Walk through Combinatorics: Regularity lemma*

The regularity lemma was conceived to prove Szemerédi's theorem on k-term arithmetic progressions, but has since grown into a powerful principle that applies to a multitude of mathematical objects. Here we restrict to graphs, which is the setting of the original Szemerédi's regularity lemma. The key notion is that of a density:

Definition. *Let $G = (V, E)$ be a graph, and $A, B \subset V$ be disjoint sets of vertices. The* edge density between A and B *is defined by*

$$d(A, B) = e(A, B)/|A||B|,$$

where $e(A, B)$ is the number of edges between A and B.

In other words, the edge density is the probability that a randomly chosen pair of vertices $(a, b) \in A \times B$ spans an edge.

For the motivation we consider a random bipartite graph with bipartition (A, B) such that each pair $(a, b) \in A \times B$ spans an edge with probability p with different edges being independent. It is easy to see that $d(A, B) \approx p$ with high probability; in fact a simple calculation with Chernoff's inequality and the union bound shows that $d(A', B') \approx p$ for *all* pairs of $A' \subset A$ and $B' \subset B$ that are not too small. This suggests the following definition:

Definition. *A pair $A, B \subset V(G)$ of disjoint sets is* ε-regular *if for each choice of sets $A' \subset A$ and $B' \subset B$ satisfying $|A'| \geq \varepsilon|A|$ and $|B'| \geq \varepsilon|B|$ we have*

$$\big|d(A', B') - d(A, B)\big| \leq \varepsilon.$$

The lower bound on sizes A' and B' is natural, for if it is absent, and we can take $|A'| = |B'| = 1$, the definition is overpowered.

The non-technical content of Szemerédi's regularity lemma is that every graph can be partitioned into only a few sets such that almost every pair of these sets is ε-regular. The precise statement is captured in the next definition:

Definition. *An* ε-regular partition *of a graph $G = (V, E)$ is a partition of the vertex set $V = J \cup V_1 \cup \cdots \cup V_k$ that satisfies the three conditions:*

*These notes are available from `http://www.borisbukh.org/DiscreteMath12/notes_regularity.pdf`.

1. *(Junk is small)* $|J| \le \varepsilon |V|$,
2. *(Equipartition)* $|V_1| = \cdots = |V_k|$,
3. *(Regularity) All but at most* εk^2 *pairs* (V_i, V_j) *are* ε*-regular.*

Theorem 1 (Szemerédi's regularity lemma)**.** *For every* $\varepsilon > 0$ *and every* m *there is an* M *such that each graph admits an* ε*-regular partition into* k *parts, where* $m \le k \le M$*.*

The regularity lemma does not say anything about the behavior of the edges that are wholly contained in a single part of a partition. It is thus the role of parameter m to ensure that there only a few such edges. Indeed, we have

$$\binom{|J|}{2} + \sum_i \binom{|V_i|}{2} \le \varepsilon^2 n^2 + k(n/k)^2 \le (\varepsilon^2 + 1/m)n^2.$$

The idea of the proof is simple and powerful: starting with an arbitrary partition, successively refine the partition into more and more regular partition. To measure the progress toward regularity we introduce the function

$$f(A,B) \stackrel{\text{def}}{=} \frac{|A|}{|V|} \cdot \frac{|B|}{|V|} d(A,B)^2.$$

For a partition $\mathcal{P}$ of V we define

$$f(\mathcal{P}) \stackrel{\text{def}}{=} \sum_{\substack{A,B \in P \\ A \ne B}} f(A,B).$$

[The choice of f is not unique. We could have replaced $d(A,B)^2$ by another strictly convex function of $d(A,B)$. Our choice is the simplest. For information-theoretic interpretation of the proof that uses the entropy function, see "Regularity lemma revisited" by Terence Tao.]

An important property of f is that $f(\mathcal{P}) \le 1$ for every partition $\mathcal{P}$. Indeed,

$$f(\mathcal{P}) = \sum_{\substack{A,B \in P \\ A \ne B}} \frac{|A|}{|V|} \cdot \frac{|B|}{|V|} d(A,B)^2 \le \sum_{A,B \in \mathcal{P}} \frac{|A|}{|V|} \cdot \frac{|B|}{|V|} = 1.$$

The second property is that f increases under refinements[1]. We first state this for a single pair (A,B):

Lemma 2. *Suppose* $A, B \subset V$ *are two disjoint sets, and* $A = A_1 \cup \cdots \cup A_k$ *and* $B = B_1 \cup \cdots \cup B_l$ *are their respective partitions, then*

$$\sum_{i,j} f(A_i, B_j) = f(A,B) + \sum_{i,j} \bigl(d(A,B) - d(A_i,B_j)\bigr)^2 \frac{|A_i||B_j|}{|V|^2}.$$

[1] A partition $\mathcal{P}'$ is a refinement of $\mathcal{P}$ if each $A' \in \mathcal{P}'$ is fully contained in some $A \in \mathcal{P}$

Proof. Mindless calculation gives

$$\begin{aligned}\sum_{i,j}\bigl(d(A,B)-d(A_i,B_j)\bigr)^2|A_i||B_j| &= d(A,B)^2\sum_{i,j}|A_i||B_j| - 2d(A,B)\sum_{i,j}d(A_i,B_j)|A_i||B_j| \\ &\quad + \sum_{i,j}d(A_i,B_j)^2|A_i||B_j| \\ &= d(A,B)^2|A||B| - 2d(A,B)e(A,B) \\ &\quad + \sum_{i,j}d(A_i,B_j)^2|A_i||B_j|, \\ &= -f(A,B)|V|^2 + \sum_{i,j}d(A_i,B_j)^2|A_i||B_j|. \end{aligned}$$ □

The identity in Lemma 2 is a form of Cauchy–Schwarz inequality, for the latter is the glorification of the inequality $\langle x-y, x-y\rangle \geq 0$ that is true for any vectors x and y.

Corollary 3. *Suppose $A, B \subset V$ are two disjoint sets, and $\mathcal{A}$ and $\mathcal{B}$ are partitions of A and B respectively. Suppose further than $\mathcal{A}'$ and $\mathcal{B}'$ are refinements of $\mathcal{A}$ and $\mathcal{B}$. Then*

$$\sum_{\substack{A'\in\mathcal{A}'\\ B'\in\mathcal{B}'}} f(A',B') \geq \sum_{\substack{A\in\mathcal{A}\\ B\in\mathcal{B}}} f(A,B).$$

Proof. The inequality follows by applying Lemma 2 to each pair (A,B). □

While in the definition of ε-regular partition there is only one junk part, in the proof it is convenient to use several. We will ensure that the total size of the junk is kept small, and lump all the junk parts into one at the end of the proof.

At each step of the proof we will work with a partition $\mathcal{P}$ of V, say $V = J_1 \cup \cdots \cup J_l \cup V_1 \cup \cdots \cup V_k$, satisfying $|V_1| = \cdots = |V_k|$, and $\sum |J_i| \leq \varepsilon|V|$. Suppose that the partition is not regular, i.e., there are more than εk^2 irregular pairs (V_i, V_j). For each irregular pair (V_i, V_j) we pick two sets $V_{i,1}^{(i,j)} \subset V_i$ and $V_{j,1}^{(i,j)} \subset V_j$ such that $|V_{i,1}^{(i,j)}|, |V_{j,1}|^{(i,j)} \geq \varepsilon|V_i|$ and

$$\left|d(V_{i,1}^{(i,j)}, V_{j,1}^{(i,j)}) - d(V_i, V_j)\right| > \varepsilon.$$

Let $V_{i,2}^{(i,j)} = V_i \setminus V_{i,1}^{(i,j)}$ and $V_{j,2}^{(i,j)} = V_j \setminus V_{j,1}^{(i,j)}$. By Lemma 2 we have

$$\sum_{1\leq s,t\leq 2} f\bigl(V_{i,s}^{(i,j)}, V_{j,s}^{(i,j)}\bigr) \geq f(V_i,V_j) + \varepsilon^2\frac{|V_i^{(i,j)}||V_j^{(i,j)}|}{|V|^2} \geq f(V_i,V_j) + \varepsilon^4/4k^2. \qquad (1)$$

A given V_i might form an irregular pair with many V_j, giving rising to many distinct partitions $V_i = V_{i,1}^{(i,j)} \cup V_{i,2}^{(i,j)}$. Let $\mathcal{V}_i$ be the partition of V_i that is the common refinement of all the partitions of the form $V_i = V_{i,1}^{(i,j)} \cup V_{i,2}^{(i,j)}$. Since there are only k non-junk parts, $\mathcal{V}_i$ contains at most 2^{k-1} parts. Corollary 3 and (1) imply that

$$\sum_{\substack{A\in\mathcal{V}_i\\ B\in\mathcal{V}_j}} f(A,B) \ge f(V_i,V_j) + \varepsilon^4/4k^2.$$

Define $\mathcal{P}_{\text{new}}$ to be the partition obtained from $\mathcal{P}$ by refining each V_i by $\mathcal{V}_i$. We have

$$f(\mathcal{P}_{\text{new}}) = \sum_{\substack{\mathcal{P}_{\text{new}}\\ \text{junk}}} f(A,B) + \sum_{\substack{\mathcal{P}_{\text{new}}\\ \text{non-junk}}} f(A,B),$$

where the first sum is over distinct pairs (A,B) such that either A or B is a junk part, and the second sum is over distinct pairs (A,B) such that neither A nor B is a junk. We shall show that the non-junk sum is substantially larger than its counterpart in $f(\mathcal{P})$:

$$\begin{aligned}\sum_{\substack{(A,B)\in\mathcal{P}_{\text{new}}^2\\ \text{non-junk}}} f(A,B) &= \sum_{i,j}\sum_{\substack{A\subset\mathcal{V}_i\\ B\subset\mathcal{V}_j}} f(A,B) = \sum_{(i,j)\text{ regular}} + \sum_{(i,j)\text{ irregular}}\\ &\ge \sum_{i,j} f(V_i,V_j) + \sum_{(i,j)\text{ irregular}} \varepsilon^4/4k^2\\ &\ge \sum_{\substack{\mathcal{P}\\ \text{non-junk}}} f(A,B) + (\varepsilon k^2)\varepsilon^4/4k^2.\end{aligned}$$

Since the contribution to $f\big(\mathcal{P}_{\text{new}}\big)$ are no less than those in $f(\mathcal{P})$ (by Lemma 2) we conclude that

$$f\big(\mathcal{P}_{\text{new}}\big) \ge f(\mathcal{P}) + \varepsilon^5/4.$$

The partition $\mathcal{P}_{\text{new}}$ is an improvement over $\mathcal{P}$, save for one blemish: the non-junk parts are no longer equally large. To fix this we refine the partition once again. We partition each non-junk set in $\mathcal{P}_{\text{new}}$ into sets of size exactly $\lfloor\varepsilon|V|/k4^k\rfloor$, and one leftover set of size smaller than $\lfloor\varepsilon|V|/k4^k\rfloor$. We then declare each leftover set to be a junk set. Since each part of $\mathcal{P}$ is partitioned into at most 2^{k-1} parts, the total size of leftover parts is at most $k2^{k-1}\lfloor\varepsilon|V|/k4^k\rfloor \le \varepsilon|V|2^{-k}$. The refinement of $\mathcal{P}_{\text{new}}$ can only increase value of f. We need to check that the total size of the junk does not grow above $\varepsilon|V|$ as we repeat the procedure above. Indeed, as the number of parts k increases at each step, the total size of all junk parts does not exceed $\sum_k \varepsilon|V|2^{-k} = \varepsilon|V|$.

As the value of $f(\mathcal{P})$ increases by $\varepsilon^5/4$ at each step and $f(\mathcal{P}) \le 1$, this process terminates after at most $4/\varepsilon^5$ steps. When it terminates, we obtain an ε-regular partition. At each step the number of non-junk parts grows as $k_{\text{new}} \le g(k)$ where $g(k) = k4^k/\varepsilon$. Thus, we can take $M = g^{\circ 4/\varepsilon^5}(m)$. In other words, M is a tower of exponentials of height approximately ε^{-5}. Surprisingly, this bound is essentially tight, for Timothy Gowers exhibited a graph such that the number of parts in every ε-regular partition is a tower of exponentials of height at least $\varepsilon^{-1/16}$ (in the paper "Lower bounds of tower type for Szemerédi's uniformity lemma").