Large Deviations for Random Graphs

Honours Year Project Mid-term Progress Talk

NG Say-Yao Mark[1] A/P SUN Rongfeng[2]

[1]Student, [2]Supervisor
Department of Mathematics
National University of Singapore (NUS)

January 19, 2017

A Quick Review

Project Goals

- Understand important concepts in large deviation theory, as well as the theory of dense graph sequences.
- Understand how it is applied to study large deviations in random graph models.
- Study results related to the Erdős-Rényi Random Graph model.
- Some interesting questions:
  - How to establish LDP for $G_{n,p}$.
  - Applications of this LDP, e.g. to triangle counts.
  - Conditional distributions given some rare event.

A Quick Review

The Erdős-Rényi Random Graph

Definition

The Erdős-Rényi graph $G_{n,p}$ is a random graph defined on vertex set $[n] = \{1, 2, \ldots, n\}$, with

$$\mathbf{P}((i,j) \in E(G_{n,p})) = p$$

independently for all $1 \leq i < j \leq n$.

- We can do some elementary analysis on $G_{n,p}$.
- For example, let $T_{n,p}$ denote the number of triangles in $G_{n,p}$. What is $\mathbf{E}\, T_{n,p}$?
- What about a Large Deviation Principle (LDP) for $G_{n,p}$?
- For the last question, we first need to figure out how to (meaningfully) take a limit on a sequence of graphs...

Limits of Dense Graph Sequences

Main Discussion

- Let $\{G_n\}$ be a sequence of finite simple graphs such that the number of vertices tend to infinity, i.e. $|V(G)| \to \infty$.
- Define the *homomorphism density* between any finite simple graph H and G_n to be

$$t(H, G_n) = \frac{\hom(H, G_n)}{|V(G_n)|^{|V(H)|}}.$$

- This gives us the probability that a random vertex mapping $\varphi : V(H) \to V(G_n)$ is a homomorphism.
- Think of $t(H, G_n)$ as a 'statistic' of the graph G_n.

Limits of Dense Graph Sequences

Main Discussion

Definition

$\{G_n\}$ is said to be *convergent*, if the following limit exists for every finite simple graph H

$$\lim_{n\to\infty} t(H, G_n) = t(H).$$

Definition

A *graphon* is a symmetric, measurable function of the form $f : [0,1]^2 \to [0,1]$. Denote the space of all graphons by $\mathcal{W}$.

Limits of Dense Graph Sequences

Main Discussion

Theorem (Lovász and Szegedy)

If $\{G_n\}$ is convergent, then there exists $f \in \mathcal{W}$ such that for every finite simple graph H with $V(H) = [k]$,

$$t(H) = t(H, f) := \int_{[0,1]^k} \prod_{(i,j)\in E(H)} f(x_i, x_j)\, dx_1 \cdots dx_k.$$

- An obvious implication of this result is that we have a 'limit object' for our sequence $\{G_n\}$.
- In other words, we can say $\{G_n\}$ *converges* to $f \in \mathcal{W}$, or $G_n \to f$, instead of just '$\{G_n\}$ is *convergent*'.
- *Remark.* The 'limit object' f is not necessarily unique.

Limits of Dense Graph Sequences

Distance Between Finite Simple Graphs.

- Let A, B be a $n \times n$ real-valued matrices.
- The *cut-norm*:

$$\|A\|_\square = \sup_{S,T \subseteq [n]} \left| \sum_{(i,j) \in S \times T} A_{ij} \right|.$$

- The *cut-metric*:

$$d_\square(A, B) = \|A - B\|_\square.$$

Limits of Dense Graph Sequences

Distance Between Finite Simple Graphs.

- Let G, H be simple graphs on the *same vertex set* $[n]$. Denote their *adjacency matrices* by A_G, A_H respectively.
- The *cut-metric*, distance between G and H:

$$d_\square(G, H) = \|A_G - A_H\|_\square.$$

- Not good enough.

Limits of Dense Graph Sequences

Distance Between Finite Simple Graphs.

- Let G, H be simple graphs of the *same order* n.
- The distance between G and H:

$$\widehat{\delta}_\square(G, H) = \min_{\widehat{G}, \widehat{H}} d_\square(\widehat{G}, \widehat{H}).$$

  where $\widehat{G}, \widehat{H}$ range over all possible relabellings of G, H on a common vertex set $[n]$.
- Still not good enough, what about graphs of different order?

Limits of Dense Graph Sequences

Distance Between Finite Simple Graphs.

- Let G, H be simple graphs of the *finite order* $n_G, h_K \in \mathbb{N}$.
- For a simple graph say G, define $G^{(m)}$ to be the graph constructed by replacing each vertex in G with m distinct vertices, and connecting edges between vertices iff their parents were connected.
- The distance between G and H:

$$\delta_\square(G, H) = \lim_{k \to \infty} \widehat{\delta}_\square(G^{(kn_H)}, H^{(kn_G)}).$$

- Good.

Limits of Dense Graph Sequences

Distance Between Graphons.

- Let $f, g : [0,1]^2 \to \mathbb{R}$ be bounded, symmetric, measurable functions.
- The *cut-norm*:

$$\|f\|_\square = \sup_{S,T} \left| \int_{[0,1]^2} f(x,y)\,\mathrm{d}x\mathrm{d}y \right|.$$

- The *cut-metric*:

$$d_\square(f,g) = \|f - g\|_\square.$$

Limits of Dense Graph Sequences

Distance Between Graphons.

- Recall that $\mathcal{W}$ denotes the set of all *graphons*; symmetric, measurable functions $f : [0,1]^2 \to [0,1]$
- Let G be a simple graph of finite order n. Define

$$f^G(x,y) = \begin{cases} 1 & (\lceil nx \rceil, \lceil ny \rceil) \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

- For two simple graphs G, H of finite order n, consider

$$\begin{aligned} d_\square(f^G, f^H) &= \|f^G - f^H\|_\square \\ &= \sup_{S,T \in [0,1]} \left| \int_{S \times T} f^G - f^H \, \mathrm{d}x\mathrm{d}y \right|. \end{aligned}$$

Limits of Dense Graph Sequences

Distance Between Graphons.

- For ‘relabelling’ of vertices, let $S[0,1]$ denote the set of all measure preserving bijections $\sigma : [0,1] \to [0,1]$.
- For $f \in \mathcal{W}$ define

$$f_\sigma(x,y) = f(\sigma(x),\sigma(y)).$$

- It makes sense to define

$$\delta_\square(f,g) = \inf_{\sigma \in S[0,1]} d_\square(f,g_\sigma) = \inf_{\sigma \in S[0,1]} d_\square(f_\sigma,g).$$

- *Remark.* We can further show that

$$\delta_\square(f,g) = \inf_{\sigma_1,\sigma_2 \in S[0,1]} d_\square(f_{\sigma_1},g_{\sigma_2}).$$

Limits of Dense Graph Sequences

Distance Between Graphons.

- For two simple graphs G, H of finite order we can consider

$$\begin{aligned}\delta_\square(f^G, f^H) &= \inf_{\sigma \in S[0,1]} d_\square(f^G, f^H_\sigma) \\ &= \inf_{\sigma \in S[0,1]} \|f^G - f^H_\sigma\|_\square \\ &= \inf_{\sigma \in S[0,1]} \sup_{S,T \in [0,1]} \left| \int_{S \times T} f^G - f^H_\sigma \, \mathrm{d}x\mathrm{d}y \right|.\end{aligned}$$

- This is a good analogue to $\delta_\square(G, H)$, the 'distance' between two finite simple graphs.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

Some Notes.

- It is easy to see that $(\mathcal{W}, d_\square)$ defines a metric space.
- *Remark.* We consider $f, g \in \mathcal{W}$ to be equal if they are equal almost everywhere.
- Not true for $(\mathcal{W}, \delta_\square)$, as $\delta_\square$ is a pseudo-metric.
- Two ways of resolving this:
  (a) The pair $(S[0,1], \circ)$ defines a group. Denote $\widetilde{f}$ to be the closure of orbit $\{f_\sigma : \sigma \in S[0,1]\}$ under $d_\square$. Let $\widetilde{\mathcal{W}}$ denote the collection of all such $\widetilde{f}$.
  (b) Let $f \sim g$ if $\delta_\square(f, g) = 0$. Let $\widetilde{\mathcal{W}}$ denote the quotient space under the relation $\sim$.
- It should be clear that $\delta_\square(\widetilde{f}, \widetilde{g})$ is well-defined, and $(\widetilde{\mathcal{W}}, \delta_\square)$ is a metric space.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

What Should We Look For?

- Question: Can we establish useful properties of the metric spaces $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$? What are some 'useful properties' we should be looking for?
- Regularity.
  - In a sense that we can approximate $\widetilde{f} \in \widetilde{\mathcal{W}}$ using *stepfunctions*.
  - Let $\mathcal{P} = (S_1, \ldots, S_k)$ be a partition of $[0, 1]$.
  - A stepfunction $g : [0, 1]^2 \to \mathbb{R}$ with steps in $\mathcal{P}$ refers to a linear combination of indicator functions;

$$g(x, y) = \sum_{1 \leq i,j \leq k} \alpha_{ij} \mathbf{1}_{S_i \times S_j}(x, y).$$

- Compactness.
  - Good properties: completeness, every sequence has a convergent subsequence, etc.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

Remarks.

- Most of the preliminary analysis is focused on the space $(\mathcal{W}, d_\square)$, despite the main underlying interest in $(\widetilde{\mathcal{W}}, \delta_\square)$.
- Let $\varepsilon > 0$, $f, g \in \mathcal{W}$ and consider the mapping $f \mapsto \widetilde{f}$.
- If $d_\square(f, g) < \varepsilon$ then $\delta_\square(\widetilde{f}, \widetilde{g}) < \varepsilon$ as well.
- If $f \in \mathcal{W}$ can be approximated by a stepfunction g with error ε, then the same can be done for its counterpart (equivalence class) in $\widetilde{\mathcal{W}}$.
- Convergence in $(\mathcal{W}, d_\square)$ implies convergence in $(\widetilde{\mathcal{W}}, \delta_\square)$.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

An interesting result.

Lemma

We have the equality

$$\sup_{S,T} \left| \int_{S\times T} f(x,y)\,dx dy \right| = \sup_{\phi,\psi} \left| \int_{[0,1]^2} f(x,y)\phi(x)\psi(y)\,dx dy \right|,$$

where the supremum on the left-hand side ranges over all measurable subsets $S, T \subseteq [0,1]$ and the right-hand side ranges over all measurable functions $\phi, \psi : [0,1] \to [0,1]$.

More importantly, the optima is attained for both sides.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

Useful Concepts & Tools For Proving Compactness.

Definition

Let $\mathcal{P} = (S_1, \dots, S_k)$ be a partition of $[0,1]$. The *stepping* of a function $f : [0,1]^2 \to \mathbb{R}$, with steps in $\mathcal{P}$ is defined to be

$$f_{\mathcal{P}}(x,y) = \frac{1}{\mu(S_i)\mu(S_j)} \int_{S_i \times S_j} f(x_0, y_0)\, \mathrm{d}x_0 \mathrm{d}y_0,$$

for $x \in S_i, y \in S_j, 1 \le i, j \le k$.

Lemma

Let $f : [0,1]^2 \to [0,1]$ and $1 \le m < k$. For any m-partition $\mathcal{Q}$ of $[0,1]$ there is a k-partition $\mathcal{P}$ refining $\mathcal{Q}$ such that

$$\|f - f_P\|_\square < \frac{2}{\log(k/m)}.$$

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

Compactness.

Theorem

The space $(\widetilde{\mathcal{W}}, \delta_\square)$ is compact.

Sketch of proof (1)

- We establish that every sequence $(\widetilde{f}_n)_{n=1}^{\infty}$ in $\widetilde{\mathcal{W}}$ has a convergent subsequence.
- Let $(f_n)_{n=1}^{\infty}$ denote a representative in $\mathcal{W}$.
- We can find paritions $\mathcal{P}_{n,k}$ of $[0,1]$ such that
  - $d_\square(f_n, f_{n\,\mathcal{P}_{n,k}}) \leq \frac{1}{k}$.
  - $\mathcal{P}_{n,k+1}$ refines $\mathcal{P}_{n,k}$.
  - $|\mathcal{P}_{n,k}| = m_k$, only depends on k.
- Show that for each k, we can find a subsequence of a representative, $f_{n_s\,\mathcal{P}_{n,k}}$ converging to a step function g_k with m_k steps. In other words, $d_\square(f_{n_s\,\mathcal{P}_{n,k}}, g_k) \to 0$.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

Compactness.

Sketch of proof (2)

- We get a sequence $(g_k)_{k=1}^{\infty}$.
- Show that there is a $g \in \mathcal{W}$ such that $g_k \to g$ almost everywhere in with respect to the L_1-norm.
- In other words $\|g_k - g\|_1 \to 0$, and so $d_\square(\widetilde{g}_k, \widetilde{g}) \to 0$.
- To complete the proof, show that we can find a subsequence $\widetilde{f}_{n_s}$ of $\widetilde{f}_n$ converging to g in $(\widetilde{\mathcal{W}}, \delta_\square)$.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

Compactness.

Sketch of proof (3)

- Fix any $\varepsilon > 0$.
- There exists $k > \frac{3}{\varepsilon}$ such that $\delta_\square(\widetilde{g}_k, \widetilde{g}) \leq \|g_k - g\|_1 < \frac{\varepsilon}{3}$.
- Given k, there exists N such that $\delta_\square(\widetilde{f}_{n_s \mathcal{P}_{n,k}}, \widetilde{g}_k) < \frac{\varepsilon}{3}$ for all $n_s \geq N$.
- Note that $\delta_\square(\widetilde{f}_{n_s}, \widetilde{f}_{n_s \mathcal{P}_{n_s,k}}) \leq d_\square(f_{n_s}, f_{n_s \mathcal{P}_{n_s,k}}) \leq \frac{1}{k} < \frac{\varepsilon}{3}$ for all n_s.
- Therefore

$$
\begin{aligned}
\delta_\square(\widetilde{f}_{n_s}, \widetilde{g}) &\leq \delta_\square(\widetilde{f}_{n_s}, \widetilde{f}_{n_s \mathcal{P}_{n_s,k}}) + \delta_\square(\widetilde{f}_{n_s \mathcal{P}_{n_s,k}}, \widetilde{g}_k) + \delta_\square(\widetilde{g}_k, \widetilde{g}) \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon.
\end{aligned}
$$

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

So what if $(\widetilde{\mathcal{W}}, \delta_\square)$ is compact?

Recall the following theorem:

Theorem (Lovász and Szegedy)

If $\{G_n\}$ is convergent, then there exists $f \in \mathcal{W}$ such that for every finite simple graph H with $V(H) = [k]$,

$$t(H) = t(H, f) := \int_{[0,1]^k} \prod_{(i,j)\in E(H)} f(x_i, x_j)\, dx_1 \cdots dx_k.$$

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

Proof.

We use compactness of $(\widetilde{\mathcal{W}}, \delta_\square)$ and the following lemma.

Lemma (Counting Lemma)

Let f, g be graphons and H be a finite simple graph. Then

$$|t(H, f) - t(H, g)| \leq E(H)\delta_\square(\widetilde{f}, \widetilde{g}).$$

- Suppose $\{G_n\}$ is a convergent.
- Consider the sequence $(\widetilde{f}^{G_n})_{n=1}^{\infty}$ in $\widetilde{\mathcal{W}}$.
- By compactness there a subsequence $(\widetilde{f}^{G_{n_s}})_{s=1}^{\infty}$ converging to some $\widetilde{f} \in \widetilde{\mathcal{W}}$.
- Apply the Counting Lemma; $t(H, f^{G_n}) \to t(H, f)$ as required.

Properties of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$

A Further Result.

Theorem

$G_n \to f$ if and only if $\delta_\square(\widetilde{f}^{G_n}, \widetilde{f}) \to 0$.

Proof.

$(\Longleftarrow)$ If $\delta_\square(\widetilde{f}^{G_n}, f) \to 0$, we can apply the Counting Lemma to show that $G_n \to f$.

$(\Longrightarrow)$ Let $\mathcal{H}$ denote the set of all finite simple graphs.

- Consider the mapping $\widetilde{f} \mapsto (t(H, f))_{H \in \mathcal{H}}$.
- It is a continuous and injective mapping from a compact space $(\widetilde{\mathcal{W}}, \delta_\square)$ to a Hausdorff space $\mathbb{R}^{\mathcal{H}}$; so must be an embedding (there's a continuous inverse).
- i.e. $\delta_\square(\widetilde{f}^{G_n}, \widetilde{f}) \to 0$ as $(t(H, f^{G_n}))_{H \in \mathcal{H}} \to (t(H, f))_{H \in \mathcal{H}}$.

$\square$

Further Progress

Check List.

- Principles and techniques in large deviation theory.
- Idea behind taking limits of dense graph sequences.
- Studied properties $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$.
- Relationship with convergent sequences $\{G_n\}$.
- ? What about Erdős-Rényi random graph $G_{n,p}$?
- ? How to go about establishing a Large Deviation Principle?

Further Progress

The Next Step.

- The Erdős-Rényi random graph $G_{n,p}$ induces two probability measures.
- $\mathbf{P}_{n,p}$ on the space $\mathcal{W}$, via the mapping $G_{n,p} \longmapsto f^{G_{n,p}}$.
- $\widetilde{\mathbf{P}}_{n,p}$ on the space $\widetilde{\mathcal{W}}$, via the mapping $G_{n,p} \longmapsto \widetilde{f}^{G_{n,p}}$.
- The idea is to establish LDP based on the topology of $(\mathcal{W}, d_\square)$ and $(\widetilde{\mathcal{W}}, \delta_\square)$.
- Still need some more tools (Szemerédi's Regularity Lemma).

Further Progress

In my final talk.

- Establish LDP for the Erdős-Rényi random graph.
- Present some applications, study conditional distribution, and more (hopefully!).

Questions & Answers

Thank you very much for your attention.

Selected References

1 Frank den Hollander. *Large deviations.* Fields Institute monographs, ISSN 1069-5273; 14.

2 László Lovász and Balázs Szegedy. *Limits of dense graph sequences.* J. Combin. Theory Ser. B, 96 no. 6, 933-957.

3 László Lovász. *Large Networks and Graph Limits.* American Mathematical Soc., 2012.

4 Sourav Chatterjee. *An introduction to large deviations for random graphs.* To appear in Bull. Amer. Math. Soc.

5 Sourav Chatterjee and S.R.S Varadhan. *The Large Deviation Principle for the Erdős-Rényi Random Graph.* To appear in European J. Comb. (special issue on graph limits), 2011.