6.5240 Sublinear Time Algorithms November 10, 2026

Lecture 16

Lecturer: Ronitt Rubinfeld *Scribe: Hyeongjoe Chu*

1 Szemerédi's Regularity Lemma

We begin with Szemerédi's Regularity Lemma, a powerful result from combinatorics. We will use this result to analyze a tester for triangle-freeness of dense graphs. This lemma is one of the so-called "Heavy Hammers".

Motivation: triangles in a random graph. Let A, B, C be disjoint groups of vertices. Include each possible edge independently with probability η. For any $u \in A$, $v \in B$, and $w \in C$,

$$\Pr[uvw \text{ forms a triangle}] = \eta^3.$$

Thus, if X is the number of triangles with one vertex in each group,

$$\mathbb{E}[X] = \eta^3 |A||B||C|.$$

Regularity gives a useful analogue of this behavior for deterministic graphs.

1.1 Density and regularity

Definition 1 (Density and regular pairs) Let $A, B \subseteq V$ be disjoint, nonempty sets. Write $e(A, B)$ for the number of edges between A and B. Their *density* is

$$d(A, B) \stackrel{\text{def}}{=} \frac{e(A, B)}{|A||B|}.$$

The pair (A, B) is γ-*regular* if, for every $A' \subseteq A$ and $B' \subseteq B$ satisfying

$$|A'| \geq \gamma |A|, \qquad |B'| \geq \gamma |B|,$$

we have

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

Intuitively, this definition "generalizes" random graphs (because in a random graph, the expected density $d(A, B)$ would be η). The size conditions are essential: if A' and B' each contain only one vertex, their density is always either 0 or 1.

1.2 Counting triangles in regular pairs

The next lemma shows that three dense regular pairs contain many triangles.

Lemma 2 *For every $0 < \eta < 1$, there exist $\gamma, \delta > 0$, depending only on η, with the following property. If A, B, C are disjoint, nonempty subsets of V, and each of (A, B), (A, C), and (B, C) is γ-regular with density greater than η, then G contains at least $\delta|A||B||C|$ triangles with one vertex in each set. In particular, we may take*

$$\gamma = \frac{\eta}{2}, \qquad \delta = \frac{(1-\eta)\eta^3}{8}.$$

Proof of Lemma 2. Fix $\gamma = \eta/2$. Let A^* consist of the vertices in A with at least $(\eta - \gamma)|B|$ neighbors in B and at least $(\eta - \gamma)|C|$ neighbors in C.

Claim 3 $|A^*| \geq (1 - 2\gamma)|A|$.

To prove the claim, let A' be the vertices in A with fewer than $(\eta - \gamma)|B|$ neighbors in B. Suppose that $|A'| \geq \gamma|A|$. Then

$$d(A', B) = \frac{e(A', B)}{|A'||B|} < \eta - \gamma.$$

Since $d(A, B) > \eta$, this implies $|d(A', B) - d(A, B)| > \gamma$, contradicting regularity. Hence $|A'| < \gamma|A|$.

Similarly, the set A'' of vertices with fewer than $(\eta - \gamma)|C|$ neighbors in C satisfies $|A''| < \gamma|A|$. It follows that

$$|A^*| = |A \setminus (A' \cup A'')| \geq |A| - |A'| - |A''| \geq (1 - 2\gamma)|A|,$$

which proves the claim.

Now fix $u \in A^*$, and let $B_u \subseteq B$ and $C_u \subseteq C$ be its neighborhoods. By definition,

$$|B_u| \geq (\eta - \gamma)|B| = \gamma|B|, \qquad |C_u| \geq (\eta - \gamma)|C| = \gamma|C|.$$

Regularity of (B, C) gives $d(B_u, C_u) > \eta - \gamma$. Each edge between B_u and C_u forms a triangle with u, so

$$e(B_u, C_u) \geq (\eta - \gamma)|B_u||C_u| \geq (\eta - \gamma)^3|B||C|.$$

Summing over $u \in A^*$, with no double counting, yields

$$\begin{aligned} \#\{\text{triangles across } A, B, C\} &\geq |A^*|(\eta - \gamma)^3|B||C| \\ &\geq (1 - 2\gamma)(\eta - \gamma)^3|A||B||C| \\ &= \frac{(1-\eta)\eta^3}{8}|A||B||C|. \end{aligned}$$

■

1.3 The regularity lemma

Now we are ready for the "Heavy Hammer" (Although we will skip the proof of the lemma).

Theorem 4 (Szemerédi's Regularity Lemma) *For every positive integer m and every $\varepsilon > 0$, there exists $T = T(m, \varepsilon)$ such that the following holds. Given a graph $G = (V, E)$ with $|V| > T$ and an equipartition $\mathcal{A}$ of V into m sets, there exists an equipartition $\mathcal{B}$ into k sets refining $\mathcal{A}$ such that $m \leq k \leq T$, and at most $\varepsilon\binom{k}{2}$ pairs of distinct parts are not ε-regular.*

Remark This lemma was introduced by Szemerédi for his theorem on arithmetic progressions in 1975. There is a deep connection between graph theory and additive combinatorics,

2 Testing Triangle-Freeness

We conclude by outlining the application to property testing. The remaining details will be covered in the next lecture.

Testing problem. Given a graph G on n vertices in the adjacency matrix representation, determine whether it is triangle-free. We say that G is *ε-far from triangle-free* if more than εn^2 edges must be deleted to make it triangle-free. The tester should satisfy:

- If G is triangle-free, it accepts with probability 1.
- If G is ε-far from triangle-free, it rejects with probability greater than $2/3$.

Tester. Let $\delta = \delta(\varepsilon)$ be the constant in Theorem 5.

1. Repeat $O(1/\delta)$ times:
   (a) Sample three distinct vertices uniformly at random.
   (b) Query the three possible edges. If they form a triangle, reject and halt.
2. Accept.

If $n < 3$, accept immediately.

Theorem 5 *For every $\varepsilon > 0$, there exists $\delta > 0$, depending only on ε, such that every n-vertex graph that is ε-far from triangle-free contains at least $\delta\binom{n}{3}$ distinct triangles.*

Consequently, each trial finds a triangle with probability at least δ when G is ε-far from triangle-free. After t independent trials, the probability of missing every triangle is at most

$$(1-\delta)^t \le e^{-\delta t}.$$

Taking $t > \log(3)/\delta$ gives the required rejection probability.