1 Lecture Outline

- Testing properties of dense graphs: Triangle-freeness.

2 The Goal

The goal of today's lecture is to find a property tester for dense graphs for the "Triangle free" property. We will work with a graph $G = (V, E)$ given in the adjacency matrix model. A is $n \times n$ matrix representing G.

Definition 1

$$G \text{ is triangle free } if \nexists x, y, z \in V \, such \; that \; A(x,y) = A(y,z) = A(x,z) = 1.$$

Our goal is to find a tester which runs in sublinear time, which always accepts triangle-free graphs, and rejects with high probability graphs that are ϵ-far from having the triangle-freeness property.

Before we begin, some definitions:
Let $A, B \subset V$ such that $A \cap B = \emptyset$, $|A|, |B| \geq 2$.

Definition 2 *The* number of edges between A and B

$$e(A, B) \equiv |\{(u, v) : u \in A, b \in B\}|.$$

Definition 3 *The* density between A and B

$$d(A, B) = \frac{d(A, B)}{|A||B|}.$$

3 Thought experiment

Let G be a random tripartite graph for which:

$$\text{For every two nodes } u \in V_i, v \in V_j \Pr[A(u, v) = 1] = \eta$$

3.1 How many triangles do we have in G?

Let $\sigma(u,v,w) = \begin{cases} 1 & \text{if } (u,v,w) \text{ form a triangle} \\ 0 & \text{otherwise} \end{cases}$

then,

$$\forall u \in V_1, v \in V_2, w \in V_3, Pr[\sigma(u,v,w) = 1] = \eta^3$$

Thus

$$\mathrm{E}[\sigma(u,v,w)] = \eta^3.$$

From the linearity of the expectency we get

$$\mathrm{E}[\text{number of triangles in G}] = \eta^3|A||B||C|$$

4 The Komlos-Simonovits Lemma.

Definition 4 *Let $G = (V,E)$ be a graph, and let $A, B \subset V$. We say that (A,B) are γ-regular if*

$$\forall A' \subset A, \forall B' \subset B \textit{ such that } |A'| \geq \gamma|A| \textit{ and } |B| \geq \gamma|B|$$

$$|d(A',B') - d(A,B)| < \gamma.$$

Notice that the regularity property is very similar to the randomness property. Intuitivily, it expresses an important property of random graphs: Given two sets of vertices A, B in a random graph, the expected fraction of edges between subsets of A, B is identical to the expected fraction of edges between A, B. This ensures a uniformity in the distribution of the edges.

Lemma 5 *Komlos-Simonovits' Lemma.*

$$\forall \eta > 0, \exists \gamma = \gamma^{\triangle}(\eta) = \frac{\eta}{2}, \exists \delta = \delta^{\triangle}(\eta) = (1-\eta)\frac{\eta^3}{8}$$

such that if A, B, C are disjoint subsets of V and each pair of them is γ- regular with density $\geq \eta$,
then G contains at least $\gamma|A||B||C|$ distinct triangles, each one of them with one vertex from each one of A,B,C.

Notice that this number is very similar to the number of triangles in three partite random graph. For example, if $\eta \leq \frac{1}{2}$ then $\delta(1-\eta)\frac{\eta^3}{8} \geq \frac{\eta^3}{16}$, so there is only a constant difference between the number of triangles in this graph and the expected number of triangles in the three partite random graph. Before we prove the lemma we need to prove the following claim.

Definition 6

$A^* = \{$*nodes in A with at least $(\eta - \gamma)|B|$ neighbors in B and at least $(\eta - \gamma)|C|$ neighbors in C*$\}$.

Claim 7 $|A^*| \geq (1-2\gamma)|A|$.

Proof Define

$$A' = \{\text{bad nodes in A with respect to B}\} \tag{1}$$

Formally,

$$A' = \{u \in A : |\{v \in B : A(u,v) = 1\}| < (\eta - \gamma)|B|\} \tag{2}$$

We show that,

$$|A'| \leq \gamma |A|.$$

Indeed, suppose for contradiction that

$$|A'| > \gamma |A|. \tag{3}$$

then

$$d(A', B) = \frac{e(A', B)}{|A'||B|} < \frac{\overbrace{(\eta-\gamma)|A'||B|}^{\text{because of the definition of A'}}}{|A'||B|} = \eta - \gamma. \tag{4}$$

and from this we have:

$$|d(A,B) - d(A',B)| > \eta - (\eta - \gamma) = \gamma. \tag{5}$$

We have:
$d(A,B) \geq \eta$ because (A,B) has density $\geq \eta$.
$d(A',B) < \eta - \gamma$ because the definition of A'.
And this is a contradiction to the regularity assumption. Therefore

$$|A'| \leq \gamma |A|. \tag{6}$$

Similarly if

$$A'' = \{\text{bad nodes in A with respect to C}\} \tag{7}$$

Then

$$|A''| \leq \gamma |A|. \tag{8}$$

Finally, from (8) and (6):

$$A^* = A \backslash (A' \cup A'') \Rightarrow |A*| \geq |A| - 2\gamma|A| = (1-2\gamma)|A|. \tag{9}$$

■

Now we are ready to prove Komlos-Simonovits' Lemma:

Proof
Let

$$\gamma^{\triangle}(\eta) = \frac{\eta}{2}$$

and

$$\delta^{\triangle}(\eta) = (1-\eta)\frac{\eta^3}{8}$$

Definition 8 *For each $u \in A^*$, define:*

$$B_u = \{\textit{neighbors of } u \textit{ in } B\}$$

and

$$C_u = \{\textit{neighbors of } u \textit{ in } C\}$$

Note that

$$|B_u| \geq (\eta - \gamma)|B| \textit{ and } |C_u| \geq (\eta - \gamma)|C|.\textit{(by definition of } A^*\textit{)}$$

Observation 9 *Notice that the number of edges between B_u and C_u gives exactly the number of distinct triangles with u as one of their vertices.*

We know that

$$d(B, C) \geq \eta \qquad \text{(because they have density } \eta\text{)}. \tag{10}$$

$$|B_u| \geq (\eta - \gamma)|B| \overbrace{\geq}^{\text{because } \gamma = \frac{\eta}{2}} \gamma|B| \tag{11}$$

$$|C_u| \geq (\eta - \gamma)|C| \overbrace{\geq}^{\text{because } \gamma = \frac{\eta}{2}} \gamma|C| \tag{12}$$

From regularity we have

$$d(B_u, C_u) \geq |d(B, C) - \gamma| \geq \eta - \gamma. \tag{13}$$

And from this we get

$$e(B_u, C_u) > (\eta - \gamma)|B_u||C_u| \geq (\eta - \gamma)^3|B||C|. \tag{14}$$

Therefore the total number of triangles is:

$$\#\text{triangles} = |A^*|(\eta - \gamma)^3|B||C|. \tag{15}$$

From the claim we have that

$$|A^*| \geq (1 - 2\gamma)|A|. \tag{16}$$

So

$$\#\text{triangles} \geq (1 - 2\gamma)|A|(\eta - \gamma)^3|B||C| \geq (1 - \eta)\frac{\eta^3}{8}|A||B||C|. \tag{17}$$

■

5 Szemerédi Regularity Lemma

Lemma 10 *$\forall m, \forall \epsilon' > 0, \exists T$ such that the following holds:*
Given a graph $G = (V, E)$ for which $|V| > T$ and an equi-partition A of V into m subsets, there exists an equi-partition B of V into k subsets such that:

- *B is a refinement of A*
- *$m \leq k \leq T$*
- *At most $\epsilon' \binom{k}{2}$ subset pairs are not ϵ'-regular.*

5.1 Notes

- We don't prove the lemma here.
- There is a technical issue with the factorization of n and k. In some versions, a "garbage can" subset is added in order to solve this problem. Here, for the sake of simplicity we assume $k|n$.
- T can be very large. In fact it can be as high as a tower function whose height depends on ϵ'. However, it doesn't depend on $|V|$, so we consider it as a (unfortunately large) constant. Furthermore, a lower bound has been shown, i.e. for certain graphs T must be a tower function.

6 Algorithm for testing triangle freeness

Definition 11 *Let $G = (V, E)$ be a graph given in the adjacency matrix model. We say that G is ϵ-*far from being triangle-free *if more than ϵn^2 edges need to be removed from G to make it triangle-free.*

Theorem 12 *$\forall \epsilon, \exists \delta$ such that every graph $G = (V, E)$ that is ϵ-far from being triangle free has at least $\delta \binom{n}{3}$ distinct triangles.*

Proof Let A be an arbitrary equipartition with $\frac{\epsilon}{5}$ subsets. By Szemeredi's regularity lemma, for $m = \frac{\epsilon}{5}$ and $\epsilon' = \frac{\epsilon}{10}$, there is a number $T = T(m, \epsilon') = T(\epsilon)$ and a partition B into $V_1, V_2, ...V_k$ such that at most $\epsilon' \binom{k}{2}$ set-pairs are not $\epsilon' - regular$ and:

$$\frac{5}{\epsilon} \leq k \leq T. \tag{18}$$

Equation (18) is equivalent to:

$$\frac{\epsilon n}{5} \geq \frac{n}{k} \geq \frac{n}{T} \tag{19}$$

Note that because B is an equipartition

$$|V_1| = |V_2| = ... = |V_n| = \frac{n}{k} \tag{20}$$

We define a procedure to obtain a new graph G' from G by removing less than ϵn^2 edges:

- Delete all internal edges in G, i.e. edges (u, v) such that u and v belong to the same V_i. There are at most $(\frac{n}{k})^2$ edges in each V_i, and $i = 1...k$, so at most $\frac{n^2}{k}$ edges are removed. By equation (19) this is $\leq n\frac{\epsilon n}{5} = \frac{\epsilon}{5}n^2$
- Delete edges between non-ϵ'-regular pairs. There are $\leq \epsilon'\binom{k}{2}$ such pairs, and between each pair there are at most $(\frac{n}{k})^2$ edges. So, at most $\epsilon'\frac{k(k-1)}{2}\frac{n^2}{k^2} \leq \frac{\epsilon'}{2}n^2 = \frac{\epsilon}{20}n^2 < \frac{\epsilon}{5}n^2$ edges are removed.
- Delete edges between pairs (V_i, V_j) whose density is less than $\frac{\epsilon}{5}$. The number of such edges is certainly less than $\frac{\epsilon}{5}n^2$

In all these stages together we removed $\frac{3}{5}\epsilon n^2 < \epsilon n^2$ edges. Since G is ϵ-far from being triangle-free, G' surely has at least one triangle. Furthermore, because we removed the internal edges, each triangle has its vertices in three distinct partition subsets. Let (x, y, z) be a triangle in G', and let (V_i, V_j, V_k) be the corresponding partition subsets. From the construction of G', we know that:

- $(V_i, V_j), (V_j, V_k), (V_i, V_k)$ are ϵ' regular
- $(V_i, V_j), (V_j, V_k), (V_i, V_k)$ have a density of at least $\frac{\epsilon}{5}$

From the Komlos-Simonovits lemma we conclude that there are at least $\delta'|V_i||V_j||V_k| = \delta'\frac{n^3}{k^3}$ triangles in G'. From equation (19) we know that G' has at least $\delta'\frac{n^3}{T^3}$ triangles. If we define $\delta = \frac{6\delta'}{T}$ then G has at least $\delta\binom{n}{3}$ triangles. Since every triangle of G' is also a triangle of G (We only removed edges), then G has at least $\delta\binom{n}{3}$ triangles as required.
■

Corollary 13 *Let G be a graph in the adjacency matrix model. The following algorithm is an one-sided tester for triangle-freeness:*

- *Repeat $O(\frac{1}{\delta})$ times:*
  - *If $(v1, v2, v3)$ is a triangle in G: Reject.*
- *Accept*

Proof If G is triangle-free then a triangle will never be found and we return true as required. If G is ϵ-far from being triangle-free, then by the theorem there are at least $\delta\binom{n}{3}$ triangles. Therefore:
$Pr[\text{no triangle was found}] \leq (1-\delta)^{O(\frac{1}{\delta})} \leq e^{-c} \leq \frac{1}{3}$
Where c is a constant we choose such that the last inequality holds. ■