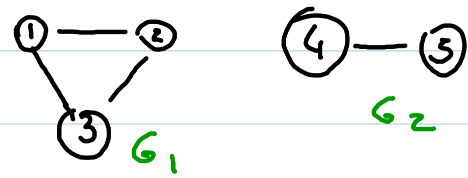
Lesson 17

Read ch 6

Graph connectivity, clustering,
Fiedler eigenvalue

Th: G is connected iff $\lambda_2 > 0$ for L_G
(i.e $\lambda = 0$ has multiplicity 1)

I will not give a general proof in class, but the discussion of the next example should give you an idea of how to prove this theorem.

Example G: 

Not connected

$$L_G = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

$\lambda_1 = 0$ is an eigenvalue for L_G

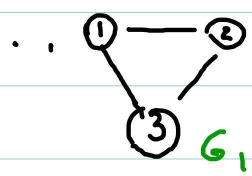
what is the multiplicity of $\lambda_1 = 0$?

suppose $u \neq 0 \quad L_G u = \vec{0}$.

$$\begin{bmatrix} L_{G_1} & 0 \\ 0 & L_{G_2} \end{bmatrix} \begin{bmatrix} \left.\begin{matrix} u_1 \\ u_2 \\ u_3 \end{matrix}\right\} = w_1 \\ \left.\begin{matrix} u_4 \\ u_5 \end{matrix}\right\} = w_2 \end{bmatrix} = \begin{bmatrix} L_{G_1} w_1 \\ L_{G_2} w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

So $\quad L_{G_1} w_1 = 0 \qquad L_{G_2} w_2 = 0$

what can w_1 look like? Recall G_1 and G_2 are connected.



$L_{G_1} w_1 = \vec{0}$

Th: If G is a connected graph then for L_G, $E_0 = \text{span}(\mathbb{1})$
ie 0 has multiplicity 1

Proof: assume $L_G w = 0$

Then $w^T L_G w = 0 = \sum_{\{i,j\} \in E} (w_i - w_j)^2$

so if there is an edge from i to j in G

$w_i = w_j$. What can w look like?

$$w = \begin{bmatrix} \vdots \\ w_h \\ \vdots \\ w_k \\ \vdots \end{bmatrix} \begin{array}{l} \\ \to \text{position } h \\ \\ \to \text{position } k \\ \\ \end{array}$$

Since G, is connected,

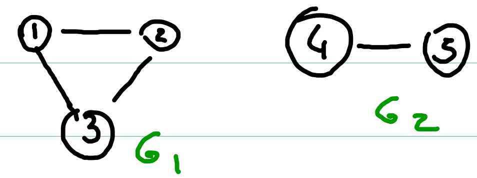
given any two vertices h, k in G_1 there is a path $h - i_1 \cdots i_m - k$

$w_h = w_{i_1} = \quad w_{i_m} = w_k \qquad$ so $w_h = w_k$

So $w = \begin{bmatrix} c \\ \vdots \\ c \end{bmatrix} = c\mathbb{1}$.

Back to our exemple

G



Not connected

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

$L_G\, u = 0 \iff u = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ and

$L_{G_1} w_1 = 0$, $L_{G_2} w_2 = 0$

By previous th:

$w_1 = \begin{bmatrix} c \\ \vdots \\ c \end{bmatrix} = c\,\mathbb{1}_3$ similarly $w_2 = \begin{bmatrix} d \\ d \end{bmatrix} = d\,\mathbb{1}_2$

so $u = \begin{bmatrix} c \\ c \\ c \\ d \\ d \end{bmatrix}$, therefore

so $u \in \operatorname{span}\left(\begin{bmatrix} \mathbb{1}_3 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathbb{1}_2 \end{bmatrix} \right) = E_0$

so $\lambda = 0$ has multiplicity 2 in L_G

Th: G has k connected components

$\iff \lambda_1 = 0$ has multeplicity k in L_G

Proof sketch: if G is connected $E_0 = \text{span}(\mathbb{1})$

n_1 vertices $\quad n_2$ vertices $\quad \ldots \quad n_k$ vertices

if $G = G_1 \quad G_2 \quad \cdots \quad G_k$

$$L_G = \begin{bmatrix} L_{G_1} & & & 0 \\ & L_{G_2} & & \\ & & \ddots & \\ 0 & & & L_{G_k} \end{bmatrix}$$ is a block metrix

$$E_0 = \text{span}\left(\begin{bmatrix} \mathbb{1}_{n_1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ \mathbb{1}_{n_2} \\ \vdots \\ 0 \end{bmatrix} \cdots \begin{bmatrix} 0 \\ \vdots \\ \mathbb{1}_{n_k} \end{bmatrix} \right)$$

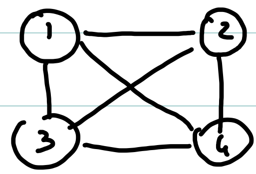
Graph clustering

Given graph G. L_G is Laplacian $D_G - A_G$.
L_G is PSD eigenvalues $0 = \lambda_1 \le \lambda_2 \cdots \le \lambda_n$
Now we focus on connected graphs and the second smallest eigenvalue of their Laplacian.
λ_2 : Fiedler value

Fact: if λ_2 is small "G is not very connected"
if λ_2 is big "G is very connected"

Not done in class you can read or skip, if you like

Example: Complete graph K_n : $\lambda_2 = n$



K_4

$$L_{K_4} = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

$$L_{K_n} = n I_n - J_n$$

J_n is the $n \times n$ matrix of all 1

- J_n has rank 1, so $\lambda = 0$ is eigenvalue with multiplicity $n-1$. Trace $(-J_n) = -n$

so the eigenvalues of $-J_n$ must be

$0, 0, \dots, 0, -n$

so $\det(-J_n - \lambda I_n) = 0$ has sol $0, \dots, 0, -n$

The eigenvalues of L_{K_n} are the solutions of $\det(n I_n - J_n - \lambda I_n) = 0$

which we can rewrite as

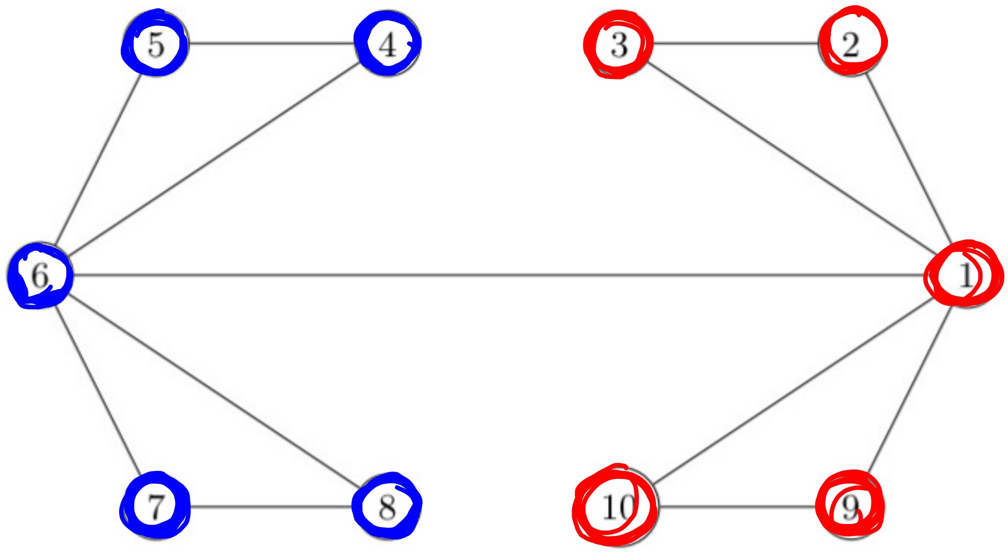
$\det(-J_n - (\lambda - n) I_n) = 0$ and therefore

has sol $(\lambda - n) = -n, 0, \dots\dots 0$

so $\lambda = 0, n, \dots\dots, n$ are the eigenvalues of L_{K_n}.

Fact: if a simple graph G has n vertexes any eigenvalue of L_G is $\leq n$

Example



G

G is connected and has 2 clusters i.e blue/red vertices are relatively well connected among themselves but not so so well connected to the rest of the graph

Def : Let G be a graph with vertices
$V = \{1 \cdots n\}$ and edges E

A cut in G is a partition of V into two
set A and V − A i.e cut is a subset $A \subseteq V$

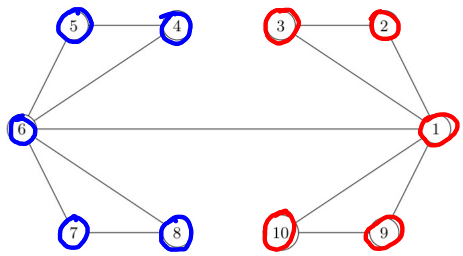
$E(A, V-A)$ is the subset of E that
contains all edges that connect a vertex in A
to a vertex in V − A.

The density of the cut is

$$\varphi_G(A, V-A) = n \, \frac{|E(A, V-A)|}{|A| \, |V-A|}$$

$|S|$ means the number of elements in the
set S

Example :



$A = \{4, 5, 6, 7, 8\}$

$V-A = \{1, 2, 3, 9, 10\}$

$E(A \quad V-A) = \{6, 1\}$

$$\varphi_G(A, V-A) = 10 \cdot \frac{|\{\{6,1\}\}|}{5 \cdot 5} = \frac{1}{5 \cdot 5} = \frac{10}{25} = \frac{2}{5}$$

Note :

In K_{10}, if $|A| = 5$ $\quad E(A, V-A) = 25$

in generel $E(A, V-A) = |A| \cdot |V-A|$ in K_n

$$\text{So } \varphi_G(A, V-A) = n \cdot \frac{E_G(A, V-A)}{E_{K_n}(A, V-A)}$$

Def $\varphi_G = \min_{A \subseteq V} \varphi_G(A, V-A)$

A cut $A, V-A$ such that $\varphi(A, V-A) = \varphi_G$ is called a sparsest cut.

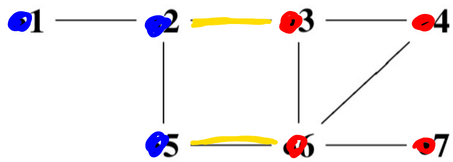
Goal: find a "sparse cut"

Note: How many different cuts are there in a graph with n vertices? V has 2^n subset, A cannot be $\emptyset$ or V, so # cuts is $\dfrac{2^n - 2}{2}$ (divide by 2 since $A, V-A$ and $V-A, A$ are the same cut)

Heuristic algorithm to find a sparse cut for G:

1. Find L_G
2. Calculate an eigenvector u_2 for λ_2
3. Sort the components of u_2 from biggest to smallest resolving ties any way you want:
   $u_{i_1} > u_{i_2} \cdots > u_{i_n}$
4. Try all the cuts $A = \{i_1\}$ $A = \{i_1, i_2\}$
   $A = \{i_1\ i_2\ i_3\}$ $\cdots$ and choose the one with the smallest density.

Example



$$L_G \quad 7\times 7 \qquad \begin{array}{c} 1 \\ \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 & -1 & 0 \\ 0 & -1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & -1 & -1 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \end{array}$$

$\lambda_2 = 0.5926$

$U = (0.71,\ 0.29,\ -0.06,\ -0.21,\ 0.04,\ -0.23,\ -0.56)$

$U_1 > U_2 > U_5 > U_3 > U_4 > U_6 > U_7$

$\varphi(\{1\}\{2,3,4,5,6,7\}) = 7 \dfrac{1}{1\cdot 6} = \dfrac{7}{6}$

$\varphi(\{1,2\}\{3,4,5\ 6,7\}) = 7 \dfrac{2}{2\cdot 5} = \dfrac{7}{5}$

$\varphi(\{1,2,5\}\{3,4,6,7\}) = 7 \dfrac{2}{3\cdot 4} = \dfrac{7}{6}$ $*$

$\varphi(\{1\ 2\ 5\ 3\}\ \{4\ 6\ 7\}) = 7 \dfrac{3}{3\cdot 4} = \dfrac{7}{4}$

$\varphi(\{1\ 2\ 5\ 3\ 4\}\ \{6\ 7\}) = 7 \dfrac{3}{5\cdot 2} = \dfrac{21}{10}$

$\varphi(\{1,2\ 5\ 3\ 4\ 6\}\ \{7\}) = 7 \dfrac{1}{6\cdot 1} = \dfrac{7}{6}$

only 6 to try vs $\dfrac{2^6-2}{2} = 31$

Th 1: $\varphi_G \geq \lambda_2$ ← Fiedler eigenvalue of L_G (hw problem)

Th 2: the algorithm in the previous page finds a cut A with $\varphi_G(A, V-A) \leq 4\sqrt{d_G \lambda_2}$ where d_G is the largest degree of a vertex in G

We will not prove th 2

Note th 1 and 2 say that

$$\underbrace{\lambda_2 \leq \varphi_G}_{\text{hw problem}} \underbrace{\leq}_{\text{for any } A} \varphi(A, V-A) \underbrace{\leq}_{\substack{\text{for the cut } A \\ \text{found by the alg}}} 4\sqrt{d_G \lambda_2}$$

where d_G is the maximum degree of a vertex in G