Lecture 5. Szemerédi Regularity Lemma

Hong Liu

31st March 2021

Szemerédi's regularity lemma is one of the most important tools in extremal graph theory dealing with dense graphs (positive edge-density). Here we give a gentle introduction to this powerful lemma and see some of its applications and other classical results related to it.

Roughly speaking, the regularity lemma states that every large graph admits a partition into bounded number of parts such that between almost all pairs of parts, the induced bipartite subgraphs behave pseudorandomly. The essence of the regularity lemma is:

Approximating large structures by small structures with low complexity.

It usually offers conceptually simple proofs for asymptotic results. For instance, the regularity lemma and its counting lemma together imply that, in terms of subgraph densities, any graph can be approximated by one of the few (weighted) graphs with bounded order (reduced graphs on $O_\varepsilon(1)$ vertices).

1 Informally...

To state the regularity lemma rigorously, we need to set up several notions. Before we do so, let us informally describe a common way of applying the regularity lemma:

- Step 1. Reduce an extremal problem A on large graphs to a problem B on small weighted graphs (using the random behaviour of the regular partition, embedding lemma, counting lemma etc.);
- Step 2. Solve problem B (using e.g. classical results in graph theory).

Let us recall the proof sketch for Erdős-Simonovits-Stone theorem that

$$\mathrm{ex}(n,H) \le \left(1 - \frac{1}{\chi(H)-1} + o(1)\right)\frac{n^2}{2}.$$

Step 1 above in this case can be done using the following consequence of the regularity lemma and counting lemma: for any graph G, there is a (weighted reduced) graph R on $O(1)$ vertices such that

P1 for any fixed H, the subgraph density of H in R is roughly the same as that in G;[1]

P2 if R contains a homomorphic copy of H, then G contains a copy of H.

By subgraphs densities, we mean the number of (not necessarily induced) copy of H in G, normalised by $\binom{|G|}{|H|}$, which can be viewed as the probability of a uniform chosen random $|H|$-set from $V(G)$ inducing a graph containing H.

[1]As we shall see in counting lemma, more precisely, here by subgraph density in R, we mean the weighted subgraph homomorphism density.

Informal proof of Erdős-Simonovits-Stone theorem. *Step 1*. Let $r := \chi(H) - 1$. By **P1** with $H = K_2$, we just need to bound the edge-density of R from above by $1 - \frac{1}{r}$.
Step 2. Note that K_{r+1} is a homomorphic image of H. Then by **P2**, R is K_{r+1}-free. The desired bound on edge-density then follows from Turán's theorem. □

Before we dive into the details, let us point out a comprehensive survey of Komlós-Simonovits on regularity lemma.[2]

2 Formal setup

The basic notion in regularity lemma is that of an *ε-regular pair* which measures the pseudorandomness/regularity of the induced bipartite subgraph between the pair. The parameter ε is the *precision* of the regularity; the smaller ε is, the more random like the pair is.

Definition 2.1 (Regular pair)**.** Given $G = (V, E)$ and disjoint vertex subsets $X, Y \subseteq V$, let $e(X, Y) := e(G[X, Y])$ and denote by

$$d(X, Y) := \frac{e(X, Y)}{|X||Y|}$$

the *density* of the pair (X, Y). For $\varepsilon > 0$, the pair (X, Y) is *ε-regular* if for any $A \subseteq X, B \subseteq Y$ with $|A| \geq \varepsilon|X|$, $|B| \geq \varepsilon|Y|$, satisfy

$$|d(A, B) - d(X, Y)| < \varepsilon.$$

Additionally, if $d(X, Y) \geq \delta$, for some $\delta > 0$, we say that (X, Y) is *(ε, δ)-regular.*

In other words, a regular pair (X, Y) has "uniform" edge distribution in the sense that the density of any pair of large (ε-proportion) subsets (A, B) is roughly the same as that of (X, Y).

Definition 2.2 (Regular partition)**.** A partition $V = V_0 \cup V_1 \cup \cdots \cup V_r$ is *ε-regular*, if
(*i*) $|V_0| \leq \varepsilon|V|$; (called *exceptional set*)
(*ii*) $|V_1| = |V_2| = \cdots = |V_r|$;
(*iii*) all but εr^2 pairs (V_i, V_j) with $1 \leq i < j \leq r$ are ε-regular.

It is worth making a couple of quick remarks.

- We do not assume that V_i, $i \in [r]$, is larger than the exceptional set V_0. In fact, quite the contrary, most of the time, we take $r \geq m \geq 1/\varepsilon$ to make the edges in V_i negligible.
- In the definition of regular partition, we can also have no exceptional set (by distributing V_0 equally to other parts) and instead have $||V_i| - |V_j|| \leq 1$ for all $1 \leq i \leq j \leq r$.
- There is a degree version in which we can require that for each $i \in [r]$, all but εr pairs involving V_i are ε-regular.

We will use mostly the version of regular partition with no exceptional set V_0, unless otherwise specified.

We can now state the lemma.

Theorem 2.3 (Szemerédi regularity lemma 1976)**.** *Given $\varepsilon > 0$ and $m \in \mathbb{N}$, there exists $M = M(\varepsilon, m)$, such that any graph G admits an ε-regular partition $V = V_0 \cup V_1 \cup \cdots \cup V_r$ with $m \leq r \leq M$.*

[2]Komlós and Simonovits, *Szemerédi's regularity lemma and its applications to graph theory*, Bolyai Math. Soc., (1996).

Remark 2.4. Let us make some remarks about the parameters in the regularity lemma.

- We usually think of ε in the regularity lemma as a very small constant, i.e. $o(1)$.
- Both the lower and upper bounds $m \le r \le M$ on the number of parts of the partition are meaningful. If there is no lower bound, then the trivial partition $V = V$ consisting of just one part is vacuously a regular partition and clearly this partition is of no use for us. The upper bound on r is also needed as we shall see shortly, the proof of the counting lemma relies crucially on the fact that the reduced graph R we use to approximate the original graph G is of bounded order.
- If the graph G does not have positive edge-density, then the regularity lemma does not say much about G.
- The εr^2 exceptional irregular pairs are needed. Consider the following example:

  Half graph. $G = (A \cup B, E)$, where $A = B = [n]$. For any $a \in A$ and $b \in B$, put $ab \in E(G)$ if and only if $a \ge b$. Notice that $d(A, B) = 1/2$. Let the top half of A be X and bottom half of B be Y, then $d(X, Y) = 0$, while $d(A - X, B - Y) = 1$. There are εr irregular pairs in any partition.
- The upper bound on the size of the partition M coming from the proof of regularity lemma is rather large, it is a tower of 2s with height $2\varepsilon^{-5}$. Gowers gave a construction showing that a tower of 2s with height $\varepsilon^{-1/16}$ is needed.

We end this section with two simple lemmas. The first one states that between a regular dense pair, almost every vertex has the "correct" degree to any large subset of the other side.

Lemma 2.5. *Let (X, Y) be an ε-regular pair with density d, and $B \subseteq Y$ with $|B| \ge \varepsilon|Y|$, then all but $2\varepsilon|X|$ vertices in X have degree $(d \pm \varepsilon)|B|$ in B.*

Proof. Let $A \subseteq X$ be the set of vertices with "small" degree in B, i.e.

$$\frac{d(v, B)}{|B|} < d - \varepsilon.$$

Suppose that $|A| > \varepsilon|X|$, consider the pair (A, B). By the choice of A, we have

$$d(A, B) = \frac{e(A, B)}{|A||B|} < \frac{|A| \cdot (d - \varepsilon)|B|}{|A||B|} = d - \varepsilon,$$

contradicting (X, Y) being ε-regular. Thus, $|A| \le \varepsilon|X|$. Similarly, the same bound holds for the set of vertices of "large" degree, i.e. $d(v, B)/|B| > d + \varepsilon$ in B. □

Given a regular pair (X, Y), one can also show that almost all pairs from one part, say X, have the "correct" codegree to large subsets of the other side.

Exercise 2.6. Formulate the above codegree statement rigorously and prove it.

The second lemma states that regularity is inherited by large subsets of pairs (with a slightly worse precision/regularity). This lemma is useful as it implies that we can further refine a regular partition to get additional properties without losing regularity.

Lemma 2.7 (Slicing lemma)**.** *Let $V_0 \cup V_1 \cup \cdots \cup V_r$ be an ε-regular partition. Further refine each part into s equal parts: $V_i = V_i^1 \cup \cdots \cup V_i^s$. The new partition (with $sr + 1$ parts) is $O(s\varepsilon)$-regular.*[3]

Exercise 2.8. Prove the slicing lemma.

[3]Note that $O(s\varepsilon)$-regular implicitly requires that in the slicing lemma, $s \ll 1/\varepsilon$.