COMS 4995-8: Advanced Algorithms (Spring'21) Feb 25, 2021

Lecture 14: Random walks, largest eigenvalue

Instructor: *Alex Andoni* Scribes: *Maryam Bahrani, Miranda Christ*

1 Introduction

We apply spectral results from last class to characterize matrices corresponding to random walks in a graph. In particular, we prove that the largest eigenvalue λ_1 of a symmetric graph derived from the adjacency matrix must equal 1. We also show that the second eigenvalue $\lambda_2 = 1$ if and only if the graph is disconnected.

2 Random walks in a graph G

We apply concepts from last class (namely the Spectral Theorem and the Rayleigh quotient) to analyze random walks.

Let $A = A_G$ be an adjacency matrix for an undirected graph G on n vertices. Let $D = D_G$ be the diagonal matrix of degrees of vertices of G. Let X^0 be a starting distribution over $[n]$, and let X^t be the distribution at time t. That is, X_i^t is the probability that a random walk is at vertex i after t random steps.

As shown last class,

$$X^t = (A \cdot D^{-1})^t \cdot X^0$$

We wish to apply the Spectral Theorem here to somehow characterize X^t; however, we can only apply the theorem to a symmetric matrix. Letting $\hat{A} = D^{-1/2} \cdot A \cdot D^{-1/2}$ (note that $\hat{A}$ is symmetric), we have:

$$\begin{aligned} X^t &= A \cdot D^{-1} \cdot A \cdot D^{-1} \cdot \dots \cdot A \cdot D^{-1} \cdot X^0 \\ &= A \cdot D^{-1/2} \cdot D^{-1/2} \cdot A \cdot D^{-1/2} \cdot D^{-1/2} \cdot \dots \cdot D^{-1/2} A \cdot D^{-1/2} \cdot D^{-1/2} \cdot X^0 \\ &= D^{1/2} \cdot \hat{A}^t \cdot D^{-1/2} \cdot X^0 \end{aligned}$$

Letting $Y^t := D^{-1/2} X^t$, we have $Y^t = \hat{A}^t \cdot D^{-1/2} \cdot X^0 = \hat{A}^t \cdot Y^0$. Observe that $\hat{A} = D^{-1/2} \cdot A \cdot D^{-1/2}$ is symmetric as desired. We can now apply the Spectral Theorem to $\hat{A}$, letting us write

$$\hat{A} = \sum_i \lambda_i v_i v_i^T \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

where (λ_i, v_i) are eigenvalue/eigenvector pairs, $\|v_i\| = 1$ for all i, and $v_i \cdot v_j = 0$ for $i \neq j$.

Fact 1.

- $\hat{A}^t = \sum_{i=1}^n \lambda_i^t v_i v_i^T$ *since for all* $i \neq j$, v_i *and* v_j *are orthogonal.*
- $Y^0 = \sum_{i=1}^n \alpha_i v_i$, *where* $\alpha_i \in \mathbb{R}$ *and the* α_i*'s are uniquely determined.*

- $Y^t = \sum_{i=1}^n \lambda_i^t \alpha_i v_i$

Proof. We sketch a proof of the last fact, that $Y^t = \sum_{i=1}^n \lambda_i^t \alpha_i v_i$. We first compute Y^1:

$$\begin{aligned} Y^1 &= \hat{A} \cdot Y^0 \\ &= \sum_{i=1}^n \lambda_i v_i v_i^T \cdot \sum_{j=1}^n \alpha_j v_j \\ &= \sum_{i,j} \lambda_i \alpha_j v_i v_i^T v_j \\ &= \sum_i \lambda_i \alpha_i v_i \end{aligned}$$

where in the last step, we use the fact that for $i \neq j$, v_i and v_j are orthogonal, so $v_i^T v_j = 0$. By inducting on t and doing a similar calculation, again using the fact that v_i and v_j are orthogonal for $i \neq j$, it follows that $Y^t = \sum_i \lambda_i^t \alpha_i v_i$. $\square$

Observation 2.

- *If $\lambda_1 > 1$, then Y_t diverges as $t \to \infty$. Since $Y^t = D^{-1/2} X^t$, where X^t represents a probability distribution, this cannot happen. The same issue arises if $\lambda_1 < -1$. Thus $\lambda_i \in [-1, 1]$ for all i.*

- *If $\lambda_i \in (-1, 1)$, as $t \to \infty$, the value $\lambda_i^t \alpha_i v_i$ goes to zero.*

Theorem 3. *Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ be the eigenvalues of $\hat{A}$ chosen according to the spectral theorem. Then $\lambda_1 = 1$.*

Proof. We first show that $\lambda_1 \geq 1$. Recall that

$$\lambda_1 = \max_{x \neq 0} R(x)$$

where $R(x)$ is the Rayleigh quotient of x. Thus to show that $\lambda_1 \geq 1$, it is sufficient to find an x such that $x \neq 0$ and $R(x) \geq 1$. Let $x = (\sqrt{d_1}, \sqrt{d_2}, ..., \sqrt{d_n})^T = D^{1/2} \cdot \mathbb{1}$, where $\mathbb{1}$ denotes the vector of all ones. We will show that this choice of x indeed achieves $R(x) = 1$.

Observe that $A \cdot \mathbb{1}$ is a vector of the degrees of each vertex. We use this fact below.

$$\begin{aligned}
R(x) &= \frac{x^T \hat{A} x}{\|x\|^2} \\
&= \frac{\mathbb{1}^T D^{1/2} D^{-1/2} A D^{-1/2} D^{1/2} \mathbb{1}}{\sum_{i=1}^n x_i^2} \\
&= \frac{\mathbb{1}^T A \mathbb{1}}{\sum_{i=1}^n d_i} \\
&= \frac{\mathbb{1}^t \cdot [d_1, ..., d_n]^T}{\sum_{i=1}^n d_i} \\
&= \frac{\sum_{i=1}^n d_i}{\sum_{i=1}^n d_i} \\
&= 1
\end{aligned}$$

Therefore, $\lambda_1 \geq 1$. It now suffices to show $\lambda_1 \leq 1$, completing the proof.

We show that for all $x \neq 0$, $R(x) \leq 1$. Let $x \neq 0$.

$$\begin{aligned}
R(x) &= \frac{x^T \hat{A} x}{\|x\|^2} \\
&= \frac{x^T D^{-1/2} A D^{1/2} x}{\|x\|^2} \\
&= \frac{\left(\frac{x_1}{\sqrt{d_1}}, ..., \frac{x_n}{\sqrt{d_n}}\right) A \left(\frac{x_1}{\sqrt{d_1}}, ..., \frac{x_n}{\sqrt{d_n}}\right)^T}{\|x\|^2} \\
&= \frac{\sum_{i,j} A_{ij} \cdot \frac{x_i}{\sqrt{d_i}} \cdot \frac{x_j}{\sqrt{d_j}}}{\|x\|^2} \\
&= \frac{\sum_{(i,j) \in E} \frac{x_i}{\sqrt{d_i}} \cdot \frac{x_j}{\sqrt{d_j}}}{\|x\|^2}
\end{aligned}$$

Here we can apply the Cauchy-Schwarz inequality: for all vectors p, q, we have $p \cdot q = \|p\| \cdot \|q\| \cdot \cos(p, q) \leq \|p\| \cdot \|q\|$. We define $p, q \in \mathbb{R}^{2|E|}$, where for i, j such that $A_{ij} = 1$, we let $p_{ij} = \frac{x_i}{\sqrt{d_i}}$ and $q_{ij} = \frac{x_j}{\sqrt{d_j}}$. Thus

$$\begin{aligned}
\frac{\sum_{(i,j) \in E} \frac{x_i}{\sqrt{d_i}} \cdot \frac{x_j}{\sqrt{d_j}}}{\|x\|^2} &\leq \frac{\left(\sum_{(i,j) \in E} \left(\frac{x_i}{\sqrt{d_i}}\right)^2\right)^{1/2} \cdot \left(\sum_{(i,j) \in E} \left(\frac{x_j}{\sqrt{d_j}}\right)^2\right)^{1/2}}{\|x\|^2} \\
&= \frac{\sum_{(i,j) \in E} \frac{x_i^2}{d_i}}{\|x\|^2} \\
&= \frac{\sum_i x_i^2}{\sum_i x_i^2} \\
&= 1.
\end{aligned}$$

Thus $R(x) \leq 1$ for all $x \neq 0$, and consequently $\lambda_1 \leq 1$. Since we previously showed $\lambda_1 \geq 1$, we have that $\lambda_1 = 1$ as desired. $\square$

Remark 1. If λ_1 is unique, then its corresponding eigenvector $v_1 = (\sqrt{d_1}, \sqrt{d_2}, \ldots, \sqrt{d_n})$.

Remark 2. If $R(x) = 1$ for some vector x, then $\frac{x_i}{\sqrt{d_i}} = \frac{x_j}{\sqrt{d_j}}$ for every edge (i, j).

Proof. Consider x such that $R(x) = 1$. Consider the $\lambda_1 \leq 1$ part of the proof above. The only inequality there is due to Cauchy-Schwarz, which we can replace with the equality $p \cdot q = \|p\| \cdot \|q\|$, to conclude that $R(x) = \cos(p, q)$ for all x. Thus $R(x) = 1$ implies $\cos(p, q) = 1$, that is $p = \alpha q$ for some positive real number α. By definition of p and q, this implies that for every edge (i, j), we have

$$\frac{x_i}{\sqrt{d_i}} = \alpha \frac{x_j}{\sqrt{d_j}}$$

and by symmetry also

$$\frac{x_j}{\sqrt{d_j}} = \alpha \frac{x_i}{\sqrt{d_i}}.$$

For both these equalities to hold simultaneously, it must be the case that $\alpha = 1$. Therefore, $\frac{x_i}{\sqrt{d_i}} = \frac{x_j}{\sqrt{d_j}}$ for every edge (i, j) as desired. $\square$

Remark 3. $R(x) \geq -1$ for all x, with equality iff $\frac{x_i}{\sqrt{d_i}} = -\frac{x_j}{\sqrt{d_j}}$ for every edge i, j.

The following theorem connects the algebraic quantity λ_2 to the combinatorial property of graph connectivity. This is rather surprising as the two properties seem unrelated on the surface but is a common theme in the Spectral graph theory, as we shall see in future lectures.

Theorem 4. *$\lambda_2 = 1$ if and only if G is disconnected.*

Proof. We first show that if G is disconnected, then $\lambda_2 = 1$. If G is disconnected, it can be thought of as two non-empty graphs G_1, G_2 with no paths between them. Reordering the rows and columns of A_G to list the vertices of G_1 first and the vertices of G_2 second, we get the following block structure on A_G

$$\hat{A}_G = \begin{bmatrix} \hat{A}_{G_1} & 0 \\ 0 & \hat{A}_{G_2} \end{bmatrix}$$

Suppose G_1 has k vertices, and let (λ_i, v_i), $i = 1, 2, \ldots, k$ be the eigenvalues and eigenvectors of $\hat{A}_{G_1}$. Note that $v_i \in \mathbb{R}^k$. Furthermore, let (λ_{i+k}, v_{i+k}), $i = k+1, \ldots, n$ be the eigenvalues and eigenvectors of $\hat{A}_{G_2}$, where $v_{i+k} \in \mathbb{R}^{n-k}$. Then the eigenvalues and eigenvectors of $\hat{A}_G$ are (λ_i, v_i'), $i = 1, \ldots, n$, where

$$v_i' = \begin{cases} (v_i, 0, \ldots, 0) & \text{if } i \leq k \\ (0, \ldots, 0, v_i) & \text{if } i > k \end{cases}$$

By Theorem 3, the first eigenvalue of both $\hat{A}_{G_1}$ and $\hat{A}_{G_2}$ is equal to 1. So in particular, $\lambda_1 = 1, v_1' = (v_1, 0, \ldots, 0)$ and $\lambda_{k+1} = 1, v_{k+1}' = (0, \ldots, 0, v_{k+1})$ are both eigenvalue-eigenvector pairs of $\hat{A}_G$, where v_1' and v_{k+1}' are orthogonal, and thus $\lambda_2 = 1$.

We now show the other direction by contradiction. Suppose G is connected and $\lambda_2 = 1$. Let v_2 be an eigenvector with eigenvalue λ_2. We know that $v_2 \perp v_1$, so $R(v_2) = \lambda_2 = 1$. By Remark 2, this implies that for every edge ij,

$$\frac{x_i}{\sqrt{d_i}} = \frac{x_j}{\sqrt{d_j}}.$$

Let $\beta = \frac{x_1}{\sqrt{d_1}}$. Since G is connected, there is a path from vertex 1 to every other vertex i, and the chain of equalities along each edge of this path implies $\frac{x_i}{\sqrt{d_i}} = \frac{x_1}{\sqrt{d_1}} = \beta$ for all vertices i. Put differently, $x_i = \sqrt{d_i} \cdot \beta$ for every $i \in [n]$. Therefore, $x = v_1 \cdot \beta$, since $v_1 = (\sqrt{d_1}, \ldots, \sqrt{d_n})$ by Remark 1. This contradicts the assumption that $v_2 \perp v_1$, completing the proof. $\square$