Szemerédi's regularity lemma and graph limits

Svante Janson

Abel Prize symposium, KVA, 25 May 2012

In honour of

Endre Szemerédi.

Partly based on papers by Szemerédi, Lovász, Szegedy, Borgs, Chayes, Sós, Vesztergombi, Simonovits, Diaconis, Janson and others.

Szemerédi's regularity lemma

Let G be a graph, with vertex set V. If X and Y are subsets of V, let $e_G(X, Y)$ be the number of edges between X and Y. If U and W are disjoint subsets of V and $\varepsilon > 0$, we say that (U, W) is ε-regular if there exists $d \in [0, 1]$ such that

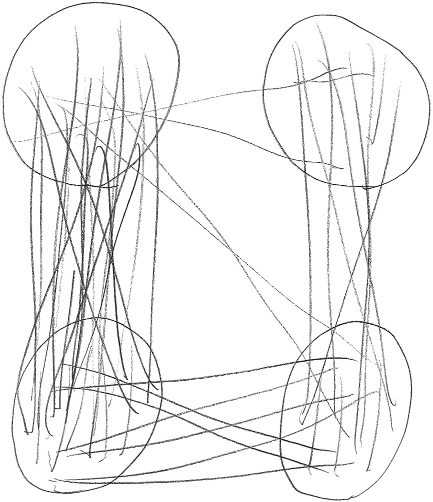
$$\left| \frac{e_G(X, Y)}{|X| \cdot |Y|} - d \right| \leq \varepsilon$$

for all $X \subseteq U$ and $Y \subseteq W$ such that $|X| \geq \varepsilon |U|$ and $|Y| \geq \varepsilon |W|$.
Finally, a partition $V_1, \ldots, V_k$ of V is a equipartition if $\big| |V_i| - |V_j| \big| \leq 1$ for all i, j.

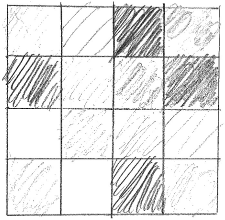
Theorem (Szemerédi's regularity lemma - one version)

Given m and $\varepsilon > 0$, there exists $M = M(m, \varepsilon)$ such that if G is any graph with at least M vertices, then for some ℓ with $m \leq \ell \leq M$, G has an equipartition into ℓ sets $V_1, \ldots, V_\ell$, such that all but $\varepsilon \ell^2$ pairs (V_i, V_j) are ε-regular.

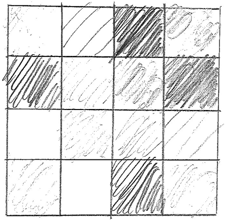
Thus, any graph can be partitioned like this:



Hence, we can approximate a large graph like this:



Hence, we can approximate a large graph like this:



0.4	0.2	1	
0.6	0.3		1
0		0.3	0.2
	0	0.6	0.4

From graphs to functions

Thus large graphs can be approximated by piecewise constant functions on $[0,1]^2$
This suggests that symmetric functions $[0,1]^2 \to [0,1]$ ("graphons") can be seen as limits of large graphs, and conversely.

Definition (in principle)

A sequence $G_1, G_2, \ldots$ of graphs with $|G_n| \to \infty$ converges to a graphon $W : [0,1]^2 \to [0,1]$ if there is a sequence of Szemerédi partitions of G_n such that the corresponding densities, regarded as functions on $[0,1]^2$ converge to W.

Subgraph densities

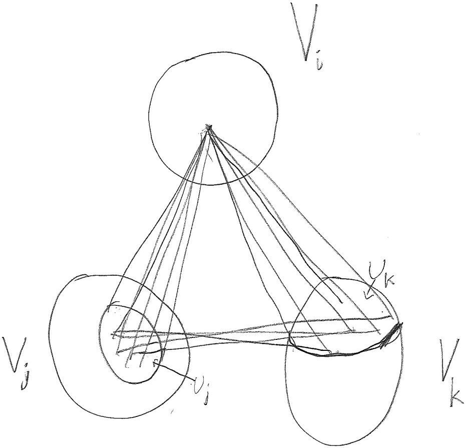
Let G be a graph with a Szemerédi partition $V_1, \dots, V_\ell$, and edge densities d_{ij}. Let $|G| = n$; thus each $|V_i| \approx n/\ell$.

Fix i and j and assume that (V_i, V_j) is ε-regular. Then most (all but $O(\varepsilon|V_i|)$) of the vertices in V_i have at least $(d_{ij} - \varepsilon)|V_j|$ neighbours in V_j.

(Otherwise there would be a large bad set $B \subset V_i$ with $e(B, V_j) < (d_{ij} - \varepsilon)|B|\,|V_j|$, violating the definition of ε-regularity.)

Similarly, most vertices in V_i have at most $(d_{ij} + \varepsilon)|V_j|$ neighbours in V_j.

Triangles



Assume that (V_i, V_j), (V_i, V_k), (V_j, V_k) are ε-regular. Then most vertices in V_i have $(d_{ij} \pm \varepsilon)|V_j|$ neighbours in V_j and $(d_{ik} \pm \varepsilon)|V_k|$ neighbours in V_k; moreover, denoting these sets of neighbours by U_j and U_k, the number of edges between them is

$$e(U_j, U_k) = (d_{jk} \pm \varepsilon)|U_j|\,|U_k| = (d_{ij}d_{ik}d_{jk} + O(\varepsilon))(n/\ell)^2.$$

This holds for most vertices in V_i, and summing over all vertices we find that the number of triangles with one vertex in each of V_i, V_j, V_k is

$$(d_{ij}d_{ik}d_{jk} + O(\varepsilon))(n/\ell)^3.$$

Summing over all i, j, k, including irregular pairs and cases with repetitions, we see that the number of triangles in G is

$$n^3 \left(\ell^{-3} \sum_{i,j,k=1}^{\ell} d_{ij} d_{ik} d_{jk} + O(\varepsilon) + O(\ell^{-1}) \right)$$

Taking the limit as $\ell \to \infty$ and $\varepsilon \to 0$, assuming that d_{ij} corresond to step functions converging to W, we find that the number of triangles is

$$n^3 \iiint W(x,y) W(x,z) W(y,z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z + o(n^3).$$

General subgraph counts

The arguments extend to counts of other subgraphs.

Definition

If F is any fixed graph, with $|F| = f$, let $t(F, G)$ denote the number of (labelled) copies of F in G, divided by $|G|^f$.

Definition

A sequence of graphs G_n with $|G_n| \to \infty$ converges to a graphon W if

$$t(F, G_n) \to t(F, W) := \int_{[0,1]^f} \prod_{ij \in E(F)} W(x_i, x_j)\, \mathrm{d}x_1 \cdots \mathrm{d}x_f.$$

Quasirandom graphs

Definition (in principle)

A quasirandom graph looks similar to a random graph where each edge appears with a fixed probability p.

Quasirandom graphs

Definition (in principle)

A quasirandom graph looks similar to a random graph where each edge appears with a fixed probability p.

Definition (with Szemerédi partitions – in principle)

A graph is quasirandom if it has a Szemerédi partition where each pair (V_i, V_j) has the same density $d_{ij} = p$.

Quasirandom graphs

Definition (in principle)

A quasirandom graph looks similar to a random graph where each edge appears with a fixed probability p.

Definition (with Szemerédi partitions – in principle)

A graph is quasirandom if it has a Szemerédi partition where each pair (V_i, V_j) has the same density $d_{ij} = p$.

Definition (with subgraph densities)

A sequence $G_1, G_2, \ldots$ of graphs with $|G_n| \to \infty$ is p-quasirandom if $t(F, G_n) \to p^{e(F)}$ for every fixed graph F.

Quasirandom graphs

Definition (in principle)

A quasirandom graph looks similar to a random graph where each edge appears with a fixed probability p.

Definition (with Szemerédi partitions – in principle)

A graph is quasirandom if it has a Szemerédi partition where each pair (V_i, V_j) has the same density $d_{ij} = p$.

Definition (with subgraph densities)

A sequence $G_1, G_2, \ldots$ of graphs with $|G_n| \to \infty$ is p-quasirandom if $t(F, G_n) \to p^{e(F)}$ for every fixed graph F.

Definition (with graph limits)

A sequence $G_1, G_2, \ldots$ of graphs with $|G_n| \to \infty$ is p-quasirandom if $G_n \to p$ (the constant graphon $W(x, y) = p$).

Random graphs from Szemerédi partitions

To construct a graph with a given Szemerédi partition $V_1, \ldots, V_\ell$ and given densities d_{ij}:

Take ℓ sets of vertices $V_1, \ldots, V_\ell$ of equal size. For each pair (V_i, V_j), add edges between them at random, with probability d_{ij} for each edge. (Independently. Toss a (biased) coin for each pair of vetices in $V_i \times V_j$ to decide whether to add an edge.)

Random graphs from graphons

Let $W : [0,1]^2 \to [0,1]$ be a graphon. Let $n \geq 1$. Gonstruct a random graph G_n with vertex set $\{1, \ldots, n\}$ as follows:

Construction 1.
Let $x_i = i/n$.
For each pair (i,j) with $i < j$, add an edge ij with probability $W(x_i, x_j)$.

Random graphs from graphons

Let $W : [0,1]^2 \to [0,1]$ be a graphon. Let $n \geq 1$. Gonstruct a random graph G_n with vertex set $\{1, \ldots, n\}$ as follows:

Construction 1.
Let $x_i = i/n$.
For each pair (i, j) with $i < j$, add an edge ij with probability $W(x_i, x_j)$.

Construction 2.
Let x_i be random (i.i.d.) in $[0, 1]$.
For each pair (i, j) with $i < j$, add an edge ij with probability $W(x_i, x_j)$.

Random graphs from graphons

Let $W : [0,1]^2 \to [0,1]$ be a graphon. Let $n \geq 1$. Gonstruct a random graph G_n with vertex set $\{1, \ldots, n\}$ as follows:

Construction 1.
Let $x_i = i/n$.
For each pair (i, j) with $i < j$, add an edge ij with probability $W(x_i, x_j)$.

Construction 2.
Let x_i be random (i.i.d.) in $[0, 1]$.
For each pair (i, j) with $i < j$, add an edge ij with probability $W(x_i, x_j)$.

Theorem
$G_n \to W$ almost surely.

Random graphs from convergent graph sequences

Suppose that $G_1, G_2, \ldots$ is a convergent sequence of graphs.

Random graphs from convergent graph sequences

Suppose that $G_1, G_2, \ldots$ is a convergent sequence of graphs.

For each m, construct a random graph H_{mn} with $|H_{mn}| = m$ by selecting m vertices from G_n at random.

Random graphs from convergent graph sequences

Suppose that $G_1, G_2, \ldots$ is a convergent sequence of graphs.

For each m, construct a random graph H_{mn} with $|H_{mn}| = m$ by selecting m vertices from G_n at random.

Let $n \to \infty$: Then $H_{mn} \xrightarrow{\mathrm{d}} H_m$ for some random graph H_m with vertex set $\{1, \ldots, m\}$.

Random graphs from convergent graph sequences

Suppose that $G_1, G_2, \dots$ is a convergent sequence of graphs.

For each m, construct a random graph H_{mn} with $|H_{mn}| = m$ by selecting m vertices from G_n at random.

Let $n \to \infty$: Then $H_{mn} \xrightarrow{\mathrm{d}} H_m$ for some random graph H_m with vertex set $\{1, \dots, m\}$.

These random graphs are consistent, so they can be regarded as subgraphs of an infinite random graph H_∞ with vertex set $\{1, 2, \dots\}$.

Random graphs from convergent graph sequences

Suppose that $G_1, G_2, \ldots$ is a convergent sequence of graphs.

For each m, construct a random graph H_{mn} with $|H_{mn}| = m$ by selecting m vertices from G_n at random.

Let $n \to \infty$: Then $H_{mn} \xrightarrow{\mathrm{d}} H_m$ for some random graph H_m with vertex set $\{1, \ldots, m\}$.

These random graphs are consistent, so they can be regarded as subgraphs of an infinite random graph H_∞ with vertex set $\{1, 2, \ldots\}$.

Theorem

Graph limits may be described by distributions of infinite random graphs H_∞.

Graphons from random graphs

Let H_∞ be the infinite random graph constructed above. Let

$$I_{ij} = \begin{cases} 1, & \text{there is an edge } ij \\ 0, & \text{there is no edge } ij. \end{cases}$$

Then I_{ij} is an *exchangeable* array of indicator random variables. (The distribution is invariant under permutations.)

General representation theorem for exchangeable arrays by Aldous and Hoover $\implies$

$$I_{ij} = f(x_i, x_j, \xi_{ij}), \qquad i < j,$$

for some function f and x_i, ξ_{ij} i.i.d. uniform in $[0, 1]$.

General representation theorem for exchangeable arrays by Aldous and Hoover $\Longrightarrow$

$$I_{ij} = f(x_i, x_j, \xi_{ij}), \qquad i < j,$$

for some function f and x_i, ξ_{ij} i.i.d. uniform in $[0, 1]$.

$\Longrightarrow$ H_m is given by the construction above from the graphon

$$W(x, y) = \mathbb{E}\, f(x, y, \xi).$$

General representation theorem for exchangeable arrays by Aldous and Hoover $\Longrightarrow$

$$I_{ij} = f(x_i, x_j, \xi_{ij}), \qquad i < j,$$

for some function f and x_i, ξ_{ij} i.i.d. uniform in $[0, 1]$.

$\Longrightarrow$ H_m is given by the construction above from the graphon

$$W(x, y) = \mathbb{E}\, f(x, y, \xi).$$

Theorem

$$G_n \to W.$$

Summary

There are strong links between the Szemerédi Regularity Lemma and graph limits.

Summary

There are strong links between the Szemerédi Regularity Lemma and graph limits.

There are some direct connections.

Summary

There are strong links between the Szemerédi Regularity Lemma and graph limits.

There are some direct connections.

In other cases they appear as alternatives:
a proof may be given either by the regularity lemma (involving combinatorial arguments, and $\varepsilon \to 0$)
or by graph limits (involving graphons and analytic arguments).

Summary

There are strong links between the Szemerédi Regularity Lemma and graph limits.

There are some direct connections.

In other cases they appear as alternatives:
a proof may be given either by the regularity lemma (involving combinatorial arguments, and $\varepsilon \to 0$)
or by graph limits (involving graphons and analytic arguments).

In any case, they both tell us that there is order in chaos.

Congratulations to the Abel Prize!