

## Lecture 4

Instructor: Salil Vadhan

Scribes: Muyuan Lin

## 0 Agenda

- Cayley graphs wrap-up
- Perron-Frobenius theory for symmetric matrices
- Random walks on undirected graphs

## 1 Cayley Graph Wrap-Up

Suppose that  $\mathcal{G}$  is a finite **group**<sup>1</sup> with a binary operation denoted by additive notation  $+$ , and  $W : \mathcal{G} \rightarrow \mathbb{R}^{\geq 0}$  a **weight function**<sup>2</sup>.

**Definition 1.** The Cayley graph  $\text{Cay}(\mathcal{G}, W)$ <sup>3</sup> is a directed graph whose vertex set is the finite group  $\mathcal{G}$ , and edge weights  $w(a, b) = W(b - a)$ , where  $a, b \in \mathcal{G}$ .

If the weight function  $W$  is an indicator of a subset  $S \subseteq \mathcal{G}$ , i.e.  $W = \mathbb{1}_S$ , then  $\text{Cay}(\mathcal{G}, W) = \text{Cay}(\mathcal{G}, S)$  with edges  $\{(a, a + s) : s \in S\}$ .

**Theorem 2.** For any Cayley graphs from a finite abelian group  $\mathcal{G}$ , they have a common orthogonal basis of complex eigenvectors of the adjacency matrix, which are given by the homomorphism<sup>4</sup>  $\mathcal{X} : \mathcal{G} \rightarrow S \subseteq \mathbb{C}^*$  (i.e.  $\mathbf{x} \in \mathcal{C}^{\mathcal{G}}$  such that  $x_a = \mathcal{X}(a)$ ), with corresponding eigenvalues  $\hat{W}_{\mathcal{X}} = W\mathcal{X}^* = \sum_{s \in \mathcal{G}} w(s)\mathcal{X}^*(s)$ .

Notes:

- We can calculate eigenvectors explicitly.
- We have this whole big family of graph (from a fixed group but arbitrary weight functions) that shares a common basis of eigenvectors (the Fourier basis).
- We have a formula to compute the corresponding eigenvalues  $\hat{W}_{\mathcal{X}} = \sum_{s \in \mathcal{G}} w(s)\mathcal{X}^*(s)$ , which is just the Fourier transform of the weight function.

**Example 3.** *Boolean hypercube.* Consider a Cayley graph<sup>5</sup> from the group  $\mathcal{G} = \mathbb{Z}_2^d$ ,  $S = \{e_1, e_2, \dots, e_d\}$  (i.e.  $W = \mathbb{1}_S$ ), with the operation of addition mod 2. The Fourier basis (eigenvectors) for each  $r \in \mathbb{Z}_2^d$

$$\mathcal{X}_r(x) = (-1)^{\langle r, x \rangle} = (-1)^{\langle r, x \rangle \text{ mod } 2}$$

<sup>1</sup>By the definition of group,  $\mathcal{G}$  satisfies requirements including *closure*, *associativity*, *identity element* and *inverse element*

<sup>2</sup>Think of it as a probability distribution on the group elements if it is normalized. A common case of the weight function  $W$  is indicator  $\mathbb{1}_S$  of some subset  $S \subseteq \mathcal{G}$  of group elements.

<sup>3</sup>The definition also applies to non-abelian groups where the operation is not commutative

<sup>4</sup>A group homomorphism is a map  $f : G \rightarrow H$  between two groups such that the group operation is preserved:  $f(g_1 * g_2) = f(g_1) \cdot f(g_2)$ , where the group operation on the left hand side of the equation is that of  $G$  and on the right hand side that of  $H$ .

<sup>5</sup>The graph is a cube in three dimensions, a square in two dimensions, and generally a  $d$ -dimensional cube.

**Table 1:** Calculation of the adjacency matrix's eigenvalues of the hypercube.  $L = D - M, W = MD^{-1}, N = LD^{-1}$ . Note, the eigenvalues of M are symmetric as the graph is bipartite.

| $ r $    | # of e-values    | e-value of M | e-value of L | e-value of W | e-value of N |
|----------|------------------|--------------|--------------|--------------|--------------|
| 0        | 1                | d            | 0            | 1            | 0            |
| 1        | d                | d-2          | 2            | $1 - 2/d$    | $2/d$        |
| 2        | $\binom{d}{2}$   | d-4          | 4            | $1 - 4/d$    | $4/d$        |
| 3        | $\binom{d}{3}$   | d-6          | 6            | $1 - 6/d$    | $6/d$        |
| $\vdots$ | $\vdots$         | $\vdots$     | $\vdots$     | $\vdots$     | $\vdots$     |
| d/2      | $\binom{d}{d/2}$ | 0            | d            | 0            | 1            |
| $\vdots$ | $\vdots$         | $\vdots$     | $\vdots$     | $\vdots$     | $\vdots$     |
| d-1      | d                | -(d-2)       | 2d-2         | $-(1-2/d)$   | $2-2/d$      |
| d        | 1                | -d           | 2d           | -1           | 2            |

The corresponding eigenvalues

$$\begin{aligned} \hat{W}_r &= \sum_{s \in \mathbb{Z}_2^d} \mathbb{1}_S(s) \cdot (-1)^{\langle r, s \rangle} \\ &= \sum_{i=1}^d (-1)^{\langle r, e_i \rangle} = \sum_{i=1}^d (-1)^{r_i} \\ &= d - 2|r| \end{aligned}$$

where  $|\cdot|$  is the Hamming weight (i.e. the number of nonzero components).

**Example 4.** Directed  $n$ -cycle.  $\mathcal{G} = \mathbb{Z}_n, S = \{1\}$ . The Fourier basis: for  $r \in \mathbb{Z}_n$ ,

$$\mathcal{X}_r = \omega^{rx}, \text{ where } \omega = e^{2\pi i/n}$$

The eigenvalues are  $\omega^0, \omega^1, \omega^2, \dots, \omega^{n-1}$ .

If the graph is undirected,  $S = \{\pm 1\}$ , then the eigenvalues are  $\omega^r + \omega^{-r} = 2 \cos(2\pi r/n)$ .

**Example 5.** Noisy Hypercube.  $\mathcal{G} = \mathbb{Z}_2^d, W(s) = p^{|s|} \cdot (1-p)^{d-|s|}$ , where  $p$  is the parameter of Bernoulli distribution.

The eigenvalues are

$$\begin{aligned} \hat{W}_r &= \sum_{s \in \mathbb{Z}_2^d} W(s) \mathcal{X}_r^*(s) \\ &= \mathbb{E}_{s \sim \text{Ber}(p)^d} [(-1)^{\langle r, s \rangle}] = \mathbb{E}_{s \sim \text{Ber}(p)^d} \left[ \prod_{i=1}^d (-1)^{r_i s_i} \right] \\ &= \prod_{i=1}^d \mathbb{E}_{b \sim \text{Ber}(p)} [(-1)^{r_i b}] \quad \Leftarrow \text{each bit is i.i.d.} \\ &= (1-2p)^{|r|} \approx \begin{cases} 1 - 2p|r| & \text{when } |r| \ll 1/p \\ e^{-k} & \text{when } |r| \gg k/p \end{cases} \end{aligned}$$

### Why Cayley Graph?

- Use algebra to understand graph properties (e.g. eigenvalues) or vice versa

- Highly symmetric (“vertex transitive”)
- Compact description of huge graphs (the set S)
- Explicit construction of large useful graphs (e.g. “expand graphs”)
- Connections to useful combinatorial objects (e.g. error correcting code)

### Why Abelian Cayley Graphs?

- Easier to analyze (Fourier analysis)
- Capture specific graphs of interest
- But have inherent limitations compared to non-abelian Cayley graphs (e.g. degree vs. expansion properties /  $\lambda_2$  / diameter)

## 2 Perron-Frobenius Theorem for Symmetric Matrices

Let  $M$  be a symmetric nonnegative real matrix (e.g. adjacency or normalized adjacency matrix <sup>6</sup>) with corresponding connected graph  $\mathcal{G}$ , edges  $\mathcal{E} = \{(a, b) : M(a, b) > 0\}$  and eigenvalues  $\mu_1 \geq \mu_2 \geq \dots \geq \mu_n$ . Then

- 1)  $\exists$  strictly positive eigenvectors  $v_1$  with eigenvalue  $\mu_1$  (and the only nonnegative eigenvalues are multiples of  $v_1$ ).
- 2)  $\mu_1 > \mu_2$  (cf. for the eigenvalues of Laplacian,  $\lambda_2 > 0 = \lambda_1$  iff  $\mathcal{G}$  is connected).
- 3)  $\mu_1 \geq -\mu_n$ , with equality iff  $\mathcal{G}$  is bipartite. Moreover, if  $\mathcal{G}$  is bipartite, then  $\mu_{n-i} = -\mu_{i+1}$ , for  $i = 0, \dots, n-1$ .

### Sketch of Proof

- 1) Let  $\phi_1$  be any eigenvector with eigenvalue  $\mu_1$ , and  $\mathbf{x} = |\phi_1|$  ( $|\cdot|$  is component-wise absolute value)

$$\mu_1 = \frac{\phi_1^T M \phi_1}{\phi_1^T \phi_1} \leq \frac{\mathbf{x}^T M \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \mu_1 \Rightarrow \mathbf{x} \text{ is an eigenvector of eigenvalue } \mu_1$$

$\mathbf{x}$  is at least non-negative. We then show that  $\mathbf{x}$  is strictly positive. Assume for contradiction  $\mathbf{x}(a) = 0$  for some vertex  $a$ . We know since  $\mathbf{x}$  is not a zero vector, there is  $\mathbf{x}(b) > 0$  for some vertex  $b$ . Let  $r$  = the length of a path from  $b$  to  $a$  (note that  $\mathcal{G}$  is connected), then

$$(M^r \mathbf{x})(a) = \mu_1^r \mathbf{x}(a) > 0 \text{ (positive value of } b \text{ “propagate” to } a)$$

which contradicts the assumption that  $\mathbf{x}(a) = 0$ .

- 2) Consider an eigenvector  $\phi_2$  with eigenvalue  $\mu_2$ . Let  $\mathbf{y} = |\phi_2|$

$$\mu_2 = \frac{\phi_2^T M \phi_2}{\phi_2^T \phi_2} \leq \frac{\mathbf{y}^T M \mathbf{y}}{\mathbf{y}^T \mathbf{y}} \leq \mu_1$$

If  $\mu_2 = \mu_1$ , then  $\mathbf{y}$  is nonnegative eigenvector of eigenvalue  $\mu_1$ . From 1),  $\mathbf{y}$  is strictly positive. Thus  $\phi_2$  does not have any zero entries. On the other hand, since  $\phi_2$  is orthogonal to  $\phi_1$ ,  $\phi_1$  is strictly positive,  $\phi_2$  has both positive and negative entries. As  $\mathcal{G}$  is connected, there must be some edge  $(a, b)$  for which  $\phi_2(a) < 0 < \phi_2(b)$ . Thus  $\phi_2^T M \phi_2 = \sum_{a,b} M(a, b) \phi_2(a) \phi_2(b)$  must be strictly smaller than  $\mathbf{y}_2^T M \mathbf{y}_2 = \sum_{a,b} M(a, b) \mathbf{y}(a) \mathbf{y}(b)$  which contradicts our assumption that  $\mu_2 = \mu_1$ .

<sup>6</sup>The associated random walk matrix of a  $d$ -regular graph is also symmetric. For a nonregular undirected graph, the random walk matrix is not symmetric.

- 3) (a) ( $\Rightarrow$ ) Observe that  $\mathcal{G}$  is bipartite iff  $\mathcal{G}^2$  is disconnected ( $\mathcal{G}^2$  is the graph obtained by having a length-2 walk on  $\mathcal{G}$ ). Also, the eigenvalues of  $\mathcal{G}^2$  are simply  $\mu_1^2, \dots, \mu_n^2 \geq 0$  (note that the order might change). Thus, if  $\mu_1 = -\mu_n$ , then  $\mathcal{G}^2$  is disconnected by the previous fact we proved and hence implies  $\mathcal{G}$  is bipartite.
- (b) ( $\Leftarrow$ ) From proposition 4.5.4 in Spielman's note, we know that if  $\mathcal{G}$  is bipartite and  $\mu$  is an eigenvalue of matrix  $M$ ,  $-\mu$  is also an eigenvalue of  $M$ , which implies  $\mu_1 = -\mu_n$ .

□

### 3 Random Walks on Undirected Graphs

Consider an undirected, weighted graph  $\mathcal{G}$  with  
the walk matrix of the graph  $W = MD^{-1}$ ,  
the initial probability distribution on vertices  $p_0 \in \mathbb{R}^n$ , and  
the probability distribution after  $t$  steps of random walks  $p_t = W^t p_0 \in \mathbb{R}^n$ .

**Claim 6.** The walk matrix  $W$  of the graph  $\mathcal{G}$  has a basis of eigenvectors  $\phi_1, \dots, \phi_n$  with eigenvalues  $1 \geq \omega_1 \geq \omega_2 \geq \omega_3 \geq \dots \omega_n \geq -1$ .

*Proof.* The normalized adjacency matrix  $A =: D^{-1/2}MD^{-1/2}$  is symmetric, and thus has orthogonal basis of eigenvectors  $\psi_1, \dots, \psi_n$  with corresponding eigenvalues  $\omega_1, \dots, \omega_n$ . Then

$$W(D^{1/2}\psi_i) = D^{1/2}AD^{-1/2}D^{1/2}\psi_i = D^{1/2}A\psi_i = \omega(D^{1/2}\psi_i)$$

Thus  $W$  has eigenvector  $D^{1/2}\psi_i$ , for  $i = 1, \dots, n$  with eigenvalue  $\omega$ . □

The eigenvalues and eigenvectors are useful for understanding the behavior of random walks. We can decompose the initial probability distribution and the distribution at time  $t$  as

$$p_0 = c_1\phi_1 + c_2\phi_2 + \dots + c_n\phi_n$$

$$W^t p_0 = c_1\omega_1^t\phi_1 + c_2\omega_2^t\phi_2 + \dots + c_n\omega_n^t\phi_n$$

where  $\omega_i$  and  $\phi_i$  are eigenvalues/eigenvectors of  $W$ . The probability distribution after  $t$  steps will be dominated by the term with eigenvalue(s) of the largest magnitude. The contribution of any terms with eigenvalues strictly smaller than that of the largest eigenvalue will decay exponentially.

**Theorem 7.** For start vertex  $a$ , and end vertex  $b$

$$|p_t(b) - \pi(b)| \leq \sqrt{\frac{d(b)}{d(a)}} \cdot (\max_{i>1} |\omega_i|)^t$$

To get to within total variation distance  $\epsilon$  of  $\pi$ , it suffices to have

$$\sqrt{\frac{d(b)}{d(a)}} \cdot (\max_{i>1} |\omega_i|)^t \leq \epsilon$$

There are two cases where the maximum absolute value of the other eigenvalues are not strictly less than one: disconnected and bipartite graphs.

1. If the graph is disconnected then  $\omega_2$  is equal to 1, then the states are not guaranteed convergence, or they do not converge to a unique stationary distributions.

---

<sup>7</sup>The eigenvectors of  $W$  are not necessarily orthogonal to each other when the graph is not regular, as opposed to  $A$ .

2. If the graph is bipartite, the states will be flipping back and forth between the two sides of the bipartite. Correspondingly, the magnitude of eigenvalues  $\omega_n$  is one and the random walks produce oscillating behavior.

The lazy random walk is to guarantee convergence even when the graph is bipartite (we already assume the graph is connected).