Oded (January 7, 2024): Three notion of regularity (overview by Shachar Lovett)

(This overview was provided as a main part of his presentation at the theory-lunch at Weizmann.)

For sake of simplicity, the presentation is in term of Boolean matrices rather than in terms of graphs. The starting point is Szemeredi's Regularity Lemma, which is then relaxed following Frieze and Kannan, at at last we reach the new notion. The point is that the relaxation allow for more efficient constructions, which suffice for some applications. In all cases a submtarix is a generalized one; that is, it is defined by a set of (not necessarily consecutive) rows and columns.

The strong notion of regularity [Szemeredi]. Loosely speaking, a matrix is called regular if all sufficiently large submatrices have the same density of 1's. The regularity lemma asserts that each matrix can be partitioned into "few" submatrices such that almost all these submatrices are regular.

- A matrix M is ϵ-regular if, for every submatrix R of measure at least ϵ of M, the density of 1's in R is within $\pm\epsilon$ of the density of 1's in M.

  For an n-by-n matrix M, the measure of the submatrix $R = X \times Y$ in M, denoted $\mu_M(R)$, is $|X| \cdot |Y|/n^2$. The density of 1's in a matrix A, denoted $\rho(A)$, is the fraction of 1's in it. Hence, the ϵ-regularity condition asserts that for every submatrix R (of M) such that $\mu_M(R) \geq \epsilon$ it holds that $\rho(R) = (1 \pm \epsilon) \cdot \rho(M)$.

- **Theorem:** There exists a function $T : (0, 1] \to \mathbb{N}$ such that each matrix can be partitioned into $T(\epsilon)$ submatrices such that at least $1 - \epsilon$ of these submatrices are ϵ-regular.

  Furthermore, $T(\epsilon)$ is a tower of poly$(1/\epsilon)$ exponents.

The weak notion of regularity [Frieze and Kannan]. While the strong version refers to a relative notion of approximation (i.e., the regularity of submatrices is relative to their size), the following notion uses an absolute notion of approximation (i.e., the comparison is w.r.t the size of the big matrix). Furthermore, the approximation is only in the average (over the partition of the big matrix) rather tna per each part.

- A matrix A, which has entries in $[0, 1]$, ϵ-approximates the (Boolean) matrix M if for each submatrix of measure at least ϵ in A the average value of the entries of A in this submatrix is within $\pm\epsilon$ of the density of 1's in the corresponding submatrix of M.

- **Theorem:** Each matrix M is ϵ-approximated by a matrix A that can be partitioned into at most $\exp(O(1/\epsilon^2))$ *uniform* submatrices (i.e., submatrices in which all entries holds the same value).

  Alternatively, M can be partitioned to $k = \exp(O(1/\epsilon^2))$ submatrices, $M_1, ..., M_k$, such that for every submatrix R of M such that $\mu_M(R) \geq \epsilon$ it holds that

  $$\rho(R) = (1 \pm \epsilon) \cdot \sum_{i \in [k]} \frac{\mu_M(R \cap M_i)}{\mu_M(R)} \cdot \rho(M_i). \quad (1)$$

  In contrast, strong regularity requires $\rho(R) = (1 \pm \epsilon) \cdot \rho(M_i)$ for at least $(1 - \epsilon) \cdot k$ indices i and any $R \subset M_i$ such that $\mu_M(R) \geq \epsilon \cdot \mu_R(M_i)$.

  The point is that the number of submatrices is $\exp(O(1/\epsilon^2))$ rather than $T(\epsilon)$.

Another weak notion of regularity (new). This notion is incomparable to the previous one (but is definizely weaker than Szemeredi's). On the one hand, we replace two-sided approximation by one-sided one; that is, we only require that the density of 1's in the submatrix is not too large w.r.t the density in the entire matrix. In addition, we allow to use covers rather than partitions, as long as on the average no entry is covered too many times. On the other hand, as in the strong notion, the size of the submatrices is realted to the size of the parts rather than to the size of the entire matrix. In addition, a different parameter is used for the density requirement.

- A matrix A is (δ, ϵ)-spread if for any submatrix R such that $\mu_A(R) \geq \delta$, it holds that $\rho(R) \leq \min(\delta, (1 + \epsilon) \cdot \rho(A))$.
- **Theorem:** Each matrix M can be covered by $k = \exp(\mathrm{poly}(\epsilon^{-1}/\log(1/\delta))$ submatrices, $M_1, ..., M_k$, such that

  1. For each $i \in [k]$, the matrix M_i is (δ, ϵ)-spread.
  2. $\sum_{i \in [k]} \mu_M(M_i) \leq \mathrm{poly}(\log(1/\epsilon))$.

  Alternatively, the first condition can be replaced by requiring that, for every $R \subset M_i$ such that $\mu_{M_i}(R) \geq \delta$, if $\rho(M_i) \geq \epsilon$ then $\rho(R) \leq (1 + \epsilon) \cdot \rho(M_i)$.

For the application to Boolean Matrix Multiplication, additional requirements are used and proved to be statisfied. In these applications, ϵ is set to a small consatnt, and so $k = \text{quasi-poly}(1/\delta)$.