

## Lecture Lecture 10: Laplacian

Instructor: *Alex Andoni*Scribes: *Jiahe Shi, Pierre Tholoni*

## 1 Reminder: Spectral Decomposition for Graphs

$A$  is the  $n \times n$  adjacency matrix.

$D$  is the diagonal matrix of degrees.

$$\hat{A} = D^{-1/2}AD^{-1/2}.$$

And we proved last time that eigenvalues  $\lambda_1 \dots \lambda_n$  satisfy  $-1 \leq \lambda_n \leq \dots \leq \lambda_1 = 1$ , where eigenvector  $v_1$  is proportional to  $(\sqrt{d_1}, \dots, \sqrt{d_n})^T$ .

**Theorem 1.**  $\lambda_2 < 1$  if and only if  $G$  is connected

*Proof.* We can prove this theorem by showing that if  $G$  is disconnected, then  $\lambda_2 = 1$ .

Since  $G$  is disconnected, we can reorder the nodes, such that the adjacency matrix has a connected component on its upper left corner, and the rest on the lower right corner. There are no edges between them, so we can treat them as two separate matrix, and do spectral analysis on them independently. Thus, each would have a eigenvalue of 1.  $\square$

**Remark:** the number of eigenvalues equal to 1 is equal to the number of connected components.

## 2 Laplacian

### 2.1 Definition

Define Normalized Laplacian:

$$\hat{L} = I - \hat{A}$$

$\hat{L}$  has eigenvalues  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_n$ . We can observe that the eigenvectors  $v_1, v_2, \dots, v_n$  of  $\hat{A}$  are also eigenvectors of  $\hat{L}$ . We can show this by considering  $v_i$ :

$$\hat{L}v_i = (I - \hat{A})v_i = v_i - \lambda_i v_i = (1 - \lambda_i)v_i$$

Thus,  $\mu_i = 1 - \lambda_i$  is an eigenvalue of  $\hat{L}$ . From this, we know that  $0 = \mu_1 \leq \mu_2 \leq \dots \leq \mu_n \leq 2$ .

$$\begin{aligned} \hat{L} &= I - D^{-1/2}AD^{-1/2} \\ D^{1/2}\hat{L}D^{1/2} &= D - A \\ &= L \end{aligned}$$

$$L = \begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -1 & \dots & d_n \end{pmatrix}$$

↓ edge
↓ no edge

Here, we define  $L$  as Laplacian.

**Claim 2.**  $\forall x \in \mathbb{R}^n, x^T Lx = \sum_{e \in E} (x_i - x_j)^2$

*Proof.* Let's define  $L_e$  for some edge  $e = (i, j)$  to be laplacian of the graph whose only component is  $e$ . Then  $L$  is the summation of all such sub graphs:  $L = \sum_{e \in E} L_e$ . This means, we can rewrite the above expression  $x^T Lx$  as:

$$\begin{aligned} x^T Lx &= \sum_{e \in E} x^T L_e x \\ &= \sum_{e=(i,j)} x_i^2 + x_j^2 - x_i x_j - x_j x_i \\ &= \sum_e (x_i - x_j)^2 \end{aligned}$$

□

## 2.2 Intuition of Laplacian

We can think about a network of resistors. Suppose every vertex is a resistor. Let  $K$  be the current, and  $k_{ij}$  be the current between  $i$  and  $j$ . Let  $x_i$  be the potential at  $i$ , and assuming all the resistances are one. Then, we can show that  $k_{ij} = x_j - x_i$ , and the energy dissipation, which is potential difference times current,  $E_{i,j} = (x_j - x_i) * k_{ij} = (x_j - x_i)^2$ . The total dissipation is the sum of  $E_{ij}$  for all the edges, which corresponds exactly to the Laplacian  $x^T Lx = \sum_e (x_i - x_j)^2$ . So the quadratic form  $x^T Lx$  can be thought of as some energy dissipation, where  $x$  is the potential at different nodes.

## 2.3 Normalized Laplacian

The quadratic form with normalized Laplacian is:

$$\begin{aligned} x^T \hat{L}x &= x^T D^{-1/2} L D^{-1/2} x \\ &= \sum_{e=(i,j)} \left( \frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}} \right)^2 \end{aligned}$$

which by definition is always positive, and the smallest eigenvalue  $\mu_1$  by definition is:

$$\mu_1 = \min \frac{x^T \widehat{L}x}{\|x\|_2^2} \geq 0$$

**Claim 3.** *If  $G$  is connected, then  $\mu_2 > 0$*

*Proof.* We will instead prove that if  $\mu_2 = 0$ , then  $G$  is disconnected. If  $\mu_2 = 0$ , it implies that  $\exists x, x \perp v_1, R(x) = \frac{x^T \widehat{L}x}{\|x\|_2^2} = 0$ , which implies  $\forall e = (i, j), \frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}} = 0$ . Define  $\beta = \frac{x_1}{\sqrt{d_1}}$ , this quantity  $\beta$  for every node must be equal.

If  $G$  is connected, then  $\beta = \frac{x_1}{\sqrt{d_1}}$  for all  $i \in [n]$ , which implies  $x$  is proportional to  $(\sqrt{d_1}, \dots, \sqrt{d_n})^T = v_1$ .

But this cannot be the case, because we defined  $x$  to be perpendicular to  $v_1$ .

Thus,  $G$  must be disconnected. □

**Theorem 4.**  $\mu_2 > 0$  if and only if  $G$  is connected

### 3 Drawing graphs: spectral representation

**Goal of this section.** The eigenvectors of the normalized Laplacian  $v_1, \dots, v_n$  are vectors in  $\mathbb{R}^n$  that assign a real value for each node. We know that  $v_1 = (\sqrt{d_1}, \dots, \sqrt{d_n})$ . We also know that  $v_2$  tells us something about whether  $G$  is connected. But what does  $v_2$  look like? And what about  $v_3$ ?

#### 3.1 Matlab experiments

**Representing a graph in 2D.** Suppose we want to represent a graph  $G$  in 2 dimensions. We can assign a point  $(x_i, y_i)$  of the plane  $\mathbb{R}^2$  for each node  $i$ , and draw a line between each pair of points that are associated an edge  $(i, j) \in E$ .

**Representing using the Laplacian eigenvectors.** We are free to choose any representation for  $G$ . Let us try the following one, that associates to each node  $i$  the  $i$ th coordinate of the second and the third eigenvectors of the normalized Laplacian:

$$\forall i \in [n], (x_i, y_i) = (v_{2i}, v_{3i})$$

Surprisingly, this choice of representation gives intuitive 2D drawings of graphs without any *a priori* geometrical knowledge of it – the graph is only described by its combinatorial properties.

**Simple geometric examples.** For example:

- If  $A$  is a “square” graph defined by  $V = [4]$  and  $E = \{(1, 2), (2, 3), (3, 4), (4, 1)\}$ , the representation defined by  $Lap(A)$ ’s eigenvectors looks like an actual square.
- If  $A$  corresponds to a cube, its 2D representation looks like a cube squashed on a plan, *i.e.* a cube in perspective.

- For an hypercube, or a cube in dimension 5, the 2D representation is still nice.
- If  $A$  corresponds to a grid, its 2D representation is an actual grid, albeit slightly tilted.

See the Matlab code (when uploaded with the course material) for more examples.

**What about 3D representations?** We can plot a graph in 3D by using  $v_2, v_3, v_4$  instead of  $v_2, v_3$  only. For example, the spectral representation of a dodecahedron in 3D corresponds to the canonical representation in 3D.

### 3.2 Drawing a graph in 2D

After these experiments, let us now try to understand why spectral representations make sense. This section is not very formal.

**General problem.** How to draw a graph nicely in a plane? For example, suppose that we have some abstract graph such as a social network, given as an adjacency matrix. How to represent this graph in the most human readable form?

**Formalization.** We are looking for a function that associates a point of the plane  $\mathbb{R}^2$  to each node in  $V = [n]$ :

$$i \in V \mapsto (x_i, y_i) \in \mathbb{R}^2$$

Instead of reasoning about functions, *i.e.* elements of  $(\mathbb{R}^2)^n$ , we can reason about vectors as elements of  $(\mathbb{R}^n)^2$ . Thanks to this bijection, the problem is equivalent to finding two vectors  $X, Y \in \mathbb{R}^n$  that represent the graph nicely.

**Error minimization.** Intuitively, we would like to draw  $G$  such that for all nodes  $i, j$ , if  $(i, j) \in E$  then the  $i$  and  $j$  are close in the representation.

In other words, we want to minimize the error  $E(x, y)$  defined by:

$$E(X, Y) := \sum_{(i,j) \in E} \|(X_i, Y_i) - (X_j, Y_j)\|_2^2 = \sum_{(i,j) \in E} (X_i - X_j)^2 + (Y_i - Y_j)^2$$

We choose the euclidean distance quite arbitrarily, but it is a common choice (Mean Squared Error) and has nice mathematical properties.

**Problem 1.**

$$\min_{X, Y} E(X, Y)$$

A simple solution to this problem is to put all the points together at  $(0, 0)$ . It is not a very useful representation of a graph, we would like some tension that keeps points apart.

**Problem 2.** Let us reason in dimension 1 first. We can try to avoid the first problem by fixing the norm of  $X$ .

$$\min_{\|X\|_2^2=1} E(X)$$

Once again, we can reach  $E = 0$  with an uninteresting solution where  $X = \left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)$ .

Note that  $E(X) = \sum_{(i,j) \in E} (X_i - X_j)^2 = X^T L X$ : it is the Laplacian quadratic form (see claim in Section 1). The solution  $X = \left(\frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}}\right)$  is actually the first eigenvector of  $L$ .

**Problem 3.** We need to force  $X$  to be even more spread out. Let us fix  $\|X\|_2^2 = 1$  and also force  $\sum_i X_i = 0$  i.e.  $\mathbf{1} \cdot X = 0$  i.e.  $\mathbf{1} \perp X$ , to make sure that the coordinates of  $X$  are not all equal.

$$\min_{\|X\|_2^2=1, \mathbf{1} \perp X} E(X)$$

We have  $E(X) = X^T L X$ . By the theorem on Rayleigh quotients, for  $R(X) = \frac{X^T L X}{\|X\|^2}$  for  $X \neq 0$ , we have  $\lambda_i = \max_{X \neq 0, X \perp v_1, \dots, v_{i-1}} R(X)$ . Hence, a solution to Problem 3 is  $v_2$ , the second eigenvector of  $L$ .

**Problem 4.** Let us go back in 2D.

$$\min_{\substack{\|X\|_2^2=1, \mathbf{1} \perp X \\ \|Y\|_2^2=1, \mathbf{1} \perp Y}} E(X, Y) = X^T L X + Y^T L Y$$

In this problem,  $X$  and  $Y$  are optimized independently so a solution is to put them both equal to  $v_2$ .

**Problem 5.** We want  $Y$  to give a maximum of information after we conditionalize by  $X$ .

$$\min_{\substack{\|X\|_2^2=1, \mathbf{1} \perp X \\ \|Y\|_2^2=1, \mathbf{1} \perp Y, X \perp Y}} E(X, Y) = X^T L X + Y^T L Y$$

The optimal solution is to take  $v_2$  for  $X$  and  $v_3$  for  $Y$  (if we minimize  $X$  first, otherwise we might have to swap  $v_2$  and  $v_3$ ).

**Remark.** This problem is not the only way to represent nicely a graph, but it is reasonable.