Probabilistic method in Combinatorics — matingale and chromatic number

Yuheng Zhang, 518030910438

July 4, 2020

1 Introduction

I come up with this topic because I have learned basic probabilistic method in combinatorics in CS477, which I spent lots of time, of course, so did this course. Martingale are great innovations in Probabilistic. Can we use it in combinatorics? So I looked up a lot of materials and take a short note of application of probabilistic method about chromatic number.

This note refers to notes of SJTU [Che], MIT [Lin] and Berkeley [Sin].

2 Chromatic number

In this section, I may talk about some definitons and proofs more "combinatorial".

Definition 2.1. *Let $G = (V, E)$ be a graph, and C a set of colors. A vertex colouring of G by C is a function $s : V \to C$. A coloring is proper if $s(u) \neq s(v)$ whenever $uv \in E$. The graph is said to be k-colorable if it has a proper coloring by $[k]$. The chromatic number $\chi(G)$ is defined to be the smallest k such that G is k-colorable.*

Definition 2.2. *The girth of a graph is the length of the shortest cycle in the graph, or ∞ if there is no cycle in the graph.*

Definition 2.3. *Let $G = (V, E)$ be a graph, $S \subseteq V$ is called an independent set if $xy \notin E$ for any $x, y \in S$. The independence number $\alpha(G)$ is defined as the size of the largest independent set in G.*

Fact 2.1.

$$\chi(G) \geq \frac{n}{\alpha(G)}$$

Here's a classic application of probabilistic method. The theorem is by Paul Erdős. He was one of the most prolific mathematicians and producers of mathematical conjectures of the 20th century. He devoted his waking hours to mathematics, even into his later years—indeed, his death came only hours after he solved a geometry problem at a conference in Warsaw.

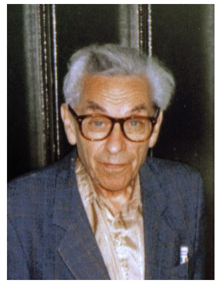

Figure 1: erdős

Theorem 2.1 (Erdős 1959)**.** *For any positive interger k and l, there is a graph G s.t. girth of G is bigger than l and $\chi(G) > k$*

Proof. For any $p \in [0,1]$, define the probability space of random graphs $\mathcal{G}(n,p)$, which means each edge is present w.p. p. For any graph G on $[n]$ with m edges, the probability mass is

$$\Pr(G) = p^m(1-p)^{\binom{n}{2}-m}$$

Define the r.v. X to be the number of cycles of length at most l, then

$$\mathrm{E}[X] \le np + (np)^2 + (np)^3 + \cdots + (np)^l$$

Define the r.v. Y to be the the number of independent sets of size $t = \frac{n}{2k}$.

$$\mathrm{E}[Y] = \binom{n}{t}(1-p)^{\binom{t}{2}} \le n^t e^{-p\binom{t}{2}} = (ne^{-p\frac{t-1}{2}})^t$$

To get a good bound, we must carefully choose a reasonable constant before t and p, we pick $p = 10k\frac{\log(n)}{n}$, then

$$\mathrm{E}[X+Y] \le (np + (np)^2 + (np)^3 + \cdots + (np)^l) + (ne^{-p\frac{t-1}{2}})^t < \frac{n}{2}$$

for big enough n.

We then remove one vertex from each short cycle or independent t-set, then we obtain a graph with $\alpha(G) < t$ and girth $> l$, then $\chi(G) > k$ by Fact 2.1. ☐

3 Martingale

Definition 3.1. *A martingale is a sequence of random variables $Z_0, Z_1, \cdots,$, such that for every n, $\mathrm{E}[Z_n] < \infty$, and*

$$\mathrm{E}[Z_{n+1}|Z_0, \cdots, Z_n] = Z_n$$

Definition 3.2. *Doob or exposure martingale: Suppose we have some random variables $X_1, \cdots, X_n$, and we have a function $f(X_1, \cdots, X_n)$. Then let*

$$Z_i = \mathrm{E}[f(X_1, \cdots, X_n)|X_1, \cdots, X_i]$$

It's easy to check that this is actually a martingale.

Proof. Use the definition of Z_i to get

$$\begin{aligned}
\mathrm{E}[Z_i|Z_1, \cdots, Z_n] &= \mathrm{E}[\mathrm{E}[f(X_1, \cdots, X_n)|Z_1, \cdots, Z_i]|Z_1, \cdots, Z_{i-1}] \\
&= \mathrm{E}[f(X_1, \cdots, X_n)|Z_1, \cdots, Z_{i-1}] \\
&= Z_{i-1}
\end{aligned}$$

□

f may also be a random variable, for example, the chromatic number of the graph, and X_i are indicator variables of the edges. In this case, $Z_0 = \mathrm{E}[f]$, the expectation when we don't know everything about edges. Z_1 revised mean after we learn about the status of an edge, and so on. This is called **edge-exposure martingale**.

Here's an example. Let's talk chromatic number of random graphs $\mathcal{G}(3, \frac{1}{2})$, which means $V = [3]$, and each edge exist w.p. $\frac{1}{2}$. There are eight possible graphs, obviously, $Z_0 = 2$. Then $Z_1 = 2.25$ or 1.75, depending on whether the first edge is present or not.

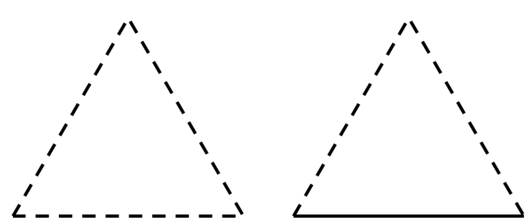

Figure 2: a simple example on triangle

Alternatively, we define **vertex-exposure martingale**: at the i-th step, expose all edges (j, i) with $j < i$.

4 Azuma's inequality

Let's see an important inequality by Kazuoki Azuma, a Japanese mathematician. Unfortunately, I didn't find his photo, and there are very little information about him.

Theorem 4.1. *Given a martingale $Z_0, \cdots, Z_n$ with bounded differences*

$$|Z_i - Z_{i-1}| \leq 1 \quad \forall i \in [n]$$

Then we have a tail bound for all lambda

$$\Pr(Z_n - Z_0 \geq \lambda\sqrt{n}) \leq e^{\frac{-\lambda^2}{2}}$$

More genera state: if

$$|Z_i - Z_{i-1}| \leq c_i \quad \forall i \in [n]$$

and $Z_0 = 0$, then for all $a > 0$

$$\Pr(Z_n - Z_0 \geq a) \leq \exp\left(\frac{-a^2}{2\sum_{i=1}^n c_i^2}\right)$$

Proof. Using convexity, we maximize e^X by having a $\pm c$ Bernoulli variable

$$\mathrm{E}[e^X] \leq \frac{e^c + e^{-c}}{2} + \frac{e^c - e^{-c}}{2c}x \leq e^{\frac{-t^2}{2}}$$

The latter inequality follows by directly comparing the Taylor series.
Now let $\lambda \geq 0$

$$\begin{aligned}
\Pr(Z_n \geq a) &= \Pr(\exp(\lambda Z_n) \geq \exp(\lambda a)) \\
&\leq \exp(-\lambda a)\mathrm{E}[\exp(\lambda Z_n)] \\
&= \exp(-\lambda a)\mathrm{E}\left[\exp\left(\lambda\sum_{i=1}^n (Z_i - Z_{i-1})\right)\right]
\end{aligned}$$

$$\mathrm{E}\left[\exp(\lambda\sum_{i=1}^n (Z_i - Z_{i-1}))\right] = \mathrm{E}\left[\mathrm{E}\left[\exp(\lambda(Z_n - Z_{n-1}))\exp\left(\lambda\sum_{i=1}^{n-1}(Z_i - Z_{i-1})\right)|\mathcal{F}_{n-1}\right]\right]$$

Since Z_i are meansureble, then

$$\begin{aligned}
&\mathrm{E}\left[\exp(\lambda(Z_n - Z_{n-1}))\exp\left(\lambda\sum_{i=1}^{n-1}(Z_i - Z_{i-1})\right)|\mathcal{F}_{n-1}\right] \\
&= \exp\left(\lambda\sum_{i=1}^{n-1}(Z_i - Z_{i-1})\right)\mathrm{E}[\exp(\lambda(Z_n - Z_{n-1}))|\mathcal{F}_{n-1}] \\
&\leq \exp\left(\lambda\sum_{i=1}^{n-1}(Z_i - Z_{i-1})\right)\left(\exp\left(\frac{\lambda^2 c_n^2}{2}\right) + \frac{\exp(\lambda c_i) - \exp(-\lambda c_i)}{2}\mathrm{E}[Z_n - Z_{n-1}|\mathcal{F}_{n-1}]\right)
\end{aligned}$$

note that $\mathrm{E}[Z_n - Z_{n-1}|\mathcal{F}_{n-1}] = 0$, then we have

$$\Pr(Z_n \geq a) \leq \exp(-\lambda a) \exp\left(\frac{\sum_{i=1}^{n} \lambda^2 d_i^2}{2}\right)$$

By setting $\lambda = \frac{a}{\sum_{i=1}^{n} c_i^2}$, we get

$$\Pr(Z_n \geq a) \leq \exp\left(-\frac{a^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

finishing the proof.

□

5 Concentration of the chromatic number

How concentrated is the chromatic number $\chi(G)$ of a random graph $G \in \mathcal{G}(n, p)$?

Theorem 5.1 (Shamir-Spencer 1987). *Let $G = \mathcal{G}(n, p)$ be a random graph. Then*

$$\Pr(|\chi(G) - \mathrm{E}[\chi(G)]| \leq \lambda\sqrt{n-1}) \leq 2e^{-\frac{\lambda^2}{2}}$$

Proof. Consider the vertex-exposure martingale. At stage i, we learn the edges from v_i to the previous vertices. Whenever a new vertex is reaveled, the worse case is assign the vertex a new color. Then $\chi(G - v_i) \leq \chi(G) \leq \chi(G - v_i) + 1$. So the expected chromatic numbers can't differ by more than 1. Apply Azuma's inequality and it is done. □

Further, we can get even better concentration if we have sufficiently small p.

Theorem 5.2. *Let $\alpha > \frac{5}{6}$ be a fixed constant. If $p < n^{-\alpha}$, then the chromatic number $\chi(\mathcal{G}(n, p))$ is concentrated among four values with high probability. That is, there exist a function $u(n, p)$, such that as $n \to \infty$*

$$\Pr(u \leq \chi(\mathcal{G}(n, p)) \leq u + 3) = 1 - o(1)$$

Proof. It suffices to show that for any fixed $\epsilon > 0$, we can find a sequence $u = u(n, p, \epsilon)$ such that as $n \to \infty$,

$$\Pr(u \leq \chi(\mathcal{G}(n, p)) \leq u + 3 > 1 - \epsilon - o(1)$$

Pick u to be the smallest positive integer such that $\Pr(\chi(\mathcal{G}(n, p)) \leq u) > \epsilon$.

Let $Y = Y(G)$ be the minimum size of a subset of the vertices $S \subset V(G)$ such that $G - S$ may be properly colored with u colors. Note that Y is 1-Lipschitz with respect to the vertex-exposure martingale. So by Azuma's inequality,

$$\Pr(Y \leq \mathrm{E}[Y] - \lambda\sqrt{n}) \leq e^{\frac{-\lambda^2}{2}}$$

$$\Pr(Y \geq \mathrm{E}[Y] + \lambda\sqrt{n}) \leq e^{\frac{-\lambda^2}{2}}$$

Then,

$$\epsilon < \Pr(Y \leq \mathrm{E}[Y] - \mathrm{E}[Y]) \leq \exp\left(-\frac{\mathrm{E}[Y]^2}{2n}\right)$$

Simplifying this, λ is a function of ϵ, to get a bound,

$$\mathrm{E}[Y] \leq \sqrt{2\log(\frac{1}{\epsilon})n} = \lambda\sqrt{n}$$

We can do an upper-tail bound to show that Y is rarely too big relative to the mean

$$\Pr(Y \geq 2\lambda\sqrt{n}) \leq \Pr(Y \geq \mathrm{E}[Y] + \lambda\sqrt{n}) \leq e^{-\frac{\lambda^2}{2}} = \epsilon$$

By the definition of λ and Y, we know that with probability at least $1-\epsilon$, we can color all but $2\lambda\sqrt{n}$ vertices.

Next, we'll show that with high probability, we can color the remaining vertices with just 3 colors. We want to find a constant C, with high probability, every subset of size at most $C\sqrt{n}$ vertices in $\mathcal{G}(n,p)$ can be properly 3-colored.

Suppose that we fault for some graph G. Choose a minimal size $T \subset V(G)$ that is not 3-colorable. Consider the induced subgraph $G[T]$. This has minimum degree 3, because if there's a vertex x with $d(x) < 3$, then $T - x$ is also not 3-colorable, which contradicts the minimality of T.

So $G[T]$ has at least $\frac{3|T|}{2}$ edges.

$$\Pr[\exists such T] \leq \sum_{t=4}^{C\sqrt{n}} \binom{n}{t} \binom{\binom{t}{2}}{\frac{3t}{2}} p^{\frac{3t}{2}}$$

Now we want to show that this is $o(1)$ as $n \to \infty$

$$\leq \sum_{t=4}^{C\sqrt{n}} \left(\frac{ne}{t}\right)^t \left(\frac{te}{3}\right)^{\frac{3t}{2}} n^{-\frac{3t\alpha}{2}} = \sum_{t=4}^{C\sqrt{n}} (O(n^{1-\frac{3\alpha}{2}+\frac{1}{t}}))^t = o(1)$$

if $\alpha > \frac{5}{6}$

The final step is that take ϵ arbitrarily small to show the result.

$\square$

6 Related wokrs

Shamir and Spencer [SS87] proved in 1987 that $\chi(G)$ is concentrated in an interval of length $\omega(n)\sqrt{n}$. For sparse random graphs, much stronger concentration results are known: Shamir and Spencer [SS87] showed that for $-p < n^{\frac{5}{6}-\epsilon}$, the chromatic number is concentrated on 5 consecutive integers, which this notes introduced.

Luczak [Sco08] sharpened this to a 2-point concentration result.

Alon and Krivelevich [AK97] extended 2-point concentration to the larger range $p < n^{-\frac{1}{2}-\epsilon}$.

Alex Scott [Sco08] improvement on the concentration result of Shamir and Spencer from $\omega(n)\sqrt{n}$ to $\frac{\omega(n)\sqrt{n}}{\log n}$ in 2008.

References

[AK97] Noga Alon and Michael Krivelevich. The concentration of the chromatic number of random graphs. *Combinatorica*, 17(3):303–313, 1997.

[Che] Xiaomin Chen. Lecture notes of cs477 of sjtu. `https://acm.sjtu.edu.cn/w/images/5/57/Notes.pdf`.

[Lin] Andrew Lin. Lecture notes of probabilistic method in combinatorics of mit. `https://ocw.mit.edu/courses/mathematics/18-218-probabilistic-method-in-combinatorics-spring-2019/lecture-notes/`.

[Sco08] Alex Scott. On the concentration of the chromatic number of random graphs. *arXiv preprint arXiv:0806.0178*, 2008.

[Sin] Alistair Sinclair. Lecture notes of cs271 of berkeley. `https://people.eecs.berkeley.edu/~sinclair/cs271/`.

[SS87] Eli Shamir and Joel Spencer. Sharp concentration of the chromatic number on random graphs g n, p. *Combinatorica*, 7(1):121–129, March 1987.