Lecture Notes Spring 2006 M690I: Extremal Graph Theory

Scribe: Olga Pryporova

Week 11(T): March 28, 2006

Proposition 1 *Let $(X, Y; E)$ be a bipartite graph with density d, and let $X' \subseteq X$, $Y' \subseteq Y$. If $d, \delta_1, \delta_2, \delta_3$ satisfy the following:*

- $\displaystyle \frac{1}{|Y|} \sum_{y \in Y} \left(\frac{|N(y)|}{|X|} - d \right)^2 \leq \delta_1$
- $\displaystyle \left| \left\{ (y_1, y_2) \in Y \times Y : \left| \frac{|N(y_1) \cap N(y_2)|}{|X|} - \frac{|N(y_1)|}{|X|} \frac{|N(y_2)|}{|X|} \right| \geq \delta_2 \right\} \right| \leq \delta_3 |Y|^2$

then with $\delta_4 = \delta_1 + (\delta_2^2 + \delta_3)^{1/2}$:

$$|d(X', Y') - d| \leq \left(\delta_4 \frac{|X||Y|}{|X'||Y'|} \right)^{1/2}$$

Corollary 1 *Let $\epsilon > 0$ be given, and let $|X'| \geq \epsilon |X|$, $|Y'| \geq \epsilon |Y|$. If $\delta_1, \delta_2, \delta_3$ (satisfying the conditions of the proposition 1) are small enough, so that $\delta_4 = \delta_1 + (\delta_2^2 + \delta_3)^{1/2} \leq \epsilon^4$, then (X, Y) is $\epsilon-$regular.*

Proof: By the proposition 1

$$|d(X', Y') - d| \leq \left(\epsilon^4 \frac{|X||Y|}{\epsilon^2 |X||Y|} \right)^{1/2} = \epsilon.$$

Therefore, (X, Y) is $\epsilon-$regular. ∎

Definition 1 (Frobenius Norm) *Given a matrix N*

$$\|N\|_F = \left(\sum_{i,j} N_{i,j}^2 \right)^{1/2}$$

Definition 2 (L^2-norm) $\|N\|_2 = \max_{\|\overrightarrow{x}\|_2 = 1} \|N \overrightarrow{x}\|_2 = \max_{\|\overrightarrow{x}\|_2 = \|\overrightarrow{y}\|_2 = 1} |\overrightarrow{y}^T N \overrightarrow{x}|$

Definition 3 (Trace) *Let N be $n \times n$ matrix. The trace $tr(N) = \sum_i (N)_{i,i}$*

Proposition 2 *Let S be a symmetric matrix and $\lambda_{\max}$ its largest eigenvalue. Then $\lambda^2_{\max}(S) \leq \|S\|^2_F$*

Proof:

$\|S\|^2_F = \sum_{i,j}(S)^2_{i,j} = \sum_{i,j}(S)_{i,j}(S)_{j,i} = \sum_i \left[\sum_j (S)_{i,j}(S)_{j,i}\right] = \sum_i (S^2)_{i,j} = tr(S^2) = \sum_i \lambda^2_i(S) \geq \lambda^2_{\max}(S)$ ∎

Proposition 3 *Let N be an $m \times n$ real matrix. Then $\lambda^2_{\max}(N^TN) \leq \|N^TN\|^2_F$*

Proposition 4 *Let N be an $m \times n$ real matrix. Then $\|N\|^2_2 = \lambda_{max}(N^TN)$.*

Proof:

$$\begin{aligned} |N\|^2_2 &= \max_{\|\overrightarrow{x}\|_2=1} \|N\overrightarrow{x}\|^2_2 = \max_{\|\overrightarrow{x}\|_2=1} \langle N\overrightarrow{x}, N\overrightarrow{x}\rangle \\ &= \max_{\|\overrightarrow{x}\|_2=1} |(N\overrightarrow{x})^T(N\overrightarrow{x})| = \max_{\|\overrightarrow{x}\|_2=1} |\overrightarrow{x}^TN^TN\overrightarrow{x}| \end{aligned}$$

Let $\{\overrightarrow{v}_1, \dots \overrightarrow{v}_n\}$ be an orthonormal eigen basis of N^TN.
Then $\overrightarrow{x} = \sum_i a_i\overrightarrow{v}_i$, where $\sum_i a^2_i = 1$.

$$\begin{aligned} |\overrightarrow{x}^TN^TN\overrightarrow{x}| &= \left|\overrightarrow{x}^T\sum_i a_i\lambda_i\overrightarrow{v}_i\right| = \left|\sum_i a^2_i\lambda_i\overrightarrow{v}^T_i\overrightarrow{v}_i\right| \\ &= \sum_i a^2_i|\lambda_i| \leq \lambda_{\max}\sum_i a^2_i = \lambda_{\max} \end{aligned}$$

(since N^TN is positive semidefinite, all its eigenvalues λ_i's are nonnegative)
To prove that $\|N\|^2_2 = \lambda_{max}(N^TN)$, take $\overrightarrow{x} = \overrightarrow{v}_j$, where $\overrightarrow{v}_j$ is an element of the orthonormal eigen basis that corresponds to the eigenvalue $\lambda_{\max}$.
Then

$$|\overrightarrow{x}^TN^TN\overrightarrow{x}| = |\overrightarrow{v_i}^T\lambda_{\max}\overrightarrow{v_i}| = \lambda_{\max}\|v_i\| = \lambda_{\max},$$

and since

$$\|N\|^2_2 = \max_{\|\overrightarrow{x}\|_2=1} |\overrightarrow{x}^TN^TN\overrightarrow{x}| \geq \lambda_{\max}$$

we have $\|N\|^2_2 = 1$ ∎

Proof of Proposition 1: Consider the bipartite adjacency matrix $A = (a_{i,j})$ of a bipartite graph $(X, Y; E)$:

$$a_{i,j} = \begin{cases} 1 & \text{if} \quad x_i \sim y_j \\ 0 & \text{otherwise} \end{cases}$$

Let $D = A - dJ$, where J is $|X| \times |Y|$ matrix with all entries equal to 1.
For any subsets X_1, $X_2 \subseteq X$ define the covariance:

$$\text{cov}(X_1, X_2) = \frac{|X_1 \cap X_2|}{|X|} - \frac{|X_1|}{|X|} \cdot \frac{|X_2|}{|X|}$$

$$\begin{aligned}
(D^T D)_{y_1,y_2} &= \left((A - dJ)^T (A - dJ)\right)_{y_1,y_2} \\
&= \left(A^T A\right)_{y_1,y_2} - d\left(J^T A\right)_{y_1,y_2} - d\left(A^T J\right)_{y_1,y_2} + d^2\left(J^T J\right)_{y_1,y_2} \\
&= \left(|N(y_1) \cap N(y_2)| - \frac{|N(y_1)||N(y_2)|}{|X|}\right) + \frac{|N(y_1)||N(y_2)|}{|X|} \\
&\quad -d|N(y_2)| - d|N(y_1)| + d^2|X| \\
&= |X|\text{cov}(N(y_1), N(y_2)) \\
&\quad +|X|\left[\frac{|N(y_1)|}{|X|} \cdot \frac{|N(y_2)|}{|X|} - d\frac{|N(y_1)|}{|X|} - d\frac{|N(y_2)|}{|X|} + d^2\right] \\
&= |X|\text{cov}(N(y_1), N(y_2)) + |X|\left(\frac{|N(y_1)|}{|X|} - d\right)\left(\frac{|N(y_2)|}{|X|} - d\right)
\end{aligned}$$

$$\begin{aligned}
\|D^T D\|_F^2 &= \sum_{y_1,y_2}\left[|X|\text{cov}(N(y_1), N(y_2)) + |X|\left(\frac{|N(y_1)|}{|X|} - d\right)\left(\frac{|N(y_2)|}{|X|} - d\right)\right]^2 \\
&= |X|^2 \sum_{y_1,y_2} \left[A_{y_1,y_2} + B_{y_1,y_2}\right]^2
\end{aligned}$$

Where

$$\begin{aligned}
A_{y_1,y_2} &= \text{cov}(N(y_1), N(y_2)) \\
B_{y_1,y_2} &= \left(\frac{|N(y_1)|}{|X|} - d\right)\left(\frac{|N(y_2)|}{|X|} - d\right)
\end{aligned}$$

$$\begin{aligned}
\|D^T D\|_F &= |X|\left(\sum_{y_1,y_2}\left[A_{y_1,y_2} + B_{y_1,y_2}\right]^2\right)^{1/2} \\
&\le |X|\left[\left(\sum_{y_1,y_2} A_{y_1,y_2}^2\right)^{1/2} + \left(\sum_{y_1,y_2} B_{y_1,y_2}^2\right)^{1/2}\right]
\end{aligned}$$

(by the triangle inequality)

Note that

$$\left(\sum_{y_1,y_2} B_{y_1,y_2}^2\right)^{1/2} = \left(\sum_{y_1,y_2} \left(\frac{|N(y_1)|}{|X|} - d\right)^2 \left(\frac{|N(y_2)|}{|X|} - d\right)^2\right)^{1/2} =$$

$$= \left(\sum_{y_1}\sum_{y_2} \left(\frac{|N(y_1)|}{|X|} - d\right)^2 \left(\frac{|N(y_2)|}{|X|} - d\right)^2\right)^{1/2} =$$

$$= \sum_{y} \left(\frac{|N(y)|}{|X|} - d\right)^2$$

Therefore

$$\frac{\|D^T D\|_F}{|X|} \le \left[\sum_{y_1,y_2} \operatorname{cov}^2(N(y_1), N(y_2))\right]^{1/2} + \sum_{y} \left(\frac{|N(y)|}{|X|} - d\right)^2 \le$$

$$\le \left[\sum_{y_1,y_2} \operatorname{cov}^2(N(y_1), N(y_2))\right]^{1/2} + \delta_1|Y|$$

Recall that $|\{(y_1, y_2) \in Y \times Y : |\operatorname{cov}(N(y_1), N(y_2))| \ge \delta_2\}| \le \delta_3|Y|^2$.
Thus

$$\frac{\|D^T D\|_F}{|X|} \le \left[\delta_2^2|Y|^2 + 1 \cdot \delta_3|Y|^2\right]^{1/2} + \delta_1|Y| = \delta_4|Y|$$

So $\|D^T D\|_F \le \delta_4|X||Y|$.

Let $\overrightarrow{x}'$ and $\overrightarrow{y}'$ be the indicator vectors for the subsets X' and Y' respectively, i.e.

$$(\overrightarrow{x}')_i = \begin{cases} 1 & \text{if} \quad x_i \in X' \\ 0 & \text{otherwise} \end{cases}$$

$$(\overrightarrow{y}')_i = \begin{cases} 1 & \text{if} \quad y_i \in Y' \\ 0 & \text{otherwise} \end{cases}$$

Then

$$|e(X', Y') - d|X'||Y'|| = |\overrightarrow{x}'^T D \overrightarrow{y}'| \le \|D\|_2 \|\overrightarrow{x}'\|_2 \|\overrightarrow{y}'\|_2 \le$$

$$\le \sqrt{\|D^T D\|_F}\, \|\overrightarrow{x}'\|_2 \|\overrightarrow{y}'\|_2 \le$$

$$\le \sqrt{\delta_4|X||Y|}\sqrt{|X'||Y'|}$$

Therefore

$$|d(X', Y') - d| \leq \left(\delta_4 \frac{|X||Y|}{|X'||Y'|} \right)^{1/2}$$

■