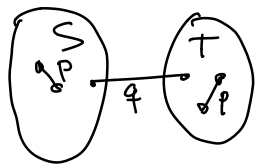
Partitioning a random graph that has a good partition

Model: partition V into $S \cup T$, $|S| = |T| = \frac{n}{2}$

Include (S,S) and (T,T) edges with prob p

(S,T) with prob q, $q < p$



Key: perturbation theory for eigenvectors

Weyl: If A has eigvals $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$

B has eigvals $\beta_1 \geq \cdots \geq \beta_n$

$R = A - B$ then $|\alpha_i - \beta_i| \leq \|R\|$.

(Recall, $\|R\| = \max_{\|x\|=1} \|Rx\|$ = max abs of eigval of R)

What about the eigenvectors?

$A\phi_i = \alpha_i \phi_i$, $B\psi_i = \beta_i \psi_i$ $\theta = \text{ang}(\phi_i, \psi_i)$

Davis & Kahan: $\sin 2\theta \leq \dfrac{2\|R\|}{\min_{j \neq i} |\alpha_i - \alpha_j|}$

Will prove $\sin\theta \leq \dfrac{2\|R\|}{\min_{j \neq i} |\alpha_i - \alpha_j|}$

proof. Assume, wolog $\alpha_i = 0$. Say by $A - \alpha_i I$, $B - \alpha_i I$

Let $\delta = \min_{j \neq i} |\alpha_i - \alpha_j|$ (note i is fixed here)

Compute $\|A \psi_i\|$, two ways

$$\|A \psi_i\| \leq \|(B + R)\psi_i\| \leq \underbrace{\|B \psi_i\|}_{\substack{= \beta_i \\ \leq \|R\|}} + \underbrace{\|R \psi_i\|}_{\leq \|R\|} \leq 2\|R\|$$

Setting $c_j = \phi_j^T \psi_i$, $\psi_i = \sum_j c_j \phi_j$

$$\|A \psi_i\|^2 = \sum_j c_j^2 \alpha_j^2 = \sum_{j \neq i} c_j^2 \alpha_j^2 \geq \sum_{j \neq i} c_j^2 \delta^2$$

$$= \delta^2 \sum_{j \neq i} c_j^2 = \delta^2 (1 - c_i^2) = \delta^2 \sin^2 \theta$$

so, $\delta \sin\theta \leq 2\|R\| \Rightarrow \sin\theta \leq \dfrac{2\|R\|}{\delta}$

Why dependence on difference between eigvals?

Consider $\begin{pmatrix} 1+\varepsilon & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 1+\varepsilon \end{pmatrix}$ with eigvecs $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Model. $A = \begin{array}{cc} \begin{pmatrix} p J_{n/2} & q J_{n/2} \\ q J_{n/2} & p J_{n/2} \end{pmatrix} & \begin{array}{c} S \\ T \end{array} \\ \begin{array}{cc} S & T \end{array} & \end{array}$ $M(a,b) = \begin{cases} 1 & \text{prob } A(a,b) \\ 0 & \text{prob } 1 - A(a,b) \end{cases}$ for $a \neq b$

independently.

$$M(a,a) = p = A(a,a)$$

So, add pI to adjacency matrix. Does not change eigenvectors.

$M = A + R$ where $R(a,b)$ chosen from $1-p, -p$ same group
$1-q, -q$ diff group.

By slight extension of last lecture,
$\|R\| \leq 3\sqrt{p(1-p)n}$ with high probability, if $p \geq C\frac{(\log n)^4}{n}$

some constant c.

Will recover S, T from ψ_2, if $p-q$ big enough.

$\alpha_1 \geq \alpha_2 \geq \alpha_3 = \cdots = \alpha_n = 0$ be eigvals of A.
$\mu_1 \geq \cdots \geq \mu_n$ eigvals of M.

$\alpha_1 = \left(\frac{n}{2}\right)(p+q)$ $\quad A \cdot \mathbf{1} = \frac{n}{2}(p+q) \cdot \mathbf{1}$ $\quad \phi_1 = \mathbf{1}$.

$\phi_2 = \begin{pmatrix} \mathbf{1}_{n/2} \\ -\mathbf{1}_{n/2} \end{pmatrix} = \mathbf{1}_S - \mathbf{1}_T$ $\quad A\phi_2 = \frac{n}{2}(p-q)\phi_2$

ϕ_2 gives the groups. if q close to p,

$$\alpha_2 - \alpha_3 = \frac{n}{2}(p-q) < \alpha_1 - \alpha_2 = 2q\frac{n}{2}$$

Let ψ_2 be second eigvec of M.

Relate $\arg(\phi_2, \psi_2)$ to # misclassified,
where classify a by $\operatorname{sign}(\psi_2(a))$

really $\phi_2(a) = \begin{cases} \frac{1}{\sqrt{n}} & a \in S \\ -\frac{1}{\sqrt{n}} & a \in T \end{cases}$ so $\|\phi_2\| = 1$

So, when misclassify a, $|\phi_2(a) - \psi_2(a)| \geq \frac{1}{\sqrt{n}}$

$\Rightarrow$ if misclassify k, $\|\phi_2 - \psi_2\| \geq \sqrt{\frac{k}{n}}$

As $\|\phi_2 - \psi_2\| \leq \sqrt{2} \sin \theta$, get

$$\sin \theta \geq \sqrt{\frac{k}{2n}}$$

By Davis-Kahan, $\sin \theta \leq \dfrac{2\|R\|}{\frac{n}{2}(p-q)}$

$$\Rightarrow \sqrt{\frac{k}{2n}} \leq \frac{4\|R\|}{n(p-q)} \leq \frac{12\sqrt{p(1-p)n}}{n(p-q)}$$

$$\sqrt{k} \leq \sqrt{2} \cdot 12 \cdot \frac{\sqrt{p(1-p)}}{p-q}$$

$$k \leq 288 \cdot \frac{p(1-p)}{(p-q)^2}$$

If p and q are constants and n grows, probably only missclassify a constant #.

If $p = \frac{1}{2}$, $q = \frac{1}{2} - \frac{17}{\sqrt{n}}$

$$k \le 288 \cdot \frac{1}{4} \cdot \frac{n}{17^2} \le \frac{n}{4}.$$

So, missclassify at most a constant fraction.