Local algorithms on random graphs and graph limits

Endre Csóka
['ɛndrɛ t͡ʃ'oːkɒ]

MTA Alfréd Rényi Institute of Mathematics
Budapest, Hungary

Dense graph limit

(by Lovász, Szegedy, Borgs, Chayes, Sós, Vesztergombi et al.)

Sampling from a graph: we choose a constant number of uniform random vertices and observe the spanned subgraph.

Accordingly, we say that a sequence of graphs (G_n) is convergent if for all $k \in \mathbb{Z}$, the probability distribution of the subgraphs spanned by k uniform random vertices $s_k(G_n)$ is convergent for $n \to \infty$.

Dense graph limit

(by Lovász, Szegedy, Borgs, Chayes, Sós, Vesztergombi et al.)

Sampling from a graph: we choose a constant number of uniform random vertices and observe the spanned subgraph.

Accordingly, we say that a sequence of graphs (G_n) is convergent if for all $k \in \mathbb{Z}$, the probability distribution of the subgraphs spanned by k uniform random vertices $s_k(G_n)$ is convergent for $n \to \infty$.

Graphon is a symmetric measurable function $W\colon [0,1] \times [0,1] \to [0,1]$. Graphs $G \in \{0,1\}^{n \times n}$ are represented by $W(x,y) = G(\lfloor nx \rfloor, \lfloor ny \rfloor)$.

The topology on the set of graphons are defined by the topology induced by $s_k(W)$, or by the metric

$$\inf_{\substack{P_1,P_2\colon [0,1]\to[0,1] \text{ bijection,} \\ P_1^{-1},P_2^{-1} \text{measure-preserving}}} \ \sup_{\substack{X,Y\subset[0,1] \\ \text{measurable}}} \int_{X\times Y} \left| W_1(x,y) - W_2(x,y) \right| \,\mathrm{d}(x,y).$$

Theorem. The two definitions are equivalent.

Theorem. The closure of the set of graphs is the set of graphons.

Theorem. The closure of the set of graphs is the set of graphons.

This language simplifies many problems, e.g.

- Szemerédi Regularity Lemma $\Leftrightarrow$ The space of graphons is compact.

Theorem. The closure of the set of graphs is the set of graphons.

This language simplifies many problems, e.g.

- Szemerédi Regularity Lemma $\Leftrightarrow$ The space of graphons is compact.
- Let $N_G(F)$ = number of labelled copies of G' in G, and $t(F, G)$ the homomorphism densities. Theorem (Chung, Graham, Wilson):

  $\forall p \in [0,1],\ \forall F \in \mathcal{G},\ \forall \varepsilon > 0,\ \exists \delta > 0$:

  $$\text{if } N_G(P_1) \geq pn^2 \text{ but } N_G(C_4) \leq (1+\delta)pn^4, \text{ then}$$

  $$N_G(F) \in (1 \pm \varepsilon)p^{V(F)}$$

Theorem. The closure of the set of graphs is the set of graphons.

This language simplifies many problems, e.g.

- Szemerédi Regularity Lemma $\Leftrightarrow$ The space of graphons is compact.
- Let $N_G(F)$ = number of labelled copies of G' in G, and $t(F, G)$ the homomorphism densities. Theorem (Chung, Graham, Wilson):

$$\forall p \in [0,1],\ \forall F \in \mathcal{G},\ \forall \varepsilon > 0,\ \exists \delta > 0:$$

$$\text{if } N_G(P_1) \geq pn^2 \text{ but } N_G(C_4) \leq (1+\delta)pn^4, \text{ then}$$

$$N_G(F) \in (1 \pm \varepsilon)p^{V(F)}$$

$$\Updownarrow$$

$$\text{If } t(C_4, W) = t(P_1, W)^4, \text{ then } W \text{ is constant.}$$

Sparse graph limits

(by Benjamini, Schramm, Elek, Hatami, Lovász, Szegedy, Gamarnik, Sudan et al.)

Sampling from a graph with degrees bounded by d: we choose a constant number of uniform random vertices v and observe $B_r(v)$, the constant-radius neighborhood of v.

Accordingly, we say that a sequence of graphs (G_n) with degree bound d is convergent if for all $r \in \mathbb{Z}$, the distribution of $B_r(G_n)$ is convergent for $n \to \infty$.

Graph limits

- Probability distributions on rooted connected (finite or) infinite trees.
  Necessary: unimodularity (some consistency condition).
  Aldous–Lyons: is it sufficient? (Elek: probably not.)
  Theorem (Cs). The question whether there is a unimodular random rooted graph supported on a set of neighborhoods is undecidable.
  The same applies for limits of finite graphs.

Sampling from a graph with degrees bounded by d: we choose a constant number of uniform random vertices v and observe $B_r(v)$, the constant-radius neighborhood of v.

Accordingly, we say that a sequence of graphs (G_n) with degree bound d is convergent if for all $r \in \mathbb{Z}$, the distribution $B_r(G_n)$ is convergent for $n \to \infty$.

Graph limits

- Probability distributions on rooted connected (finite or) infinite trees. Necessary: unimodularity (some consistency condition). Aldous–Lyons: is it sufficient? (Elek: probably not.) Theorem (Cs). The question whether there is a unimodular random rooted graph supported on a set of neighborhoods is undecidable. The same applies for limits of finite graphs.
- Graphing or measurable graph is a finite union of measure-preserving partial bijections on a measurable vertex set. It is a richer structure: Theorem. Every convergent sequence of graphs local-globally tends to a graphing.

Local-global limit and local algorithms

For a finite set of colors $C = \{1, 2, ..., c\}$, consider all possible vertex colorings of a graph G. Consider the colored r-neighborhood distributions, let us call now them structures. These define the local-global topology on graphs. For graphings, we want measurable colorings.

Lemma. For each unimodular random graph, there exists a graphing (Bernoulli-graphing) with the weakest possible structure: only those structures which can be made by local algorithms.

Questions:

- Is there a graph sequence local-globally converging to each graphing?
- Is the Bernoulli-graphing the local-global limit of a graph sequence?
- Is the sequence of random graphs asymptotically the least structured?

Theorem. (Gamarnik, Sudan) For large d, random graphs have larger independence ratio as what can be constructed by any local algorithm.

Local (graph) algorithms

Local algorithm is a function $l\colon [0,1]^{V(B_r(v))} \to C$. We assign a random variable q to each vertex (or edge, etc.) and the output at v is $f\big(q(\mathcal{B}_r(v)), [g, B_r(G)]\big)$.

- g is a global randomization
- $B_r(G)$ is the exact statistics of r-neighborhoods

Theorem. (Cs) Access to $B_r(G)$ does not help.

Theorem. (Bollobás) A random 3-regular graph has an independence ratio < 0.46. (Later improved to 0.455.)

Rephrased theorem. Let $H(\cdot)$ denote the entropy of the output at a random vertex, and $H(-)$ is the entropy on two neighboring vertex. Then

$$H(-) \geq \frac{4}{3} H(\cdot).$$

Theorem. (Cs) We can construct an independence ratio 0.445 by a local algorithm.

Entropy bounds for local algorithms

Reveal the random seeds one by one. The Shapley–Shannon information $i(x, v)$ of a seed $s(x)$ to the output $c(v)$ is the expected mutual information between them when the seed is revealed. We know that for neighboring vertices v, w,

$$
\begin{aligned}
i(x, v) &\geq 0 \\
H(c(v)) &= \sum_x i(x, v) \\
H(c(v), c(w)) &\geq \sum_x \max\left(i(x, v), i(x, w)\right).
\end{aligned}
$$

All negative results about local algorithms are followed by such inequalities (+ graph automorphisms).

If we exchange the entropy function to other functions, we can get correlation-inequalities and other bounds.