Eigenvalues and Structures of Graphs

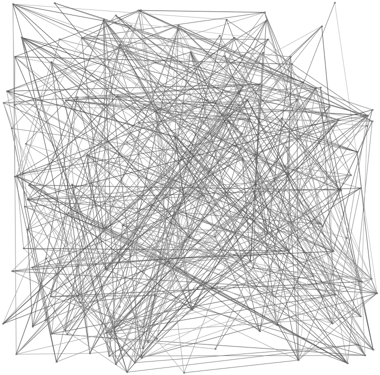
Steve Butler[1]

[1]Department of Mathematics
University of California, San Diego
www.math.ucsd.edu/~sbutler

Final Defense
27 May 2008

Graph theory.

Graphs are composed of vertices (or elements) and edges (which connect different elements). With the growing interest in networks (i.e., social networks and communication networks) the study of graphs have gained increase importance.



BUT most graphs people are interested in studying are large. So it becomes infeasible to keep track of the entire structure.

Spectral graph theory.

Spectral graph theory takes a “snapshot” of a graph by studying the eigenvalues of a matrix associated with a graph. The set of eigenvalues is fairly small when compared to the graph as a whole.

Questions about spectral graph theory:

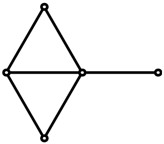
- How do we associate a graph with a matrix?
- What can the eigenvalues of a matrix say about a graph?
- What are the limitations of the approach?

The adjacency matrix.

The most studied matrix associated with a graph is the adjacency matrix A. The vertices index the rows and columns and the entries of A are used to indicate whether there is an edge between vertices.

Example:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$



Eigenvalues of A:
$2.6855\ldots,\ 0.3349\ldots,\ 0,\ -1.2713\ldots,\ -1.7491\ldots.$

An application of the adjacency matrix.

Eigenvalues can be used to find the trace of a matrix raised to a power. When raising the adjacency matrix to a power the entries count the number of closed walks.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be eigenvalues of A.

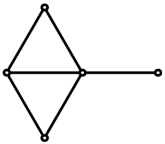
- $\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2$ is the trace of A^2 so is equal to twice the number of edges.
- $\lambda_1^3 + \lambda_2^3 + \cdots + \lambda_n^3$ is the trace of A^3 so is equal to six times the number of triangles.
- In general, if A is connected and not bipartite and λ_n is the largest eigenvalue then for large k the number of closed walks of length k is $\approx \alpha\lambda_n^k$ (for some α independent of k).

The combinatorial Laplacian.

A commonly studied matrix is the combinatorial Laplacian L. This matrix is $L = D - A$ where A is the adjacency matrix and D the diagonal degree matrix. This matrix is closely associated with the incidence matrix.

Example:

$$L = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 4 & -1 & -1 & -1 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & -1 & -1 & 3 & -1 \\ 0 & -1 & 0 & -1 & 2 \end{pmatrix}$$



Eigenvalues of L: 5, 4, 2, 1, 0.

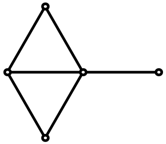
The application of the combinatorial Laplacian.

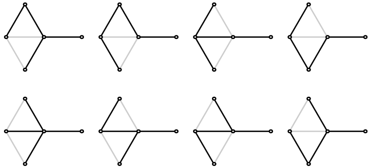
There is a classic application for the combinatorial Laplacian, namely counting the number of spanning trees of a graph. Many (if not most) applications involving the combinatorial Laplacian rely on this fact.

Kirchoff's Matrix Tree Theorem

If $\sigma_0 = 0 \leq \sigma_1 \leq \cdots \leq \sigma_{n-1}$ are the eigenvalues of L then the number of spanning trees of the graph is

$$\frac{\sigma_1 \sigma_2 \cdots \sigma_{n-1}}{n}.$$



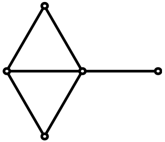


The normalized Laplacian.

A useful matrix for studying nonregular graphs is the normalized Laplacian $\mathcal{L}$. This matrix is $\mathcal{L} = D^{-1/2}LD^{-1/2}$. While it might not be intuitive it combines many of the best features of the adjacency matrix and the combinatorial Laplacian.

Example:

$$\mathcal{L} = \begin{pmatrix} 1 & \frac{-1}{2} & 0 & 0 & 0 \\ \frac{-1}{2} & 1 & \frac{-1}{\sqrt{8}} & \frac{-1}{\sqrt{12}} & \frac{-1}{\sqrt{8}} \\ 0 & \frac{-1}{\sqrt{8}} & 1 & \frac{-1}{\sqrt{6}} & 0 \\ 0 & \frac{-1}{\sqrt{12}} & \frac{-1}{\sqrt{6}} & 1 & \frac{-1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{8}} & 0 & \frac{-1}{\sqrt{6}} & 1 \end{pmatrix}$$



Eigenvalues of $\mathcal{L}$: $1.7287\ldots,\ 1.5,\ 1,\ 0.7712\ldots,\ 0.$

An application of the normalized Laplacian

A random walk involves starting at a vertex and then at each time step move to a vertex adjacent to the one that you are currently at.

Question: How long until we are at a “random” vertex?

The random walk is controlled by the probability transition matrix $D^{-1}A$ which has eigenvalues $1-\lambda_0, 1-\lambda_1, \ldots, 1-\lambda_{n-1}$ where the λ_i are the eigenvalues of the normalized Laplacian. In particular, the distance between the random walk after k steps and the stationary distribution is bounded above by

$$\max_{i\neq 0} |1-\lambda_i|^k \frac{\max_i \sqrt{d_i}}{\min_j \sqrt{d_j}}.$$

Comparing spectrums.

Different matrices can have different spectrums and so it becomes important in applications to make sure that we choose the right matrix. There is one well known exception.

If G is a regular graph of degree d then $\mathcal{L} = I - \frac{1}{d}A = \frac{1}{d}L$. In particular the spectrums are related to one another by scaling and shifting.

For graphs which are almost regular then the spectrums should also almost match.

Let G be a graph with λ_i and σ_i the eigenvalues of the normalized Laplacian and combinatorial Laplacian respectively. Then

$$\frac{1}{d_{\max}}\sigma_i \le \lambda_i \le \frac{1}{d_{\min}}\sigma_i.$$

Random graphs.

Random graphs are useful in graph theory, for instance, they can be used to show existence of graphs satisfying some property without actually producing a graph with the property. The problem is that while random graphs are nice, how do we know that any single graph is “random-like”, i.e., behaves like a random graph.

An important feature of random graphs is that edges are chosen independently. So we can assign some measure of how independently edges are placed in a graph. This leads to the notion of discrepancy of a graph.

Definition of discrepancy.

We need to compare the total number of edges between subsets X and Y of the vertices (denoted $e(X, Y)$) with how many we would expect if placed randomly. To estimate the number of expected edges we use volume.

$$\operatorname{vol} X = \sum_{v \in X} d_v.$$

Discrepancy.

The discrepancy of a graph is the minimal α so that for all subsets X and Y of the vertices

$$\left| e(X, Y) - \frac{\operatorname{vol} X \operatorname{vol} Y}{\operatorname{vol} G} \right| \leq \alpha \sqrt{\operatorname{vol} X \operatorname{vol} Y}.$$

Discrepancy for directed graphs.

This idea easily generalized to directed graphs. Let $e(X \to Y)$ count the number of directed edges from X to Y and let

$$\mathrm{vol}_{in}\, X = \sum_{v \in X} d_{in}(v) \qquad \mathrm{vol}_{out}\, X = \sum_{v \in X} d_{out}(v).$$

Directed discrepancy.

The discrepancy of a directed graph is the minimal α so that for all subsets X and Y of the vertices

$$\left| e(X \to Y) - \frac{\mathrm{vol}_{out}\, X \,\mathrm{vol}_{in}\, Y}{\mathrm{vol}\, G} \right| \le \alpha \sqrt{\mathrm{vol}_{out}\, X \,\mathrm{vol}_{in}\, Y}.$$

Relating discrepancy and eigenvalues.

Theorem

Let α be the discrepancy of a directed graph G and let D_{out} and D_{in} be the diagonal out- and in-degree matrices. Then

$$\alpha \leq \sigma_2\left(D_{out}^{-1/2} A D_{in}^{-1/2}\right) \leq 150\alpha(1 - 8\log\alpha).$$

Proof has three main ingredients.

- The definition of discrepancy can be rewritten as an inner product.
- For vectors x, y and a matrix M we have $|\langle x, My\rangle| \leq \sigma_1(M)\|x\|\|y\|$.
- Lots of bookkeeping.

Rewriting discrepancy.

For a set X let ψ_X be a 0-1 indicator vector.

$$\left| e(X \to Y) - \frac{\mathrm{vol}_{out} X \, \mathrm{vol}_{in} Y}{\mathrm{vol}\, G} \right| = \left| \langle \psi_X, A\psi_Y \rangle - \frac{\langle \psi_X, A\mathbf{1} \rangle \langle \mathbf{1}, A\psi_Y \rangle}{\langle \mathbf{1}, A\mathbf{1} \rangle} \right|$$

$$= \left| \langle \psi_X, A\psi_Y \rangle - \langle \psi_X, \left(\frac{D_{out} J D_{in}}{\langle \mathbf{1}, A\mathbf{1} \rangle}\right) \psi_Y \rangle \right|$$

$$= \left| \langle \psi_X, \left(A - \left(\frac{D_{out} J D_{in}}{\langle \mathbf{1}, A\mathbf{1} \rangle}\right) \right) \psi_Y \rangle \right|$$

$$= \left| \langle D_{out}^{1/2} \psi_X, \left(D_{out}^{-1/2} A D_{in}^{1/2} - \left(\frac{D_{out}^{1/2} J D_{in}^{1/2}}{\langle \mathbf{1}, A\mathbf{1} \rangle}\right) \right) D_{in}^{1/2} \psi_Y \rangle \right|$$

Quasirandom graphs.

Quasirandom graph properties are a collection of properties so that if a graph satisfies one of the properties it must satisfy all of them. Examples of quasirandom properties (for undirected graphs) include

- There are $\approx pn^2/2$ edges in the graph and for the adjacency matrix the largest eigenvalues is $\approx pn$ while all other eigenvalues are $o(n)$.
- For any "small" graph H on s vertices the number of copies of H in the graph is $\approx p^{\#E(H)}n^s$.
- For all but $o(n^2)$ pairs of vertices u and v the number of vertices w which have the same adjacency relationship with u and v is $\approx (1-2p+2p^2)n$.

Directed quasirandom graphs.

Very little is known about directed quasirandom graphs. The problem is that the obvious generalizations of the properties for undirected graphs are not equivalent for directed graphs.

Where is the problem? The "sameness" condition in undirected graphs is a condition on the entries of $A^2 + (J - A)^2$ which can be easily controlled using eigenvalues. For directed graphs the sameness condition is much more convoluted.

On the other hand, we saw that a small discrepancy implies a small second singular value and vice versa. So these two properties are in the same quasirandom class for directed graphs.

Localness of eigenvalues.

The condition $A\mathbf{x} = \lambda\mathbf{x}$ translates at each vertex into

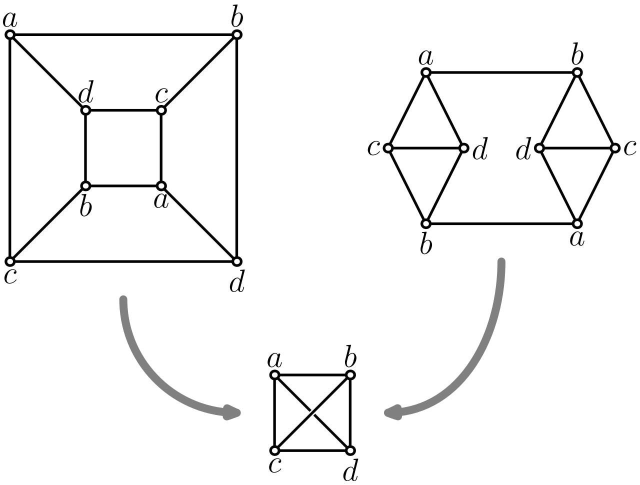
$$\sum_{u:u\sim v} w(u,v)\mathbf{x}(u) = \lambda\mathbf{x}(v).$$

The condition $\mathcal{L}\mathbf{y} = \lambda\mathbf{y}$ can be changed by letting $\mathbf{y} = D^{1/2}\mathbf{x}$ into $L\mathbf{x} = \lambda D\mathbf{x}$ which translates at each vertex into

$$d(v)\mathbf{x}(v) - \sum_{u:u\sim v} w(u,v)\mathbf{x}(u) = \lambda d(v)\mathbf{x}(v).$$

If two graphs share the same "local" structure then they should also share eigenvalues. This leads to the notion of covering.

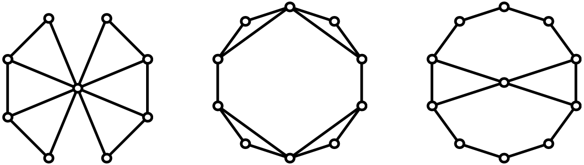
An example of coverings.



Since the two graphs on the top share the same local structure as K_4 they also will have all of the eigenvalues of K_4 in its spectrum.

Three related graphs.

Graphs can share eigenvalues for reasons other than covering a common graph. Consider the three graphs below.



These share eigenvalues for both the adjacency and normalized Laplacian but do not cover some common graph.

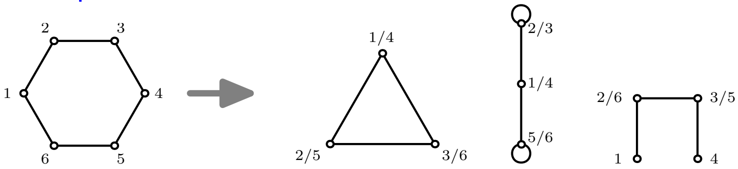
In this case the common eigenvalues comes from sharing the same anti-cover.

Definition of 2-edge-coverings.

The graph G is a 2-edge-covering of H if there is an onto map $\pi : V(G) \to V(H)$ such that

- if $u{\sim}v$ then $\pi(u){\sim}\pi(v)$, further $w(\pi(u), \pi(v)) = w(u, v)$;
- if $\pi(u){\sim}\bar{w}$ then there is some w so that $u{\sim}w$ and $\pi(w) = \bar{w}$;
- for edge $(\bar{w}, \bar{z})$ there are exactly two edges (p, q) and (r, s) so that $(\pi(p), \pi(q)) = (\pi(r), \pi(s)) = (\bar{w}, \bar{z})$.

Examples:



Basic property.

If $\pi(u) = \bar{u}$ then one of two possibilities can happen.

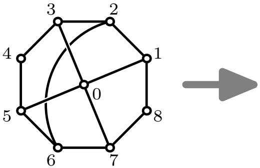
- If there is some $v \neq u$ so that $\pi(v) = \pi(u)$ then the neighborhood of u can be mapped 1-to-1 to the neighborhood of $\bar{u}$.
- If there is no $v \neq u$ so that $\pi(v) = \pi(u)$ then the neighborhood of u "folds" to form the neighborhood of $\bar{u}$, i.e., there is a 2-to-1 mapping.

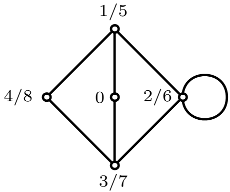
By our convention there cannot be two folding vertices next to each other. (This can be relaxed.)

2-edge-coverings and the adjacency matrix.

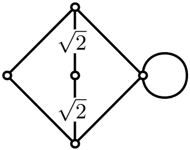
We can find the eigenvalues of the adjacency matrix of G by using the adjacency matrices of modified forms of H.

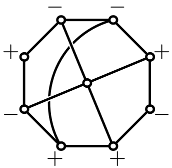
Example:

$G =$ 

$H =$ 

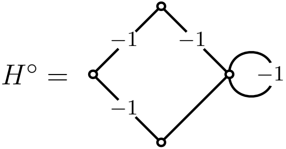
First, we will form the modified cover. This is done by multiplying any edge of H incident to a "folding" vertex by $\sqrt{2}$. This gives us the graph $H_\circ$.

$H_\circ =$ 



Using the covering, give a signing of vertices of G from $\{0, 1, -1\}$ so that (i) if a vertex folds it has sign 0, (ii) otherwise if $\pi(u) = \pi(v)$ then u and v have opposite sign. (Signings are not unique.)

Now we form the anti-cover H° by removing any folding vertex and incident edges from H. For the remaining edges, let $w(\bar{u}, \bar{v}) = w(u, v)\,\mathrm{sgn}(u)\,\mathrm{sgn}(v)$.



Theorem

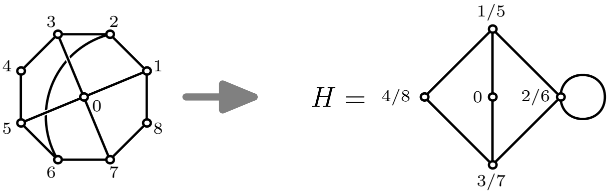
The eigenvalues of the adjacency matrix of G are the union of the eigenvalues of the adjacency matrices of $H_\circ$ and H°.

2-edge-coverings and the Laplacian matrix.

We can find the eigenvalues of the normalized Laplacian matrix of G by using the normalized Laplacian matrices of modified forms of H.

Example:

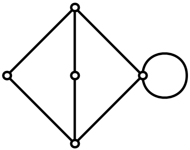
$G =$



$H =$

The first graph we use is H.
(Which we will also denote H_Δ.)

$H_\Delta =$

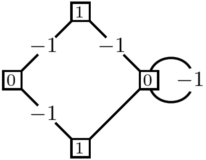


We again give a signing of vertices of G, remove any folding vertices in H and incident edges and modify the weight of the edges as before. We need one additional modification. Namely removing edges changes degrees, but for eigenvectors to lift to G we need to have comparable degrees. So we introduce weights on the vertices

$$w(\bar{u}) = \sum_{\substack{v:v\sim u \\ v \text{ folds}}} w(u,v).$$

Now we form the anti-cover H^{Δ}. The weights on vertices are denoted by putting " $\boxed{k}$ " at the vertex,

$H^{\Delta} =$



We define degrees by

$$d(\bar{u}) = w(\bar{u}) + \sum_{\bar{v}:\bar{v}\sim\bar{u}} |w(\bar{u},\bar{v})|.$$

Theorem

The eigenvalues of the normalized Laplacian matrix of G are the union of the eigenvalues of the normalized Laplacian matrices of H_Δ and H^Δ.

The proof is done by showing how eigenvectors of H_Δ and H^Δ can be lifted up to eigenvectors of G. In particular this will give a full set of eigenvectors so we have all the eigenvalues.

Example of a typical argument.

Let $\mathbf{x}_\Delta$ be an eigenvector of H_Δ for eigenvalue λ. Then consider $\mathbf{x}(v) = \mathbf{x}_\Delta(\pi(v))$, we claim this is an eigenvector for G. There are two cases to check.

- v is not a folding vertex.

$$d(v)\mathbf{x}(v) - \sum_{u:u\sim v} w(u,v)\mathbf{x}(u) =$$

$$d_\Delta(v_\Delta)\mathbf{x}_\Delta(v_\Delta) - \sum_{u_\Delta:u_\Delta\sim v_\Delta} w_\Delta(u_\Delta, v_\Delta)\mathbf{x}_\Delta(u_\Delta)$$

$$= \lambda d_\Delta(v_\Delta)\mathbf{x}_\Delta(v_\Delta) \;=\; \lambda d(v)\mathbf{x}(v).$$

Example of a typical argument.

- v is a folding vertex.

$$d(v)\mathbf{x}(v) - \sum_{u:u\sim v} w(u,v)\mathbf{x}(u) =$$

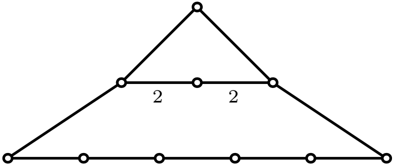
$$2d_\Delta(v_\Delta)\mathbf{x}_\Delta(v_\Delta) - 2\sum_{u_\Delta:u_\Delta\sim v_\Delta} w_\Delta(u_\Delta, v_\Delta)\mathbf{x}_\Delta(u_\Delta)$$

$$= 2\lambda d_\Delta(v_\Delta)\mathbf{x}_\Delta(v_\Delta) \;=\; \lambda d(v)\mathbf{x}(v).$$

Similar arguments work for $H_\circ$, H° and H^Δ.

Can you hear the shape of a graph?

Consider the two graphs below.



For the normalized Laplacian these graphs have *four* nontrivial eigenvalues in common and these come from the anti-cover graphs (folding in half along the vertical axis). Which are shown below.

But these graphs are not the same...

Can you hear the shape of a graph?

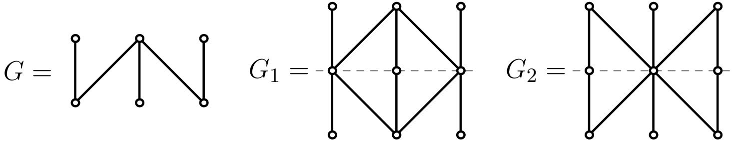
Computing the normalized Laplacian of the two graphs gives

$$\mathcal{L} = \begin{pmatrix} 1 & \frac{-1}{2\sqrt{2}} & 0 & 0 \\ \frac{-1}{2\sqrt{2}} & 1 & \frac{-1}{2} & 0 \\ 0 & \frac{-1}{2} & 1 & \frac{-1}{2} \\ 0 & 0 & \frac{-1}{2} & \frac{3}{2} \end{pmatrix}.$$

So these two anti-covers are not only cospectral but they have the same normalized Laplacian.

Application: constructing cospectral graphs.

Starting with a bipartite graph G with an equal number of vertices in each part then "unfold" along each part to form graphs G_1 and G_2.

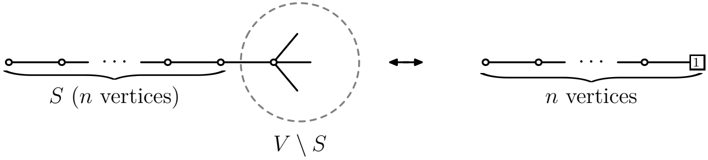


The eigenvalues of G_1 and G_2 come from the eigenvalues of G along with some isolated vertices. In particular, G_1 and G_2 are cospectral with respect to both the adjacency matrix and the normalized Laplacian.

Application: computing Dirichlet eigenvalues.

Dirichlet eigenvalues are related to random walks on a subset of vertices of a graph (i.e., set with *boundary*).

To compute the Dirichlet eigenvalues of $S \subset V$ we take the eigenvalues of $\mathcal{L}_S$ which is $\mathcal{L}$ restricted to the entries of S. The difference between $\mathcal{L}_S$ and the normalized Laplacian of the induced subgraph in S comes from how we interpret degrees.



So we can compute the Dirichlet eigenvalues by thinking of $\mathcal{L}_S$ as an anti-cover of some graph.

Interlacing eigenvalues.

By removing a few edges in a large graph the eigenvalues will change, but we might expect that they would not change by much.

Theorem

Let H be a subgraph of G, and H has t nonisolated vertices. If λ_i and θ_i are the eigenvalues of $\mathcal{L}(G)$ and $\mathcal{L}(G-H)$ respectively then

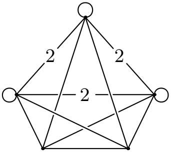
$$\lambda_{k-t+1} \leq \theta_k \leq \begin{cases} \lambda_{k+t-1} & H \text{ is bipartite;} \\ \lambda_{k+t} & \text{otherwise;} \end{cases}$$

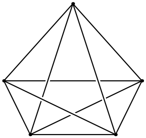
where $\lambda_{-t+1} = \cdots = \lambda_{-1} = 0$ and $\lambda_n = \cdots = \lambda_{n+t-1} = 2$.

Bipartiteness is important.

In the statement of the Theorem the upper bound for the interlacing is different for bipartite then for nonbipartite H. When a graph is bipartite we essentially have one more degree of freedom which lets us improve the bound.

Example:

$G =$ 

$G - H =$ 

In the two graphs above H is nonbipartite on three vertices. We also have that $\theta_1(G - H) = 5/4 > 8/7 = \lambda_3(G)$.

Courant-Fischer Theorem

If M is a real symmetric matrix with eigenvalues $\lambda_0 \le \cdots \le \lambda_{n-1}$, and $\mathcal{X}^k$ denotes a k dimensional subspace of $\mathbf{R}^n$. Then

$$\lambda_i = \min_{\mathcal{X}^{n-i-1}} \Big(\max_{x \perp \mathcal{X}^{n-i-1}, x \neq 0} \frac{x^T M x}{x^T x} \Big) = \max_{\mathcal{X}^i} \Big(\min_{x \perp \mathcal{X}^i, x \neq 0} \frac{x^T M x}{x^T x} \Big).$$

For the normalized Laplacian if we let $x = D^{1/2} y$ rearranging[1] gives

$$\lambda_i = \min_{\mathcal{Y}^{n-i-1}} \left(\max_{y \perp \mathcal{Y}^{n-i-1}, y \neq 0} \frac{\sum_{u \sim v} (y_u - y_v)^2 w(u,v)}{\sum_u y_u^2 d(u)} \right)$$
$$= \max_{\mathcal{Y}^i} \left(\min_{y \perp \mathcal{Y}^i, y \neq 0} \frac{\sum_{u \sim v} (y_u - y_v)^2 w(u,v)}{\sum_u y_u^2 d(u)} \right).$$

[1] Only works when there are no isolated vertices

A typical argument in the proof.

If $\{u_1, \ldots, u_t\}$ are vertices of H let $\mathcal{Z} = \{e_{u_1} - e_{u_2}, \ldots, e_{u_1} - e_{u_t}\}$.

$$
\begin{aligned}
\theta_k &= \min_{\mathcal{Y}^{n-k-1}} \left(\max_{y \perp \mathcal{Y}^{n-k-1}, y \neq 0} \frac{\sum_{u \sim v} (y_u - y_v)^2 w_{G-H}(u,v)}{\sum_u y_u^2 d_{G-H}(u)} \right) \\
&= \min_{\mathcal{Y}^{n-k-1}} \left(\max_{y \perp \mathcal{Y}^{n-k-1}, y \neq 0} \frac{\sum_{u \sim v} (y_u - y_v)^2 w_G(u,v) - \sum_{u \sim v} (y_u - y_v)^2 w_H(u,v)}{\sum_u y_u^2 d_G(u) - \sum_u y_u^2 d_H(u)} \right) \\
&\geq \min_{\mathcal{Y}^{n-k-1}} \left(\max_{y \perp \mathcal{Y}^{n-k-1}, y \perp \mathcal{Z}, y \neq 0} \frac{\sum_{u \sim v} (y_u - y_v)^2 w_G(u,v)}{\sum_u y_u^2 d_G(u) - \sum_u y_u^2 d_H(u)} \right) \\
&\geq \min_{\mathcal{Y}^{n-k-1}} \left(\max_{y \perp \mathcal{Y}^{n-k-1}, y \perp \mathcal{Z}, y \neq 0} \frac{\sum_{u \sim v} (y_u - y_v)^2 w_G(u,v)}{\sum_u y_u^2 d_G(u)} \right) \\
&\geq \min_{\mathcal{Y}^{n-k+t-2}} \left(\max_{y \perp \mathcal{Y}^{n-k+t-2}, y \neq 0} \frac{\sum_{u \sim v} (y_u - y_v)^2 w_G(u,v)}{\sum_u y_u^2 d_G(u)} \right) = \lambda_{k-t+1}.
\end{aligned}
$$

In conclusion...

Spectral graph theory has become an increasingly important way to study graphs. The material that has been presented here is just a small beginning of what can be said about graphs. As always, there still remains much more to be done.

THANK YOU!!