Notes on random sequential independent set

Spring 2017

We begin by recalling Turán's Theorem: If G is graph then

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d(v)+1}.$$

We prove this theorem by considering a random ordering of $V(G)$ and the independent set formed by including every vertex that appears before all of its neighbors in the ordering. Given this classical result, it is natural to ask the following:

> Are there situations in which we can improve on Turán by following the same proof and keeping track of whether or not the neighbors of v that appear before v in the random ordering are actually in the independent set?

In this note we use the dynamic concentration method to answer this question in the case that G is a sparse random graph.

Let $G = G_{n,p}$ with $p = \alpha/n$ and consider the following algorithm:

Set $I = \emptyset, X = [n]$
While $X \neq \emptyset$
  Choose $v \in X$ uniformly at random
  Add v to I
  Remove v and $N(v)$ from X

Let X_i be $|X|$ after i steps of this process. Our goal in this note is to track X_i over the evolution of this algorithm.

The main idea here is that we are going to track X_i by introducing a continuous function that serves as an approximation, or *scaling limit*, for the sequence $X_0, X_1, \ldots$. In particular, we hope to prove

$$X_i \approx x(i/n)n = x(t)n \tag{1}$$

where t is a continuous time variable that we relate to the process by setting $t = i/n$. Note that we scale both $|X|$ and time linearly.

We begin by deriving the function $x(t)$. Lets assume (1) holds. Then, appealing to the law of large numbers, we should have

$$\begin{aligned} nx(t+\epsilon) &\approx X_{i+\epsilon n} \\ &= X_i + \sum_{j=0}^{\epsilon n - 1} X_{i+j+1} - X_{i+j} \\ &\approx X_i + \epsilon n E\left[X_{i+1} - X_i \mid \mathcal{F}_i\right] \\ &\approx nx(t) + \epsilon n E\left[X_{i+1} - X_i \mid \mathcal{F}_i\right] \end{aligned}$$

where $\mathcal{F}_i$ is the filtration determined by the first i choices made by this algorithm. Rearranging terms we conclude that we should have

$$x'(t) = E\left[X_{i+1} - X_i \mid \mathcal{F}_i\right]$$

And therefore

$$\begin{aligned} x'(t) &= E\left[X_{i+1} - X_i \mid \mathcal{F}_i\right] \\ &= -1 - \frac{\alpha}{n}(X_i - 1) \\ &\approx -1 - \frac{\alpha}{n}(x(t)n) \\ &= -1 - \alpha x(t). \end{aligned}$$

As (1) imposes the initial condition $x(0) = 1$, we conclude that the function $x(t)$ that gives the scaling limit is

$$x(t) = -\frac{1}{\alpha} + \frac{\alpha+1}{\alpha} e^{-\alpha t}.$$

If we can prove (1) then the algorithm terminates when $x(t)$ reached zero. If we let I_{final} be the size of the independent set produced by our random sequential algorithm, we should be able to conclude

$$|I_{\text{final}}| \geq \frac{\log(\alpha+1)n}{\alpha}. \tag{2}$$

Note. *As the degree distribution in $G_{n,p}$ is Poisson, Turán's Theorem implies the following*

$$\frac{\alpha(G_{n,p})}{n} \geq \sum_{k=0}^{\infty} e^{-\alpha} \frac{\alpha^k}{k!} \cdot \frac{1}{k+1} = \frac{e^{-\alpha}}{\alpha} \sum_{k=0}^{\infty} \frac{\alpha^{k+1}}{(k+1)!} = \frac{e^{-\alpha}}{\alpha}[e^{\alpha} - 1] = \frac{1}{\alpha}[1 - e^{-\alpha}].$$

So, when α is a large constant, we conclude that (2) indeed gives a substantial improvement over Turán's Theorem.

Theorem 1. *For $i = 0, \ldots, \frac{\log(\alpha+1)}{\alpha} \cdot n - n^{3/4}$ we have*

$$X_i = x(i/n) \pm n^{2/3}. \tag{3}$$

Proof. Set $M = \frac{\log(\alpha+1)}{\alpha} \cdot n - n^{3/4}$. We define two stopping times S and T. Let S be the first step at which we choose a vertex v such that v has at least $n^{1/20}$ neighbors in X. Let T be the minimum of M and the first step where (3) does not hold.

The proofs of the upper and lower bounds in (3) are essentially identical. We prove the upper bound here.

Consider the sequence of random variables $Z_0, Z_1, \ldots$ where

$$Z_i = \begin{cases} X_i - x(t)n - e(t)n^{3/5} & \text{if } i \leq T \wedge (S-1) \\ Z_{i-1} & \text{if } i > T \wedge (S-1), \end{cases}$$

$e(t)$ is an error function that is to be determined below, and $t = i/n$. Note that we are explicitly preventing a 'huge' one step change in Z_i by 'freezing' at the previous value when we encounter a high-degree vertex. We impose the condition $e(0) = 1$ and thus

$$Z_0 = -e(0)n^{3/5} = -n^{3/5}.$$

Note further that if $e(t)$ is bounded then

$$Z_i < 0 \text{ and } i < S \quad \Rightarrow \quad \text{the upper bound in (3) holds at step } i$$

We complete the proof by showing that the $Z_0, Z_1, \ldots$ is a supermartingale and concluding, by an application of Hoeffding-Azuma, that it is unlikely for Z_i to ever be positive.

If we assume $i < T$ and we have some bound on $e''(t)$ we have

$$\begin{aligned}
E[Z_{i+1} - Z_i \mid \mathcal{F}_i] &\leq E[X_{i+1} - X_i \mid \mathcal{F}_i] - [x(i/n + 1/n) - x(i/n)]\, n \\
&\qquad - [e(i/n + 1/n) - e(i/n)]\, n + Pr\left(d(v) > n^{1/20}\right) \cdot n \\
&\leq -1 - \frac{\alpha}{n}(X_i - 1) - x'(t) - e'(t)n^{-2/5} \pm O(1/n) \\
&\leq -1 - \frac{\alpha}{n}\left(x(t)n - e(t)n^{3/5}\right) - x'(t) - e'(t)n^{-2/5} \pm O(1/n) \\
&= \alpha e(t)n^{-2/5} - e'(t)n^{-2/5} \pm O(1/n)
\end{aligned}$$

We conclude that we have a supermartingale so long as $e'(t) > \alpha e(t)$. We set $e(t) = e^{2\alpha t}$.

Note that the maximum one-step change in $Z_0, Z_1, \ldots$ is $n^{1/20}$. Therefore, it follows from Hoeffding-Azuma that we have

$$Pr\left(Z_M \geq 0\right) = Pr\left(\exists i \leq M \text{ such that} Z_i \geq 0\right) \leq \exp\left\{-\Omega\left(\frac{(n^{3/5})^2}{n \cdot \left(n^{1/20}\right)^2}\right)\right\} = \exp\left\{-\Omega\left(n^{1/10}\right)\right\}.$$

We finally note that the probability that any vertex has degree larger than $n^{1/20}$ is also exponentially small.

□