TUTORIAL #3: MULTIPLICATIVE CHERNOFF BOUND AND CONVERGENCE OF W-RANDOM GRAPHS

CHRISTIAN BORGS AND HENRY COHN

For this tutorial, $W : [0,1]^2 \to \mathbb{R}$ will always be a symmetric, integrable function, i.e., a function such that $\int |W| < \infty$ and $W(x,y) = W(y,x)$. If $W \geq 0$, we call it a graphon.

For a graphon W and a sequence $\rho_n \to 0$, we define a sparse sequence of W-random graphs $G_n(\rho_n W)$ as follows: Choose $x_1, \dots, x_n$ 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 $P^{(n)}_{ij} = \min\{1, \rho_n W(x_i, x_j)\}$. The graph $G_n(\rho_n W)$ on $[n]$ is then defined by choosing, independently for all $i < j$, an edge ij with probability $P^{(n)}_{ij}$. In this tutorial, we will prove the following theorem. The analogous result with $\rho_n = 1$ and $W : [0,1]^2 \to [0,1]$ was used in Lecture 2 to prove the inverse counting lemma.

Theorem 1. If W is a graphon and ρ_n is such that $\rho_n \to 0$ and $n\rho_n \to \infty$, then

$$\mathbb{E}\left[\delta_\square\Big(\frac{1}{\rho_n} G_n(\rho_n W), W\Big)\right] \to 0.$$

The theorem relies on two lemmas, which we will prove separately. To state the second one, for an $n \times n$ matrices A and a graphon $W : [0,1]^2 \to \mathbb{R}_+$ 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 $A^\sigma_{ij} = A_{\sigma(i),\sigma(j)}$. (Recall that W^A denotes the graphon corresponding to a matrix A.) 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 (see Tutorial 1), we clearly have that $\delta_\square(W^A, W) \leq \hat{\delta}_1(W^A, W)$.

Lemma 2. Let $P \in [0,1]^{n\times n}$ be a symmetric matrix with empty diagonal, let $\tilde{\rho} = \frac{1}{n^2}\sum_{i,j} P_{ij}$, and let $A \in \{0,1\}^{n\times n}$ be the random, symmetric matrix with empty diagonal obtained from P by setting $A_{ij} = A_{ji} = 1$ with probability P_{ij}, independently for all $i < j$. If $\tilde{\rho} n \geq 6$, then

$$\|W^A - W^P\|_\square \leq \sqrt{\tilde{\rho}/n}$$

with probability at least $1 - e^{-n}$. As a consequence

$$E[\|W^A - W^P\|_\square] \leq \sqrt{\tilde{\rho}/n} + 12/n,$$

whether or not $n\tilde{\rho} \geq 6$.

Lemma 3. If $W \in L^1$, and $\rho_n \to 0$ then

$$\mathbb{E}\left[\hat{\delta}_1\Big(\frac{1}{\rho_n} P^{(n)}, W\Big)\right] \to 0.$$

(1) Prove the following multiplicative version of the Chernoff bound: Let $X = \sum_{i=1}^{N} X_i$ where $X_1, \dots X_N$ are independent random variable with values in $[-1,1]$. If

$$\sum_i E[|X_i|] \le B,$$

then

$$Pr\Big(X - E[X] \ge \alpha B\Big) \le \exp\Big(-\frac{\min\{\alpha, \alpha^2\}}{3} B\Big).$$

(2) Use (1) to prove Lemma 2. **Hint:** write the cut-norm as a maximum over sets $S, T \subset [n]$ and use that there are 4^n such pairs, together with a union bound.

(3) Prove Lemma 3 for step functions. More precisely, let P_k be the partition of $[0,1]$ into k intervals of length $1/k$, and prove Lemma 3 for functions W which are constant on the classes of $P_k \times P_k$.

- **Hint 1:** In a first step, show that if W is bounded and $\rho_n \to 0$, it is enough to show that $E\Big[\hat{\delta}_1(H_n(W), W)\Big] \to 0$, where for an arbitrary integrable function $U : [0,1]^2 \to \mathbb{R}$, $H_n(\mathrm{U})$ is the random matrix obtained from $x_1, \dots, x_n$ by setting $(H_n)_{ij} = U(x_i, x_j)$.
- **Hint 2**: Reorder $x_1, \dots, x_n$ in such a way that $x_1 < x_2 < \dots < x_n$, and use that for $n \gg k$, the fraction of variables x_i that fall into the i^{th} interval of the partition P_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 $E\Big[\hat{\delta}_1(H_n(W), W)\Big] \to 0$.

(4) Reduce Lemma 3 to the case where W is a step function. To this end, define two approximations to the graphon W: the truncated graphon $W_{\rho_n} = \min\{W, 1/\rho_n\}$, and the graphon W_{P_k} obtained by averaging W over each class of $P_k \times P_k$.

- Use the fact that $P^{(n)}(W) = \rho_n H_n(W_{\rho_n})$ to show that

$$E\Big[\hat{\delta}_1\Big(\frac{1}{\rho_n} P^{(n)}(W), W\Big)\Big] \le \|W - W_{\rho_n}\|_1 + E\Big[\hat{\delta}_1(H_n(W), W)\Big].$$

  Hint: It will we useful to calculate the expectation of $\|H_n(W - W_{\rho_n})\|_1$ and express it in terms of $\|W - W_{\rho_n}\|_1$.
- Show that for each graphon W, $\|W - W_\rho\|_1 \to 0$ as $\rho \to 0$.
- Prove that for all graphons W, $W_{P_k} \to W$ almost everywhere, and show that this implies that $\|W_{P_k} - W\|_1 \to 0$.
- Reduce the proof of Lemma 3 to the case analyzed under (3).

(5) Prove Theorem 1 from Lemmas 2 and 3.

(6) **Challenge Problem:** Prove that under the conditions of Theorem 1,

$$\delta_\square\Big(\frac{1}{\rho_n} G_n(\rho_n W), W\Big) \to 0 \quad \text{with probability 1.}$$

Hint: When proving the a.s. version of Lemma 3, use the law of large numbers for two dimensional arrays, or, in the terms of the statistics literature, the law of large numbers for U-statistics.