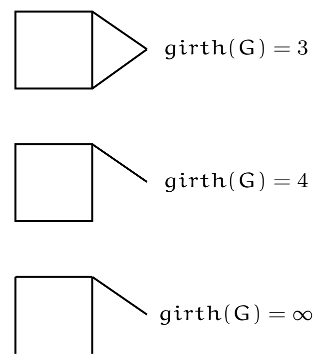
[CS1961: Lecture 10] Girth and Chromatic Number, Second-Moment Method

Instructor: Chihao Zhang;
Scribed by Haosong Liu, Shuangcheng Liu, Yuxiao Yang, Yuchen He

1 Girth and Graph Coloring

Definition 1 (Girth) *Given an undirected graph* $G = (V, E)$*, the girth of* G *is the length of the shortest cycle in* G*. Specifically, when* G *does not contain any cycle, i.e.,* G *is a forest, its girth is* ∞*.*



For example, the girth of a bipartite graph must be a even number no less than 4.

Girth reflects the connectivity of G in a sense. Intuitively, the denser G is, the smaller $\mathrm{girth}(G)$ might be.

Definition 2 (Chromatic Number) *The chromatic number of a graph* $G = (V, E)$ *is defined as* $\chi(G) = \min\{q \in \mathbb{N} \mid G$ *has a proper* q*-coloring*$\}$. [1]

[1] A proper q-coloring means we color each vertex with one of the q colors while guaranteeing no monochromatic edges.

Denser graphs tend to have larger chromatic number. The chromatic number of a complete graph $\chi(K_n)$ is n and the chromatic number of a tree is at most 2. However, this intuition is not generally correct. For example, the chromatic number of a bipartite graph is 2 while bipartite graphs can be very dense. In fact, Erdős showed that there exists graph with arbitrarily large chromatic number and large girth.

Theorem 3 (Erdős,1959) *For any* $k, \ell \in \mathbb{N}$*, there exists a graph* G *with* $\mathrm{girth}(G) \geqslant \ell$ *and* $\chi(G) \geqslant k$*.*

A tentative proof of Theorem 3.

We prove the theorem using probabilistic method by drawing graphs from $G \sim G(n, p)$ with appropriate p. [2] To prove the existence of a graph with desired property, we can turn to prove that there are respectively more than half graphs with $\mathrm{girth}(G) \geqslant \ell$ and with $\chi(G) \geqslant k$.

[2] In the Erdős-Rényi random graph model $G(n, p)$, a graph with n vertices is constructed by including each edge with probability p independently. The graph G tends to be dense when p is large and tends to be sparse for smaller p.

We compute the probability of G having each of the two properties respectively.

For the first part, we want to choose a p such that with probability larger than $\frac{1}{2}$, $\mathrm{girth}(G) \geqslant \ell$. Let X be the number of cycles with length no larger than ℓ in G. Then

$$\begin{aligned}
\mathbf{E}[X] &= \mathbf{E}\left[\sum_{i=0}^{\ell}\sum_{(v_1,v_2,\cdots,v_i)} \mathbf{1}[(v_1,v_2,\cdots,v_i) \text{ is a cycle }]\right] \\
&= \sum_{i=0}^{\ell}\sum_{(v_1,v_2,\cdots,v_i)} \mathbf{Pr}[(v_1,v_2,\cdots,v_i) \text{ is a cycle }] \\
&\overset{(1)}{=} \sum_{i=3}^{\ell}\binom{n}{i}\cdot\frac{i!}{2i}\cdot p^i \leqslant \sum_{i=3}^{\ell}\frac{n(n-1)(n-2)\cdots(n-i+1)}{2i}\cdot p^i \\
&\leqslant \sum_{i=3}^{\ell}(np)^i \leqslant (np)^{\ell+1},
\end{aligned}$$

where (1) follows from the fact that i vertices can form $i!$ cycles and each cycle is counted repeatedly for $2i$ times. By the Markov's inequality,

$$\mathbf{Pr}[X > 0] = \mathbf{Pr}[X \geqslant 1] \leqslant \mathbf{E}[X] \leqslant (np)^{\ell+1}.$$

We can choose $p = O(\frac{1}{n})$ to satisfy the requirement $\mathbf{Pr}[X > 0] < \frac{1}{2}$ and consequently we have $\mathbf{Pr}[\text{girth}(G) \geqslant \ell] > \frac{1}{2}$.

For the second part, we want to choose a p satisfying that with probability larger than $\frac{1}{2}$, $\chi(G) \geqslant k$. Recall that $\chi(G) = k$ indicates G can be divided into k independent sets. By the pigeonhole principle, there exists an independent set with size no less than $\frac{n}{k}$. Therefore,

$$\mathbf{Pr}[\chi(G) \leqslant k] \leqslant \mathbf{Pr}\left[\alpha(G) \geqslant \frac{n}{k}\right]$$

where $\alpha(G)$ is the independent number of G. Note that for any $x \in \mathbb{N}$,

$$\begin{aligned}
\mathbf{Pr}[\alpha(G) \geqslant x] &\leqslant \mathbf{Pr}\left[\exists S \in \binom{[n]}{x}, S \text{ is an independent set}\right] \\
&\leqslant \binom{n}{x}(1-p)^{\binom{x}{2}} \leqslant n^x \cdot e^{\frac{-px(x-1)}{2}} \\
&= \left(ne^{\frac{-p(x-1)}{2}}\right)^x.
\end{aligned}$$

By choosing $p \geqslant \frac{3}{x}\log n$, we have $\mathbf{Pr}[\alpha(G) \geqslant x] < \frac{1}{2}$ and thus $\mathbf{Pr}[\chi(G) \geqslant k] > \frac{1}{2}$. However, when $x = \frac{n}{k}$, we need $p = \Omega\left(\frac{\log n}{n}\right)$, which is in contradiction with the condition $p = O\left(\frac{1}{2}\right)$ we yield in the first part. □

As stated above, the simple application of the probabilistic method fails since we cannot satisfy $p = O\left(\frac{1}{2}\right)$ and $p = \Omega\left(\frac{\log n}{n}\right)$ at the same time. To fix the problem, we need the technique of alteration.

A revised proof of Theorem 3. Instead of requiring $\mathbf{Pr}[X > 0] < \frac{1}{2}$ in the first part proof, we choose $p = \frac{\log^2 n}{n}$. Then $\mathbf{E}[X] \leqslant (np)^{\ell+1} = (\log n)^{2\ell+2} = o(n)$. Therefore, by the Markov's inequality,

$$\mathbf{Pr}\left[X \geqslant \frac{n}{2}\right] \leqslant \frac{\mathbf{E}[X]}{\frac{n}{2}} < \frac{1}{2}.$$

Note that this choice of p satisfies the second condition. We can find a graph G with $\alpha(G) \leqslant \frac{n}{2k}$ and the number of cycles shorter than ℓ is less than $\frac{n}{2}$. Then we construct a G' from G by breaking the short cycles in G. We remove one vertex from each cycle shorter than ℓ in G. Then $\chi(G') \geqslant \frac{n}{2\alpha(G')} \geqslant k$ since G' contains more than $\frac{n}{2}$ vertices and $\alpha(G') \leqslant \alpha(G)$. Therefore, such G' satisfies $\mathrm{girth}(G') \geqslant \ell$ and $\chi(G') \geqslant k$. □

2 Second-Moment Method

Let $\{X_n\}_{n\in\mathbb{N}}$ be a set of random variables where each $X_n \in \mathbb{N}$. If we want to show $\mathbf{Pr}[X_n = 0] \overset{n\to\infty}{\longrightarrow} 1$, we only need to prove $\mathbf{E}[X_n] \overset{n\to\infty}{\longrightarrow} 0$ since

$$\mathbf{Pr}[X_n > 0] = \mathbf{Pr}[X_n \geqslant 1] \leqslant \mathbf{E}[X_n]$$

by the Markov's inequality. Conversely, can we yield $\mathbf{Pr}[X_n > 0] \to 1$ only from $\mathbf{E}[X_n] \to \infty$? The answer is no.[3] We need more information about how X_n is *concentrated* to its expectation. To this end, we look at its variance, or equivalently its second moment.

[3] A counterexample is

$$X_n = \begin{cases} 0, & \text{w.p. } 1 - \frac{1}{n}; \\ n^2, & \text{w.p. } \frac{1}{n}. \end{cases}$$

In this case, $\mathbf{E}[X_n] = n \to \infty$ while $\mathbf{Pr}[X_n > 0] = \frac{1}{n} \to 0$.

Theorem 4 (Chebyshev's Inequality)

$$\forall a > 0,\ \mathbf{Pr}[|X - \mathbf{E}[X]| \geqslant a] \leqslant \frac{\mathbf{Var}[X]}{a}.$$

Proof. The proof is a direct application of the Markov's inequality:

$$\mathbf{Pr}[|X - \mathbf{E}[X]| \geqslant a] = \mathbf{Pr}\left[(X - \mathbf{E}[X])^2 \geqslant a^2\right] \leqslant \frac{\mathbf{E}\left[(X - \mathbf{E}[X])^2\right]}{a^2} = \frac{\mathbf{Var}[X]}{a^2}.$$

□

Equipped with the Chebyshev's inequality, we have

$$\mathbf{Pr}[X_n = 0] \leqslant \mathbf{Pr}[|X_n - \mathbf{E}[X_n]| \geqslant \mathbf{E}[X_n]] \leqslant \frac{\mathbf{Var}[X_n]}{(\mathbf{E}[X_n])^2}.$$

Therefore, if we want to show $\mathbf{Pr}[X_n = 0] \to 0$, we only need to show that $\mathbf{E}\left[X_n^2\right] = (1 + o(1))\,(\mathbf{E}[X_n])^2$.

Then we introduce two applications of the second moment method.

2.1 Threshold Behavior

Consider the Erdős-Rényi model $G(n, p(n))$ where $p(n)\colon \mathbb{N} \to [0,1]$. A graph property $\mathcal{P}$ is said to establish *threshold behavior* if $\exists r\colon \mathbb{N} \to [0,1]$ such that

- if $p(n) \ll r(n)$, $\mathbf{Pr}_{G\sim G(n,p(n))}[G \text{ satisfies } P] \overset{n\to\infty}{\longrightarrow} 0$;
- if $p(n) \gg r(n)$, $\mathbf{Pr}_{G\sim G(n,p(n))}[G \text{ satisfies } P] \overset{n\to\infty}{\longrightarrow} 1$.

We can find a threshold function for any monotone property[4]. Theorem 5 is an example where we use the second moment method to prove the threshold behavior.

[4] We say a graph property P is monotone if a subgraph of G satisfying P implies G satisfying P.

Theorem 5 *The property "G contains a 4-clique" has a threshold function* $n^{-\frac{2}{3}}$.

Proof. Let X be the number of 4-cliques in G. If $p(n) \ll n^{-\frac{2}{3}}$, by the Markov's inequality,

$$\mathbf{Pr}_{G\sim G(n,p(n))}\left[G \text{ contains a 4-clique}\right] = \mathbf{Pr}\left[X \geqslant 1\right] \leqslant \mathbf{E}\left[X\right].$$

For $S \subseteq \binom{[n]}{4}$, let $X_S = \mathbf{1}[S \text{ is a clique}]$. Then

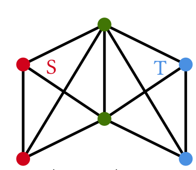
$$\mathbf{E}\left[X\right] = \mathbf{E}\left[\sum_{S\subseteq\binom{[n]}{4}} X_s\right] = \binom{n}{4}\cdot p^6 \leqslant n^4p^6 = o(1).$$

If $p(n) \gg n^{-\frac{2}{3}}$, by the Chebyshev's inequality,

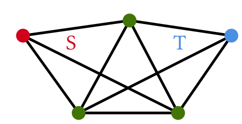
$$\mathbf{Pr}\left[X=0\right] \leqslant \mathbf{Pr}\left[|X-\mathbf{E}\left[X\right]| \geqslant \mathbf{E}\left[X\right]\right] \leqslant \frac{\mathbf{Var}\left[X\right]}{(\mathbf{E}\left[X\right])^2} = \frac{\mathbf{E}\left[X^2\right]-(\mathbf{E}\left[X\right])^2}{(\mathbf{E}\left[X\right])^2}.$$

Note that

$$\begin{aligned}
\mathbf{E}\left[X^2\right]-(\mathbf{E}\left[X\right])^2 &= \mathbf{E}\left[\left(\sum_{S\subseteq\binom{[n]}{4}} X_s\right)^2\right] - \left(\mathbf{E}\left[\sum_{S\subseteq\binom{[n]}{4}} X_s\right]\right)^2 \\
&= 2\sum_{S\neq T}\mathbf{E}\left[X_SX_T\right] + \sum_S \mathbf{E}\left[X_S^2\right] - 2\sum_{S\neq T}\mathbf{E}\left[X_S\right]\mathbf{E}\left[X_T\right] - \sum_S(\mathbf{E}\left[X_s\right])^2 \\
&= 2\sum_{|S\cap T|=2}\left(\mathbf{E}\left[X_SX_T\right]-\mathbf{E}\left[X_S\right]\mathbf{E}\left[X_T\right]\right) + 2\sum_{S\cap T=3}\left(\mathbf{E}\left[X_SX_T\right]-\mathbf{E}\left[X_S\right]\mathbf{E}\left[X_T\right]\right) \\
&\quad + \sum_S\left(\mathbf{E}\left[X_S^2\right]-(\mathbf{E}\left[X_S\right])^2\right) \\
&\leqslant 2\sum_{|S\cap T|=2}\mathbf{E}\left[X_SX_T\right] + 2\sum_{S\cap T=3}\mathbf{E}\left[X_SX_T\right] + \sum_S\mathbf{E}\left[X_S^2\right].
\end{aligned}$$



$|S\cap T| = 2$



$|S\cap T| = 3$

As the figure shows, when $|S\cap T| = 2$, $X_S = X_T = 1$ iff the 11 edges are all included. Therefore, $\mathbf{E}\left[X_SX_T\right] = \mathbf{Pr}\left[X_S = 1 \wedge X_T = 1\right] = p^{11}$. Similarly, when $|S\cap T| = 3$, $\mathbf{E}\left[X_SX_T\right] = \mathbf{Pr}\left[X_S = 1 \wedge X_T = 1\right] = p^9$. Thus,

$$\begin{aligned}
\mathbf{E}\left[X^2\right]-(\mathbf{E}\left[X\right])^2 &\leqslant 2\sum_{|S\cap T|=2}\mathbf{E}\left[X_SX_T\right] + 2\sum_{S\cap T=3}\mathbf{E}\left[X_SX_T\right] + \sum_S\mathbf{E}\left[X_S^2\right] \\
&= 2\binom{n}{2}\binom{n-2}{2}\binom{n-4}{2}p^{11} + 2\binom{n}{3}\binom{n-3}{1}\binom{n-4}{1}p^9 \\
&\quad + \binom{n}{4}p^6 \\
&\leqslant n^6p^{11} + n^5p^9 + n^4p^6 = o((\mathbf{E}\left[X\right])^2).
\end{aligned}$$

This indicates $\mathbf{Pr}\left[G \text{ contains a 4-clique}\right] \to 1$ when $p(n) \gg n^{-\frac{2}{3}}$.

□

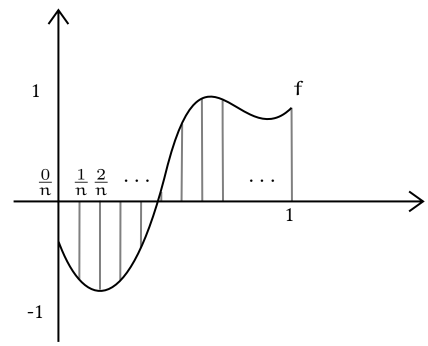
2.2 Weierstrass Approximation Theorem

Recall that we have learnt in the mathematical analysis that any continuous function on a closed interval can be approximated as closely as desired by a polynomial function. This can be proved using the second moment method.

Theorem 6 (Weierstrass Approximation Theorem) *Let* $f\colon [0,1] \to [-1,1]$ *be a continuous function. For any* $\varepsilon > 0$*, there exists a polynomial* p *such that* $\forall x \in [0,1], |p(x) - f(x)| \leqslant \varepsilon$.

Proof. Consider a random variable $Y \sim \mathrm{Bin}(n, x)$. We have $\mathbf{E}[Y] = nx$ and $\mathbf{Var}[Y] = x(1-x)n \leqslant \frac{n}{4}$. By the Chebyshev's inequality,

$$\mathbf{Pr}\left[\left|\frac{Y}{n} - x\right| \geqslant n^{-\frac{1}{3}}\right] = \mathbf{Pr}\left[|Y - nx| \geqslant n^{\frac{2}{3}}\right] \leqslant \frac{n^{-\frac{1}{3}}}{4}.$$



We use the weighted average of discrete values to get an approximation of f. Let $P_n(x) = \sum_{i=0}^{n} E_i(x) \cdot f\left(\frac{i}{n}\right)$ where $E_i(x) = \binom{n}{i} x^i (1-x)^{n-i} = \mathbf{Pr}[Y = i]$. Note that $E_i(x)$ is large when x is close to $\frac{i}{n}$ and for those i that $\frac{i}{n}$ is far from x, $\sum_i \mathbf{Pr}[Y = i] = \sum_i E_i(x)$ is small. For any $x \in [0,1]$,

$$\begin{aligned} |P_n(x) - f(x)| &\leqslant \sum_{i=1}^{n} E_i(x)\left|f\left(\frac{i}{n}\right) - f(x)\right| \\ &= \underbrace{\sum_{i\colon |i-nx| \leqslant n^{\frac{2}{3}}} E_i(x)\left|f\left(\frac{i}{n}\right) - f(x)\right|}_{A} + \underbrace{\sum_{i\colon |i-nx| > n^{\frac{2}{3}}} E_i(x)\left|f\left(\frac{i}{n}\right) - f(x)\right|}_{B}. \end{aligned}$$

Since f is continuous, there exists δ such that $\forall |x - y| < \delta, |f(x) - f(y)| < \frac{\varepsilon}{2}$. With $n^{-\frac{1}{3}} < \delta$, we have $A \leqslant \frac{\varepsilon}{2}$. Moreover, with $n^{-\frac{1}{3}} < \varepsilon$, $B \leqslant 2\sum_{i\colon |i-nx| > n^{\frac{2}{3}}} E_i(x) \leqslant \frac{n^{-\frac{1}{3}}}{2} \leqslant \frac{\varepsilon}{2}$. Therefore, choosing $n \geqslant \max\left\{\frac{1}{\varepsilon^3}, \frac{1}{\delta^3}\right\}$, we have $|P_n(x) - f(x)| \leqslant \varepsilon$ for any $x \in [0,1]$.

□