

Toolkit

Laplacians ...

1. Linear Algebra review.
2. ~~Contextualize first~~ Define the Laplacian
3. Example using K_n and path.
4. Bipartiteness, connectivity.
5. Cheeger and robust generalization.

Linear Algebra

~~Fast forward~~ So we've been working w/ undirected graphs which are represented by symmetric matrices. If we want a linear algebraic understanding of graph theory then it bodes well to recall some theorems regarding symmetric matrices.

(Spectral Theorem): Let $A \in \mathbb{R}^{n \times n}$ be symmetric. Then ...

1. A has n not necessarily distinct, but real eigenvalues.
2. If $i \neq j$, $\langle v_i, v_j \rangle = 0$ i.e. $\{v_1, \dots, v_n\}$ form an orthonormal basis.
3. A admits a spectral decomposition

$$A = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^\top$$

⑪ Variational Characterization of Eigenvalues.

So we know for $A \in \mathbb{R}^{n \times n}$ symmetric, λ is an eigenvalue w/ eigenvalue ...

$$Av = \lambda v.$$

This isn't a very robust formulation. i.e. the algebra of this equality is in some sense rigid. This characterization is in some sense rigid and not too amenable for designing algorithms. Instead we will often use the variational characterization.

To discuss this we first define the Rayleigh Quotient.

(Rayleigh Quotient) : The Rayleigh quotient of A given $x \neq 0$ is ...

$$\frac{x^T A x}{x^T x}.$$

The Variational characterization of eigenvalues is given by.

Theorem: Let $A \in \mathbb{R}^{n \times n}$ Symmetric matrix w/ eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Then ...

$$\lambda_n = \max_{x \in \mathbb{R}^n, x \neq 0} \frac{x^T A x}{x^T x} = \max_{\|x\|=1} x^T A x$$

Proof: The spectral theorem tells us A has an orthonormal basis $\{v_1, \dots, v_n\}$ of eigenvectors. Thus $\forall x \in \mathbb{R}^n$

$$x^T A x = \left(\sum_{i=1}^n c_i v_i \right)^T A \left(\sum_{i=1}^n c_i v_i \right)$$

$$= \left(\sum_{i=1}^n c_i v_i \right)^T \underbrace{\left(\sum_{i=1}^n \lambda_i v_i v_i^T \right)}_{\text{but } A \text{ also has a spectral decomposition}} \left(\sum_{i=1}^n c_i v_i \right)$$

~~but A also has a spectral decomposition,~~

$$= \left(\sum_{i=1}^n c_i v_i \right)^T \left(\sum_{i=1}^n c_i A v_i \right)$$

$$= \left(\sum_{i=1}^n c_i v_i \right)^T \left(\sum_{i=1}^n \lambda_i c_i v_i \right)$$

$$= \sum_{i=1}^n c_i^2 \lambda_i$$

Wt a similar calculation we can deduce

$$x^T x = \sum_{i=1}^n c_i^2$$

Thus the Rayleigh quotient is.

$$\frac{x^T A x}{x^T x} = \frac{\sum_{i=1}^n c_i^2 \lambda_i}{\sum_{i=1}^n c_i^2}$$

$\Rightarrow \leq \lambda_n$: Since $\lambda_i \leq \lambda_n \ \forall i \dots$

$$\frac{\sum_{i=1}^n c_i^2 \lambda_i}{\sum_{i=1}^n c_i^2} \leq \lambda_n \cdot \frac{\sum_{i=1}^n c_i^2}{\sum_{i=1}^n c_i^2} = \lambda_n$$

because

$$V_i \perp V_j \quad \forall i \neq j$$

$\Rightarrow \geq \lambda_n$: The max of $\frac{x^T A x}{x^T x}$ is bounded below for $x = v_n$.

$$\frac{V_n^T A V_n}{V_n^T V_n} = V_n^T \left(\sum_{i=1}^n \lambda_i v_i v_i^T \right) V_n = \lambda_n (V_n^T V_n)^2 = \lambda_n \ \square$$

And in general... let V_K be space orthogonal to $\{v_{n+1}, v_{n+2}, \dots, v_m\}$

$$\lambda_K = \max_{x \in V_K} \frac{x^T A x}{x^T x} \text{ or } \min_{x \in \{v_1, \dots, v_m\}} \frac{x^T A x}{x^T x}.$$

Similarly...

$$\lambda_1 = \min_{\|x\|=1} x^T A x.$$

The way you can think about this is as...

- An optimization problem to find the length of the axes of the ellipsoid formed as the image of A
~~(since A is symmetric)~~

provided $\lambda_1 \geq 0$.

Actually when ~~$\lambda_i > 0$~~ $\lambda_i \geq 0$ we say A is PSD.

~~PSD~~

(Positive Semidefinite) : $A \in \mathbb{R}^{n \times n}$ symmetric is PSD if

- $\forall i \quad \lambda_i \geq 0$.

Equivalently...

- $A = \sum u_i u_i^T$ are $k \times k$ one matrices.

$$\lambda_1 = \min_{\|x\|=1} x^T A x \geq x^T (\sum u_i u_i^T) x = \sum (x^T u_i)^2 \geq 0 \quad \square$$

Now we're ready to handle Capmations.

Laplacian

Let's now define the Laplacian.

(Laplacian): Given an undirected graph G , the Laplacian

$$L_G := D - A \text{ where}$$

$$D = \begin{pmatrix} d_1 & 0 \\ \vdots & \ddots \\ 0 & d_n \end{pmatrix} \quad A = \text{adjacency matrix}$$

$d_i := \text{degree of } i \in V$.

Convention...

i. We will order the Laplacian's eigenvalues as $\lambda_1 \leq \dots \leq \lambda_n$.

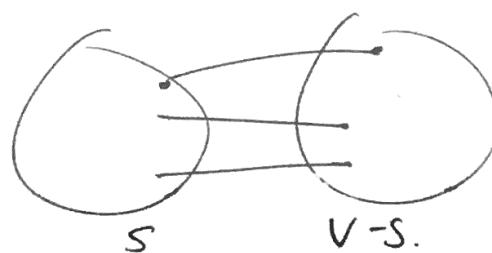
Why might you care about L_G .

Scenario 1: You're studying cut problems.

Input : Some undirected graph $G = (V, E)$

Output : $S \subseteq V$ subject to some measure of goodness

→ maybe say # of edges across $(S, V-S)$.



I can count $E(S, V-S)$ using ...

$$x \in \mathbb{R}^n \rightarrow x_i = \begin{cases} -1 & \text{if } v \notin S \\ +1 & \text{if } v \in S \end{cases}$$

of edges in $E(S, V-S) = \sum_{i \in S} (x_i - x_j)$

$$|E(S, V-S)| = \frac{1}{4} \sum_{(i,j) \in E} (x_i - x_j)^2$$

Because $(x_i - x_j)^2 = 4$ if $(i, j) \in E(S, V-S)$ o.w.

\Rightarrow Okay trust me that its natural to care about this sum of squares...

relate to

Why does this ~~exist about~~ Laplacians?

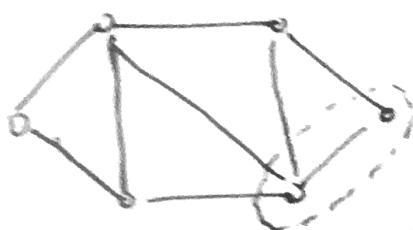
Claim: $x^T L x = \sum_{e \in E} (x_i - x_j)^2 \quad \forall x \in \mathbb{R}^n$.

To prove this let's start w/ some basic properties of Laplacians.

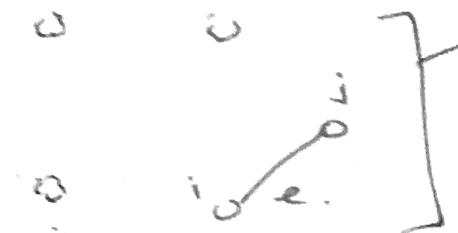
□ Thinking about a Laplacian.

You can think about a Laplacian as a sum of mini Laplacians...

Take a graph $G = (V, E)$



pick your favorite edge ...



Imagine if you defined the Laplacian ~~on~~ on graph w/ just this edge + ghost nodes.

$$b_e = \begin{pmatrix} 1 & * \\ * & -1 \end{pmatrix} \rightarrow \text{since } -A_e.$$

$$\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} + 1 \text{ along diag since } \deg(i) = \deg(j) = 1$$

everything else is 0.

b_e is a rank=1 matrix. Another way I can write this is as...

$$b_e = b_e b_e^T \rightarrow b_e = \begin{pmatrix} 0 \\ \pm 1 \\ 0 \\ \pm 1 \\ 0 \end{pmatrix} = \begin{pmatrix} +1 \\ -1 \end{pmatrix} \quad e = (i, j)$$

fix an orientation.

Claim: $L_G = \frac{1}{2} \sum_{e \in E} b_e = \sum_{e \in E} b_e b_e^T$ fixing an orientation among $e \in E$.

Proof: Think about it as $L_G = D - A$

$$L_G = \begin{pmatrix} 0 & * & * \\ * & 0 & * \\ * & * & 0 \end{pmatrix}$$

if an off diagonal is -1 then $(i, j) \in E$.

this means L_G will contribute that -1 and add +1 to the degrees of i and j. \square

Corollary : $L_G \succeq 0$ \forall graph.

Because L_G is PSD there are a couple implications.

1. All eigenvalues of L_G are non-negative
2. You can efficiently optimize our func.

(Citation: Cohen et al 2014.)

Remark : There is an $O(m\log n)$ time algorithm

for solving a linear system $Lx = b$ where L is the laplacian over some graph.

— That's faster than solving a list of m elements.

Back to original claim...

Corollary : $L_G = \mathbb{B}^T \mathbb{B}$ where $\forall i=1\dots m$

$$\mathbb{B}^T = \begin{matrix} e_1 \\ \dots \\ e_m \end{matrix} \left(\begin{array}{c|ccc} & b_{e_1} & & \\ & \dots & & \\ & b_{e_m} & & \end{array} \right)^T$$

flip it around.

\mathbb{B} is called the edge-vertex incidence matrix.

Now we can go to our original claim.

How does this relate to matrix tree theorem?

of spanning trees $\equiv \det(\underline{\mathbb{B}_0 \mathbb{B}_0^T})$

remove a row of \mathbb{B} .

Claim : $x^T L x = \sum_{e \in E} (x_i - x_j)^2 \quad \forall x \in \mathbb{R}^n$

Proof : Observe ...

$$x^T L x = x^T \left(\sum_{e \in E} L_e \right) x = \sum_{e \in E} x^T L_e x = \sum_{\{i,j\} \in E} (x_i - x_j)^2$$

$L_e = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad 0 \text{ elsewhere.}$

as required. \square

Spectrum of the Laplacian

The eigenvalues of the Laplacian encode some important information about our graph. ~~lets start from complete examples~~

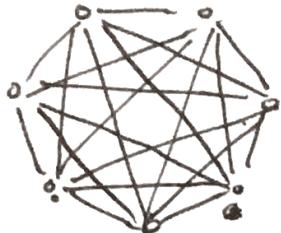
First claim that we can square away.

Claim : $\mathbf{1} \in \text{Ker } L_G$. That is $L_G \cdot \mathbf{1} = \mathbf{0}$.

Proof : You just do it.

Example:

1 The complete graph.



$$\rightarrow L = D - A = \begin{pmatrix} n-1 & -1 & & \\ -1 & n-1 & & \\ & & \ddots & \\ & & & n-1 \end{pmatrix}$$

K_n

$$= nI_n - J.$$

- $\mathbb{1}$ is an eigenvector for $\lambda = 0$.

$$(nI_n - J)\mathbb{1} = nI_n\mathbb{1} - J\mathbb{1} = n\mathbb{1} - n\mathbb{1} = 0.$$

→ *achieved*

- $n-1$ eigenvectors for $\lambda = n$.

$$V = \begin{pmatrix} +1 \\ -1 \\ 0 \\ \dots \\ 0 \end{pmatrix} \text{ such that } \begin{pmatrix} n-1 & -1 & & \\ -1 & n-1 & & \\ & & \ddots & \\ & & & n-1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ \dots \\ 0 \end{pmatrix} = \begin{pmatrix} n-1+1 \\ -1-(n-1) \\ -1+1 \\ -1+1 \\ \dots \end{pmatrix} = \begin{pmatrix} n \\ -n \\ 0 \\ 0 \\ \dots \end{pmatrix}$$

$$= UV.$$

Something to notice... $\lambda_0 = 0$ is smallest eigenvalue since $L \geq 0$. $\lambda_2 = n$ $\lambda_2 - \lambda_1 = n$ which is pretty big...

And K_n is really connected... hmmm...

Claim:

Another thing to notice.

- $\mathbb{1}$ is always an eigenvector

$$(D - A)\mathbb{1} = D\mathbb{1} - \underbrace{A\mathbb{1}}_{\text{counts degrees.}} = \cancel{\text{diag}(d_1, \dots, d_n)} - \cancel{\text{diag}(d_1, \dots, d_n)} = 0.$$

In fact for some reason we show that $\mathbb{1}$ is an eigenvector, allows us to show.

Claim : G_1 is connected $\Leftrightarrow L_G$ has an eigenvalue of 0 w/ multiplicity ≥ 1 .

Proof :

$\neg P \rightarrow \neg Q$: If G is disconnected then we can write G as the union of two subgraphs G_1, G_2 . disjoint

Because L is the sum of mini laplacians.

$$L_G = \begin{pmatrix} L_{G_1} & 0 \\ 0 & L_{G_2} \end{pmatrix}$$

The two eigenvectors that make 0 a mult ≥ 2 eigenvalue are

$$L_{G_1} \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \frac{0}{L_{G_2}} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$P \rightarrow Q$: If G is connected, consider a potential eigenvector v w/ eigenvalue 0 . ~~Since~~

$$v^T L v = 0 \rightarrow \sum_{(i,j) \in E} (v_i - v_j)^2 = 0.$$

by claim

$$\rightarrow (v_i - v_j)^2 = 0$$

$$\rightarrow v_i = v_j \quad \forall (i,j) \in E.$$

Since G is connected ...

$v_1 = v_2 = \dots = v_n$. because there are enough edges ...

This means that

$$v = c \cdot \mathbf{1}.$$

Since $Av=0 \rightarrow c=0$. There is only one eigenvector for $\lambda=0$. \square

Other properties ...



Corollary: G has k connected components \Leftrightarrow multiplicity of 0 is k .

Other properties

1. G is bipartite $\Leftrightarrow \lambda_n = 2d$ for G a d-regular graph.

Laplacian as a Robust Generalization

Okay we saw that the spectrum of L_G encodes combinatorial properties of G , but it actually allows us to generate like who comes, we could just run BFS.

Big Idea: The laplacian allows us to assign combinatorial properties to continuous values in a robust way. This allows us to generalize these properties in a very robust way.

- more time to work
- 1. λ_2 the second smallest eigenvalue is close to 0 if G is "close" to disconnected (or even worse). K_n as an example of very connected (claim)
 - 2. λ_k is small if G is close to having k connected components. (corollary)
 - 3. λ_n is close to $2d$ if G is close to bipartite. (additional claim). ← (we use this idea to come up with number non SDP soln for maxcut.)

Why is ① important? ~~but~~ Imagine a datacenter...

You want to thread ~~together~~ ^{Servers} together in a way that ensures you can be tolerant against faults.

— Do K_n ...

But that's expensive because cables!

WICIS

OR

What if you could find a graph with $\lambda_2 - \lambda_1 \gg 0$.

but $|E| = \Theta(n)$. ← you could have a well connected graph that's sparse!

Robust Generalization of Connectivity

So we're interested in formalizing what is meant when we say λ_2 is "small" when G is "close to disconnected"

① Close to disconnected.

What do we mean by G is close to disconnected.

One way to measure connectivity is via expansion.

(Expansion): The expansion of $G = (V, E)$ is ...

$$\Phi(G) = \min_{S \subseteq V, |S| \leq \frac{n}{2}} \Phi(S)$$

maybe
skip this.

where $\Phi(S)$ the expansion of S is ...

$$\Phi(S) = \frac{|E(S, V-S)|}{|S|}$$

Another way is via conductance ...

(Conductance): The conductance of $G = (V, E)$ is.

$$\phi(G) = \min_{S \subseteq V : \text{Vol}(S) \leq m} \phi(S)$$

where the conductance of S $\phi(S)$ is

$$\phi(S) = \frac{|E(S, V-S)|}{\text{Vol}(S)}$$

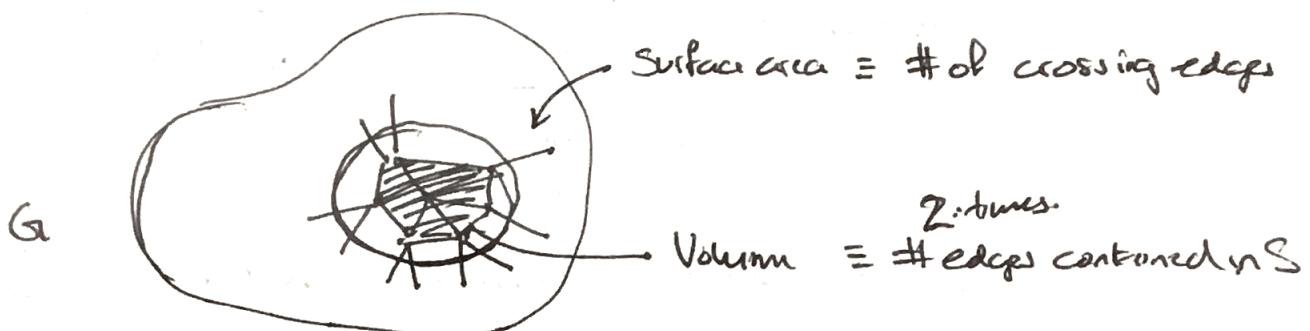
and the volume of S is ...

$$\text{Vol}(S) = \sum_{v \in S} \deg(v)$$

A way to understand the conductance is as follows...

$$\Phi(G) = \min_{S \subseteq V: |V-S| \leq m} \frac{|E(S, V-S)|}{\sum_{v \in S} \deg(v)}$$

Think about this as a "surface area" to "volume ratio"



For this reason inequalities involving $\Phi(G)$ are often called graph isoperimetric inequalities.

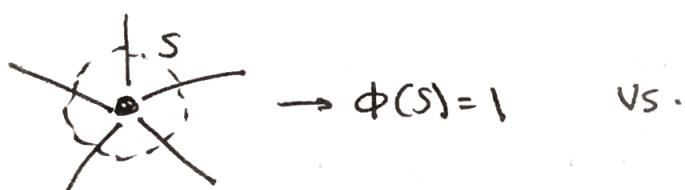
Regular isoperimetric inequality: L length of curve A is enclosed area

$$L^2 \geq 4\pi A.$$

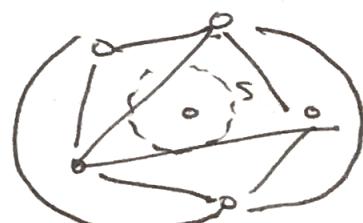
→ circle has smallest perimeter for given area.

Why does this measure connectivity?

$$\frac{|E(S, V-S)|}{\sum_{v \in S} \deg(v)} \quad \begin{array}{l} \xrightarrow{\hspace{1cm}} \\ \text{really close to 1 if a } \cancel{vertex} S \\ \text{is really well connected to} \\ \text{the outside world.} \end{array}$$



A vertex is really well connected to the outside world.



$$\Phi(S) = 0.$$

If for all $S \subseteq V$ $\phi(S)$ is large then G is really well connected. And in particular that's where we define ...

$$\phi(G) = \min_{\substack{S \subseteq V : \text{Vol}(S) \leq m}} \frac{|E(S, V-S)|}{\sum_{v \in S} \deg(v)}$$

min over
 $S \subseteq V$

$\text{Vol}(S) \leq m$ since you can't to select say $\frac{1}{2}$ vertices in S

recall $\sum_{v \in V} \deg(v) = 2m$.

- we'll usually call S st. $\phi(S)$ small a sparse cut (a cut that ~~disconnects~~
is poorly connected ~~$S \subseteq V$~~ to $V-S$)

Normalized Laplacian

So the spectrum of the laplacian could be as high as $2d_{\max}$. But $0 \leq \phi(G) \leq 1$. We want to normalize the laplacian to get rid of the dependence on d .

(Normalized Adjacency Matrix): Given A , the adjacency matrix of $G = (V, E)$.

$$cA = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$$

is the normalized adjacency matrix.

$$D^{-\frac{1}{2}} = \begin{pmatrix} y_{11} & & 0 \\ & \ddots & \\ 0 & \cdots & y_{nn} \end{pmatrix}$$

(Normalized Laplacian):

$$\mathcal{L} = I_n - cd = D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}} = D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$$

like literally just divide ~~the~~ entry of (L_G) by $\sqrt{\frac{1}{d_i d_j}}$

Claim: Let $\alpha_1 \leq \dots \leq \alpha_n$ be eigenvalues of the normalized adjacency matrix and $\beta_1 \leq \dots \leq \beta_n$ the eigenvalues of \mathcal{L} . Then

$$1. -k \leq \alpha_1 \leq \dots \leq \alpha_n \leq 1$$

$$2. 0 = \beta_1 \leq \dots \leq \beta_n \leq 2$$

Claim: Let $\alpha_1 \leq \dots \leq \alpha_n$ denote the eigenvalues of \mathcal{L} . Then --

$$0 = \alpha_1 \leq \dots \leq \alpha_n \leq 2.$$

Proof: To show $\alpha_i \geq 0$ observe \mathcal{L} is PSD.

$$\begin{aligned} x^T \mathcal{L} x &= x^T D^{-\frac{1}{2}} L D^{-\frac{1}{2}} x \\ &= \sum_{(i,j) \in E} x^T D^{-\frac{1}{2}} L e D^{-\frac{1}{2}} x \\ &= \sum_{(i,j) \in E} \left(\frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}} \right)^2 \geq 0 \end{aligned}$$

Thus $\alpha_i = \min_{x: \|x\|=1} x^T \mathcal{L} x \geq 0$.

To see $\alpha_n \leq 2$. Observe that $I_n + cd \succ 0$ as well.

$$\begin{aligned}
x^T L (I_n + cA) x &= x^T L x + 2x^T c A x \\
&= \sum_{(i,j) \in E} \left(\left(\frac{x_i}{\sqrt{d_i}} - \frac{x_j}{\sqrt{d_j}} \right)^2 + \frac{2x_i x_j}{\sqrt{d_i d_j}} \right) \\
&= \sum_{ij \in E} \left(\frac{x_i^2}{d_i} - \frac{2x_i x_j}{\sqrt{d_i d_j}} + \frac{x_j^2}{d_j} + \frac{2x_i x_j}{\sqrt{d_i d_j}} \right) \\
&= \sum_{ij \in E} \left(\frac{x_i^2}{d_i} + \frac{x_j^2}{d_j} \right) \geq 0.
\end{aligned}$$

Since $I_n + cA \succeq 0$, the smallest eigenvalue of cA has to be $\beta \geq -1$. Now look at $L = I_n - cA$. The largest eigenvalue of L is going to be $1 - \text{smallest eigenvalue of } cA$ thus ...

$$\alpha_n \leq 1 - (-1) = 2. \quad \square$$

Cheeger's Inequality

Now we're ready to state Cheeger's inequality.

Theorem (Cheeger): $\frac{1}{2}\lambda_2 \leq \phi(G) \leq \sqrt{2\lambda_2}$ where $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ are eigenvalues of L .

For simplicity we'll assume $G = (V, E)$ is d -regular. The colloquial way that Cheeger's is discussed is at an easy side and hard side.

$$\underbrace{\frac{1}{2} \lambda_2 \in \Phi(G)}_{\text{and}} + \underbrace{\text{Eqn}}_{\text{and}}$$

Let's say over what is meant that

- λ_2 is small if G is close to being disconnected.

If λ_2 is small then $\Phi(G)$ is small too... meaning
~~and~~ if G is not too well connected to the
 outside world

Because it is a large conductance \rightarrow ~~that~~ $\Phi(G)$ is
 large!

Let's prove it...

Claim - $\frac{1}{2} \lambda_2 \leq \Phi(G)$

Proof : ~~This tells us that~~ $\underline{\text{The idea is that when } \lambda_2 \text{ is small, we}}$
 $\underline{\text{that if we can pick out a sparse cut then } \lambda_2}$
 $\underline{\text{must be small.}}$

~~Sparse cut $\Rightarrow \Phi(G) \text{ is small}$~~

So to analyze, we should look at conductance of
 any cut. For intuition let's just suppose $|S| = \frac{n}{2}$

let's encode the cut as the vector ...

$$x_v = \begin{cases} +1 & \text{if } v \in S \\ -1 & \text{o.w.} \end{cases}$$

Notice that $\sum_v x_v = 0 \rightarrow \langle x, \mathbf{1} \rangle = 0$.

By \Leftrightarrow Variational characterization of eigenvalues.

$$\lambda_2 \geq \min_{\substack{x \perp \mathbf{1} \\ \|x\|=1}} \frac{\cancel{x^T L x}}{x^T x}$$

that's V_1

$$= \frac{\sum_{(i,j) \in E} (x_i - x_j)^2}{\sum_{i \in V} x_i^2}$$

normalized laplacian

because $D^{-1/2}$ is just $\frac{1}{\sqrt{d}}$ in other d regular

$$\lambda_2 \geq \min_{\substack{x \perp \mathbf{1} \\ \|x\|=1}} \frac{\cancel{x^T L x}}{x^T x} = \min_{x \perp \mathbf{1}} \frac{1}{d} \cdot \frac{x^T L x}{x^T x}$$

that's V_1

$$= \frac{\sum_{ij \in E} (x_i - x_j)^2}{d \cdot \sum_{i \in V} x_i^2}$$

$$= \frac{4 \cdot |E(S, V-S)|}{d|V|}$$

But $|V| = 2|S|$ so... (since we assume $|S| = \frac{n}{2}$)

$$= \frac{2 \cdot |E(S, V-S)|}{d|V|} = \text{Vol}(S) = \sum_{i \in S} \deg(i)$$

$$= 2\phi(S)$$

$$= \cancel{2\phi(S)} \text{ true today.}$$

The idea is that when x is ~~integral~~ a cut vector, the Rayleigh quotient of L_G behaves like the conductance of G . of the cut.

So if say we get any $S \subseteq V$, we could define one!

thus ...

$$x_i = \begin{cases} +\frac{1}{|S|} & \text{if } i \in S \\ -\frac{1}{|V-S|} & \text{o.w.} \end{cases}$$

To ensure x is integral and $\langle x, I \rangle \geq 0$.

The calculations work out but imma skip.

$$\begin{aligned} \|x\|_2^2 &\leq \frac{\sum_{i,j \in E} (x_i - x_j)^2}{d \cdot \sum_{i \in V} x_i^2} \quad \text{factor that reduces 4.} \\ &= \frac{|E(S, V-S)| \cdot \left(\frac{1}{|S|} + \frac{1}{|V-S|} \right)^2}{d \cdot \left(|S| \cdot \frac{1}{|S|^2} + |V-S| \cdot \frac{1}{|V-S|^2} \right)} \quad \|x\|_2^2 \\ &= \frac{|E(S, V-S)| \cdot |V|}{d \cdot \underbrace{|S| \cdot |V-S|}_{= |S| \cdot |V-S|}} \leq 2 \phi(S) \\ &= \frac{1}{|S|} + \frac{1}{|V-S|} \end{aligned}$$

Now that we have $\lambda_2 \leq 2\phi(S)$ $\forall S \subseteq V$ just pick

$$S^* = \underset{S \subseteq V : \text{Vol}(S) \leq m}{\text{argmin}} \frac{|E(S, V-S)|}{\text{Vol}(S)}$$

And run the argument to get $\lambda_2 \leq 2\phi(G) \Rightarrow \frac{\lambda_2}{2} \leq \phi(G)$. \square