CS 229r Spectral Graph Theory in Computer Science, Lecture 1 Sep 3, 2020

Lecture 1

Instructor: Salil Vadhan *Scribes: Pranay Tankala, Franklyn Wang*

1 Overview

Today, we're going to introduce the course's core concepts—weighted graphs and their associated matrices—in greater generality than Chapter 1 of Spielman's book [Spi19]. We will discuss the relative merits of the many different matrices that can be used to describe the same graph (a question that frequently arose on Perusall). Finally, we will review some fundamental concepts from linear algebra and walk through the proof of the Spectral Theorem from Chapter 2 of Spielman's book.

2 Logistics

- Zoom etiquette: include pronouns in your Zoom name, mute yourself when not talking, keep your video on, don't record other students or share class recordings. In case Zoom goes down, check Piazza for instructions.
- Teaching Fellows: Chi-Ning Chou, Alec Sun, Santhoshini Velusamy, + TBD
- Problem Set 0 is due **Friday, September 11$^{\text{th}}$**.
- Professor Vadhan will not hold his regular office hour on Monday, September 7$^{\text{th}}$ (Labor Day). There will be a poll regarding TF section and office hour times on Piazza.
- Let the course staff know if you need help finding other students in the class willing to discuss the course material or collaborate on Problem Set 0.

3 Graphs

In this course, the word "graph" will refer to a *weighted directed graph* (a.k.a. weighted digraph), which is a more general type of graph than the type defined in Chapter 1 of Spielman's book.

Definition 1. *A weighted directed graph G consists of:*

- *A set V of vertices.*
- *A set $E \subseteq V \times V$ of directed edges.*
- *A weight function $w : E \to \mathbb{R}^+$.*

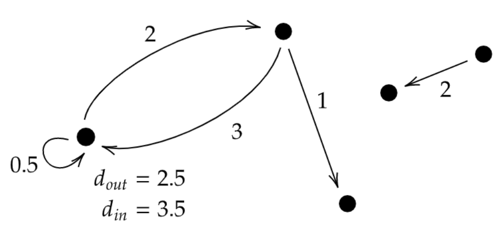

Figure 1: A Typical Weighted Digraph

Unlike in Spielman, our definition accommodates self-loops, which are directed edges of the form (a, a). We can arrive at an equivalent definition of weighted directed graphs using only a vertex set V and a weight function $w : V \times V \to \mathbb{R}^{\geqslant 0}$. In such a graph, the edge set is understood to be

$$E = \{e \in V \times V : w(e) > 0\}.$$

When studying weighted directed graphs, there are many important special cases:

- A graph is *undirected* if $w(a, b) = w(b, a)$ for all vertices $a, b \in V$. One should think of an undirected graph as a directed graph in which each edge has a corresponding edge in the opposite direction.
- An *unweighted multigraph* is a graph with a weight function $w : V \times V \to \mathbb{N}$. The non-negative integer $w(a, b)$ is called the *multiplicity* of the edge (a, b).
- A graph is *regular* if there exists an integer d such that $d_{\text{out}}(a) = d_{\text{in}}(a) = d$ for all vertices $a \in V$. Here, $d_{\text{out}}(a)$ denotes the *weighted out-degree* of a, defined by

  $$d_{\text{out}}(a) = \sum_{b \in V} w(a, b),$$

  and $d_{\text{in}}(a)$ denotes the *weighted in-degree* of a, defined by

  $$d_{\text{in}}(a) = \sum_{b \in V} w(b, a).$$

- A graph is *Eulerian* if $d_{\text{out}}(a) = d_{\text{in}}(a)$ for all vertices $a \in V$. The class of Eulerian graphs includes all undirected graphs, as well as all regular graphs. The name "Eulerian" comes from the observation that connected Eulerian graphs are precisely those that can be traversed with a *directed Eulerian tour*.

  Note: You may be familiar with *undirected Eulerian tours* from a previous graph theory course, but it's important to note that these concepts are different. For example, the graph K_4, a complete graph on 4 vertices, can be viewed as a connected Eulerian digraph. Therefore, it can be traversed with a directed Eulerian tour. However, it *cannot* be traversed with an undirected Eulerian tour since it has vertices of odd degree.
- A graph is *simple* if it is undirected, unweighted, has no parallel edges, and has no self-loops.

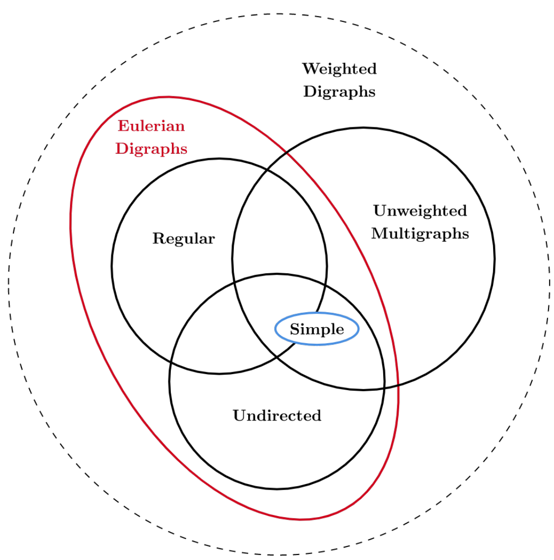

Figure 2: Types of Weighted Digraphs

4 Matrices

There are several different matrices that can be used to describe a graph G:

- The *weighted adjacency matrix* M is given by $M(a,b) = w(b,a)$. That is, the $(a,b)^{\text{th}}$ entry of M is the weight of the edge from vertex b to vertex a in G.
- The *random-walk matrix*, also known as the *random-walk transition matrix* or *diffusion matrix*, is

$$W = MD_{\text{out}}^{-1},$$

where D_{out} is the diagonal matrix

$$D_{\text{out}} = \begin{pmatrix} d_{\text{out}}(1) & & 0 \\ & \ddots & \\ 0 & & d_{\text{out}}(n) \end{pmatrix}.$$

The random-walk matrix scales the columns of M by the reciprocals of the vertices' out-degrees. Thus, each column of W sums to 1, defining a probability distribution over the out-neighbors of a vertex b:

$$W(a,b) = \frac{w(b,a)}{d_{\text{out}}(b)}.$$

- We already defined D_{out}, but D_{in} can be defined similarly. In an Eulerian digraph, in-degrees and out-degrees match, so $D_{\text{in}} = D_{\text{out}} = D$, and D is called the *degree matrix.*
- For an Eulerian digraph, the *Laplacian matrix* is

$$L = D - M,$$

the *random-walk Laplacian* is

$$\begin{aligned} L_{\text{rw}} &= I - W \\ &= LD^{-1} \end{aligned}$$

and the *normalized Laplacian* is

$$\begin{aligned} N &= I - D^{-1/2} M D^{-1/2} \\ &= D^{-1/2} L D^{-1/2} \\ &= D^{-1/2} L_{\text{rw}} D^{1/2}. \end{aligned}$$

Matrix	Eigenvector	Eigenvalue	Symmetric if G is Undirected?
L	$\mathbf{1}$	0	✓
L_{rw}	$D\mathbf{1}$	0	✗
W	$D\mathbf{1}$	1	✗
N	$D^{1/2}\mathbf{1}$	0	✓
A	?	?	✓

Table 1: Properties of Matrices for an Eulerian Digraph

Table 1 summarizes various properties that the above matrices are guaranteed to have when G is an Eulerian digraph. Proofs can be found in Appendix B. Below we collect some observations regarding Table 1:

- The vector $D\mathbf{1}$ is the vector of vertex degrees, as is $M\mathbf{1}$.
- The Eulerian condition is important. For example, if G is Eulerian, then $D\mathbf{1}$ is an eigenvector of W. However, if G is not Eulerian, then $D_{\text{out}}\mathbf{1}$ is not an eigenvector of W. Rather, $WD_{\text{out}}\mathbf{1} = D_{\text{in}}\mathbf{1}$.
- You may have heard that the stationary distribution of a random walk on an undirected graph is proportional to the vertex degrees. The third row of Table 1 generalizes this fact from the class of undirected graphs to the broader class of Eulerian directed graphs.
- The "?" entries in the last row sheds some light on why Spielman prefers to study graph matrices other than the adjacency matrix. However, this is not to say that the adjacency matrix is useless: we will encounter proofs in future lectures that analyze the adjacency matrix directly.
- If G is regular, then all of the matrices in Table 1 are essentially equivalent. This is because the degree matrix is $D = dI$ for a d-regular graph (here, I denotes the identity matrix).

5 The Spectral Theorem

An important theorem in linear algebra is the Spectral Theorem. In order to state the theorem, recall that a basis $v_1, \ldots, v_n$ of $\mathbb{R}^n$ is *orthonormal* if

$$v_i^\top v_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

and that an $n \times n$ matrix V is *orthogonal* if its columns comprise an orthonormal basis, i.e. $V^\top V = I$. Equivalently, V is orthogonal if its rows comprise an orthonomal basis, i.e. $VV^\top = I$.

Theorem 2 (Spectral Theorem)**.** *If M is a symmetric real matrix, then*

$$M = V\Lambda V^\top = \sum_{i=1}^{n} \lambda_i v_i v_i^\top$$

for an orthogonal matrix V with columns $v_1, \ldots, v_n$ and a diagonal matrix Λ with entries $\lambda_1, \ldots, \lambda_n$.

The second expression for M in the above theorem, namely $M = \sum_{i=1}^{n} \lambda_i v_i v_i^\top$, makes it clear that λ_i are the eigenvalues of M, and v_i the corresponding eigenvectors. If M is asymmetric, the Spectral Theorem no longer applies, but the following (weaker) theorem still holds.

Theorem 3 (Singular Value Decomposition)**.** *If M is a real matrix (not necessarily square), then*

$$M = U\Sigma V^\top$$

for an orthogonal matrix U, an orthogonal matrix V, and a non-negative, diagonal matrix Σ.

Intuitively, the Spectral Theorem says that for any symmetric matrix $M \in \mathbb{R}^{n\times n}$, we can find an orthonormal basis of $\mathbb{R}^n$ that diagonalizes M. In contrast, the Singular Value Decomposition says that for any matrix $M \in \mathbb{R}^{m\times n}$, we can find orthonormal bases of $\mathbb{R}^m$ and $\mathbb{R}^n$ (possibly different, even when $m = n$) that diagonalize M.

Proof of the Spectral Theorem (Outline). The proof has two steps. First, we will identify one eigenvalue λ_1 of M, along with a corresponding eigenvector v_1. Next, we will consider the restriction of M to the $(n-1)$-dimensional subspace orthogonal to v_1 and proceed by induction.

1. In order to identify an eigenvalue of M, we define the *Rayleigh quotient*

$$\frac{x^\top Mx}{x^\top x}$$

for non-zero vectors x. It's easy to check that when x is an eigenvector of M, the Rayleigh quotient equals the corresponding eigenvalue. This gives us an idea: let's choose a vector $x = v_1$ that maximizes the Rayleigh quotient, and then prove that v_1 is actually an eigenvector of M with eigenvalue

$$\lambda_1 = \max_{x \neq 0} \frac{x^\top Mx}{x^\top x} = \max_{\|x\|=1} x^\top Mx.$$

This proof, which can be found in Spielman (and Appendix A), uses the fact that M is symmetric.

2. Once we've found λ_1 and v_1, we'd like to shift our attention to the subspace

$$v_1^\perp = \{w \in \mathbb{R}^n : w^\top v_1 = 0\}.$$

Using the symmetry of M, one can show that $Mv_1^\perp \subseteq v_1^\perp$. In other words, the restriction of M to $v_1^\perp$ (denoted $M|_{v_1^\perp}$) is an operator on the $(n-1)$-dimensional subspace $v_1^\perp$. By induction on the dimension, we know that $M|_{v_1^\perp}$ has an orthonormal basis of eigenvectors $v_2, \ldots, v_n$ spanning $v_1^\perp$. Adding v_1 to this list yields the desired orthonormal basis of all of $\mathbb{R}^n$.

□

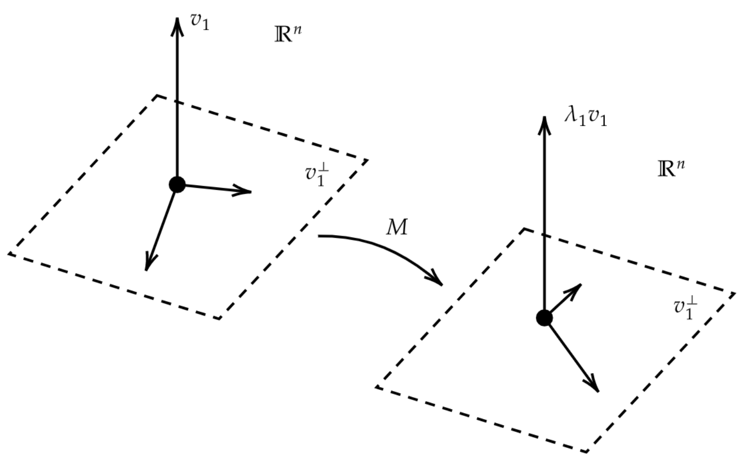

Figure 3: M maps v_1 to $\lambda_1 v_1$ and maps all of $v_1^\perp$ into $v_1^\perp$

References

[Spi19] Daniel A. Spielman. *Spectral and Algebraic Graph Theory (Incomplete Draft)*. 2019.

A Proof that the Rayleigh Quotient is maximized at a maximum eigenvector

As posed, this is a maximization problem over a potentially unbounded set, so *a priori* no maximum is guaranteed to exist. However, since the quantity is scale invariant we may restrict our domain to the unit sphere, which is a closed and compact set where the function must obtain a maximum. Thus, the function obtains a global maximum at x^* for some x^*.

For the maximizer x^*, we can use a first order condition which shows that the gradient of the Rayleigh quotient must be zero at that point: explicitly, since

$$\nabla \frac{x^\top M x}{x^\top x} = \frac{(x^\top x)(2Mx) - (x^\top M x)(2x)}{(x^\top x)^2}$$

we must have

$$(x^{*\top} x^*) M x^* = (x^{*\top} M x^*) \boldsymbol{x}^*,$$

which implies $Mx^* = \frac{x^{*\top} M x^*}{x^{*\top} x^*} x^*$. Since the right hand side is a scalar quantity times x, the gradient of the Rayleigh quotient is zero if and only if the vector x is an eigenvector. (One can check the second case directly) And since x^* maximizes the Rayleigh quotient, x^* must correspond to the largest eigenvalue.

B Proofs of Properties of Matrices

B.1 Properties of the Laplacian

Note that $L\mathbf{1} = \text{diag}(D) - M\mathbf{1} = \text{diag}(D) - \text{diag}(D) = \mathbf{0}$.

B.2 Properties of the Random Walk Laplacian

Note that $L_{\text{rw}}(D\mathbf{1}) = (LD^{-1})D\mathbf{1} = L\mathbf{1} = \mathbf{0}$.

B.3 Properties of the Random Walk Matrix

Note that $WD\mathbf{1} = MD^{-1}D\mathbf{1} = M\mathbf{1} = D\mathbf{1} = \text{diag}(D)$.

B.4 Properties of Normalized Laplacians

Note that $ND_{\text{out}}^{1/2}\mathbf{1} = D_{\text{out}}^{-1/2} L D_{\text{out}}^{-1/2} D_{\text{out}}^{1/2}\mathbf{1} = D_{\text{out}}^{-1/2} L\mathbf{1} = \mathbf{0}$. Note also that if G is symmetric, then

$$(D_{\text{out}}^{-1/2} L D_{\text{out}}^{-1/2})^\top = D_{\text{out}}^{-1/2} L^\top D_{\text{out}}^{-1/2} = D_{\text{out}}^{-1/2} L D_{\text{out}}^{-1/2}.$$