as well (and $v_0 = \frac{1}{2m}\vec{d}$ is a distribution already). Therefore,

$$D_t \underset{t\to\infty}{\to} \frac{1}{2m}\vec{d}$$

□

We are also interested in knowing how quickly D_t converges

Claim 2. *For any ϵ, we have $|D_t - \frac{1}{2m}\vec{d}| < \epsilon$ when $t = O\left(\frac{\log(n/\epsilon)}{1-\mu^*}\right)$*

Proof.

$$\left|D_t - \frac{1}{2m}\vec{d}\right| = \left|\sum_{i=2}^{n} \alpha_i \mu_i^t v_i\right| \leq \sum_{i=2}^{n} |\alpha_i \mu_i^t v_i| \leq (\mu^*)^t \sum_{i=2}^{n} |\alpha_i| \cdot |v_i| \leq (\mu^*)^t \sum_{i=2}^{n} n \leq n^2 (\mu^*)^t$$

Its enough to require $n^2(\mu^*)^t < \epsilon$, which happens when $t = \Theta\left(\frac{\log(n/\epsilon)}{1-\mu^*}\right)$ because $(\mu^*)^{\frac{1}{1-\mu^*}} \approx \frac{1}{e}$ □

Observation 3. *In the above proof, $1 - \mu^*$ is the spectral gap. As a reminder to the first lecture on expanders, we saw that the diameter of a φ-expander is at most $O\left(\frac{\log n}{\varphi}\right)$. Now we have proved something*

stronger: after a similar number of random steps, we are already in a "uniform (up to degree) random vertex" in the graph.

Therefore, in an expander graph, a random walk of logarithmic length will end in a random vertex

Observation 4. *What have we seen? We saw that if we start from some initial distribution D_0 and look at $(AD^{-1})D_0$ then we converge to v_1 - the eigenvector with the largest eigenvalue. Additionally, we have seen last week that from v_2 (the second largest eigenvector) we can get an approximation to the sparsest cut in $O(m + n \log n)$. Is it possible to approximate v_2 just as we did with v_1? Consider what happens when $\alpha_1 = 0$ but $\alpha_2 \neq 0$. In that case, the algorithm above would converge to v_2 (faster than the rest of the elements). But how can we find an initial vector with $\alpha_1 = 0$? Choose an initial vector $v \perp \vec{d}$ since $v_1 = \frac{1}{2m}\vec{d}$. To also get $\alpha_2 \neq 0$ its enough to choose a random $v \perp \vec{d}$ and with very high probability $\alpha_2 \neq 0$*

Sketch 5. *We can build an algorithm in time $\tilde{O}(m)$ to approximate the sparsest-cut problem:*

1. *Calculate $N_G = I - D^{-\frac{1}{2}}AD^{\frac{1}{2}}$ (takes $O(m)$)*
2. *Randomly select a nonzero v with $v \perp \vec{d}$ (takes $O(n)$)*
3. *Calculate $v' := (AD^{-1})^t v$ for $t = O(\log \frac{n}{\epsilon})$ (takes $O(t \cdot m)$), we can expect v' to be an approximation of v_2 (actually, the only thing we need is just $\frac{v^t N v}{v^t v}$)*
4. *On v', run the algorithm from last week that in time $O(m+n \log n)$ checks all the cuts that correspond to the coordinate ordering, and returns the best one it finds*

We got an algorithm in time $O(m \log \frac{n}{\epsilon})$ that returns a $\sqrt{2(\varphi + \epsilon)}$-approximation for sparsest cut

we won't see: Its also possible in $\varphi \log n$ rather than $\sqrt{\varphi}$, and it is also possible to find the approximate balanced cut. ***Question:*** *Complete the details (the above algorithm was just a sketch).*

Observation 6. *Its possible to use random walks and what we have seen using expanders to obtain samples from very complex distributions.*

Example 7. *We know that a graph G with maximal degree δ is $\delta + 1$ colorable. How is it possible to uniformly select a valid 2δ coloring for G?*

To do so, we will define a "Markov chain" / random walk on the valid colorings that starts from some valid initial coloring, and iteratively chooses a vertex randomly and chooses a new valid coloring for this vertex.

Consider the graph defined by all valid colorings, where two are connected if they are different only in a single vertex. Its enough to prove that this graph is an expander to complete the proof (we will not prove it here).

Observation 8. *What happens for our random walks if we consider directed graphs instead of undirected graphs? In directed graphs, if the stationary distribution (the distribution we converge towards) exists, then it is not necessarily "constant" for all graphs. Instead, this distribution tells us more information about the graph and roughly approximates the "importance" of the vertices in the graph.*

References