CS294-179 Network Structure and Epidemics
Fall 2020
Homework 3: Convergence of inhomogeneous random graphs

1 Concentration once more

We will need the following generalization of the Chernoff bound for i.i.d. Bernoulli random variables: Let $X_1, \dots, X_N$ independent r.v. with $X_i \sim Be(p_i)$ and let $n_i \in \{0, 1, 2\}$ be non-random.

(a) Prove that

$$\Pr\Big(\sum_i n_i(X_i - p_i) \geq 2\delta \sum_i p_i\Big) \leq e^{-\frac{\delta^2}{2\delta+2}\sum_i p_i}.$$

Hint: Bound $E[e^{tn_i(X_i-p_i)}]$ by $\exp((e^{n_i t} - 1)p_i - tn_i p_i) \leq \exp((e^{2t} - 1 - 2t)p_i)$.

(b) In a similar way, show that

$$\Pr\Big(\sum_i n_i(X_i - p_i) \leq -2\delta \sum_i p_i\Big) \leq e^{-\frac{\delta^2}{2}\sum_i p_i}.$$

2 $G(n, p)$ converges to the constant function p

Recall that the empirical graphon of a graph with $n \times n$ adjacency matrix A (or more general, any $n \times n$ matrix A) is defined as

$$W_A(x, y) = \sum_{i,j=1}^{n} A_{ij} \mathbb{I}_{x \in I_i, y \in I_j}$$

where $I_1, \dots, I_n$ is a partition of $[0, 1]$ into adjacent intervals of width $1/n$, and $\mathbb{I}_B$ is the indicator function that the event B happening, i.e., it is one if B holds, and 0 otherwise.

(a) Show that for any matrix with real valued entries, the cut-norm of W_A is in fact a maximum over a finite number of sets:

$$\|W_A\|_\square := \sup_{U,V \subset [0,1]} \left| \int_{U \times V} W_A \right| = \frac{1}{n^2} \max_{S,T \subset [n]} \left| \sum_{\substack{i \in U \\ j \in V}} A_{ij} \right|.$$

Hint: Consider a fixed set V, and the function $f_V(x) = \int_V dy W_A(x, y)$ and let $U_\pm$ be the sets where $f < 0$ and $f > 0$, respectively. Show that for all $U \subset [0, 1]$,

$$\left| \int_U f_V(x) \right| \leq \max\left\{ \int_{U_+} f_V, \left| \int_{U_-} f_V \right| \right\}$$

(b) Consider now the case where A is the adjacency matrix of $G(n,p)$, and use the above expression to express $\|W_A - P\|_\Box$ (where P is the constant function that assigns p to $(x,y) \in [0,1]^2$) as

$$\|W_A - P\|_\Box = \frac{1}{n^2} \max_{S,T \subset [n]} \left| \sum_{\substack{i \in S \\ j \in T}} (A_{ij} - p) \right|$$

Calculate the expectation of the sums for fixed S and T, and then use the Chernoff bound from Exercise 1 (with $N = \binom{n}{2}$ and $p_i = p$) to show that the probability that this sum is larger than ϵn^2 in absolute value is bounded by $2e^{-Kn^2}$ for some constant $K = K(\epsilon, p)$. Use a union bound to prove that $G(n,p)$ converges to the constant graphon P in probability.

(c) Bonus 1: if you know the needed probability theory, show that you get convergence with probability one.

(d) Bonus 2: consider now the case that $p = p_n \to 0$. Show that if it does not go to zero too fast, you get that $\|\frac{1}{p_n} W_A - 1\|_\Box \to 0$ in probability. If you do, you will have shown that up to rescaling, all random graphs have the same limit, provided p_n is larger enough.

3 Inhomogneous random graphs converge to the generating graphon

Recall that a graphon is a symmetric function from $[0,1]^2$ into $[0,1]$, i.e., a function $W : [0,1]^2 \to [0,1]$ such that $W(x,y) = W(y,x)$. We define the sequence $G_n(W)$ of inhomogeneous random graphs generated by W as follows:

1. Choose $x_1, \ldots, x_n$ i.i.d. uniformly at random from $[0,1]$, and define a matrix $P^{(n)} = P^{(n)}(W) \in [0,1]^{n \times n}$ by setting $P^{(n)}_{ij} = W(x_i, x_j)$.
2. The graph $G_n(W)$ on $[n]$ is then defined by choosing, independently for all $i < j$, an edge ij with probability $P^{(n)}_{ij}$.

In this exercise, we will prove the following theorem, which implies that $\delta_\Box\Big(G_n(W), W\Big) \to 0$ in probability.

Theorem 1 *If W is a graphon, then*

$$\mathbb{E}\left[\delta_\Box\Big(G_n(W), W\Big)\right] \to 0.$$

The theorem relies on two lemmas, which we will prove separately.

Lemma 1 *There exists a constant $D < \infty$, such that if $P \in [0,1]^{n \times n}$ is a symmetric matrix with empty diagonal, and $A \in \{0,1\}^{n \times n}$ is the random, symmetric matrix with empty diagonal obtained from P by setting $A_{ij} = A_{ji} = 1$ with probability P_{ij}, independently for all $i < j$, then*

$$\Pr(\|W_A - W_P\|_\Box \geq D/n) \leq 2e^{-n}.$$

As a consequence

$$E[\|W_A - W_P\|_\square] \leq D/n + 2e^{-n}.$$

To state the second lemma, for an $n \times n$ matrices A and a graphon W, we define

$$\hat{\delta}_1(A, W) = \min_\sigma \|W - W_{A^\sigma}\|_1,$$

where the min is taken over all permutations of $[n]$, and $A^\sigma_{ij} = A_{\sigma(i),\sigma(j)}$. Since such a permutation induces a measure preserving bijection on $[0,1]$ in the obvious way, and since the cut-norm is bounded by the L_1 norm, we clearly have that $\delta_\square(W_A, W) \leq \hat{\delta}_1(W_A, W)$.

Lemma 2 *For all symmetric $W : [0,1]^2 \to [0,1]$, define $P^{(n)} = P^{(n)}(W)$ to be the $n \times n$ random matrix with entries $P^{(n)}_{ij} = W(x_i, x_j)$, where $x_1, \ldots, x_n$ are chosen iid uniformly at random in $[0,1]$. Then*

$$\mathbb{E}\left[\hat{\delta}_1\left(P^{(n)}, W\right)\right] \to 0.$$

(a) Use the Chernoff bound from Exercise 1 to prove Lemma 1. The proof is very similar to that for $G(n,p)$, and is only more complicated by the fact that instead of $\binom{n}{2}$ i.i.d. $Be(p)$ random variables, you will have to deal with $\binom{n}{2}$ independent random variables distributed according to $Be(P_{ij})$.

(b) Prove Lemma 2 for step functions. More precisely, let Q_k be the partition of $[0,1]$ into k intervals of length $1/k$, and prove Lemma 2 for functions W which are constant on sets of the form $Y \times Y'$, $Y, Y' \in Q_k$.

- **Hint:** Reorder $x_1, \ldots, x_n$ in such a way that $x_1 < x_2 < \cdots < x_n$, and use that for $n \gg k$, the fraction of variables x_i that fall into the i^{th} interval of the partition Q_k is concentrated around $1/k$. Determine how large n has to be (as a function of k), to get enough concentration to imply that $E\left[\hat{\delta}_1(P^{(n)}(W), W)\right] \to 0$.

(c) Reduce Lemma 2 to the case where W is a step function. To this end, use the following approximation to the graphon W: the function W_{Q_k} obtained by averaging W over the blocks $Y \times Y'$, where Y and Y' are classes in Q_k.

- Prove that for any two graphons U, W,

$$\frac{1}{n^2} E[\sum_{i \neq j} |P^{(n)}_{ij}(U) - P^{(n)}_{ij}(W)| = \frac{n-1}{n}\|U - W\|_1.$$

- Use the fact that $\|W_{Q_k} - W\|_1 \to 0$ to reduce the proof of Lemma 2 to the case analyzed under (b). (For people familiar with measure theory: you can use the Lebesgue differentiation theorem to prove almost sure convergence, which implies L_1 convergence. You are not asked to prove this fact - you can just use it).

(d) Prove Theorem 1 from Lemmas 1 and 2.