Probabilistic Methods in Graph Theory	Instructor: Paddy

Lecture 2: High Girth and High Chromatic Number

Week 1 of 1 *Mathcamp 2010*

In this talk, we will develop Markov's inequality, a remarkably useful tool, and show how its application can give us (with relatively little effort, especially when compared to constructive methods!) graphs of arbitrarily high girth and chromatic number.

Theorem 1 *(Markov's Inequality) For a random variable X and any positive constant a, $Pr(|X| > a) \leq \mathbb{E}(|X|)/a$.*

Solution. So: let $I_{(|X|\geq a)}$ be the indicator random variable defined by

$$I_{(|X|\geq a)}(\omega) = \begin{cases} 1 & |X(\omega)| \geq a, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for any ω in Ω, we trivially have that $a \cdot I_{(|X|\geq a)}(\omega) \leq |X(\omega)|$; consequently, we have

$$\begin{aligned} \mathbb{E}(a \cdot I_{(|X|\geq a)}) &= \sum_{\omega\in\Omega} a \cdot I_{(|X|\geq a)} Pr(\omega) \\ &= \sum_{\omega\in\Omega:|X(\omega)|\geq a} aPr(\omega) \\ &\leq \sum_{\omega\in\Omega:|X(\omega)|\geq a} |X(\omega)|Pr(\omega) \\ &\leq \sum_{\omega\in\Omega} |X(\omega)|Pr(\omega) \\ &= \mathbb{E}(X). \end{aligned}$$

However, on the other hand, we have

$$\begin{aligned} \mathbb{E}(a \cdot I_{(|X|\geq a)}) &= \sum_{\omega\in\Omega} a \cdot I_{(|X|\geq a)} Pr(\omega) \\ &= a \cdot \sum_{\omega\in\Omega:|X(\omega)|\geq a} Pr(\omega) \\ &= a \cdot Pr(|X| \geq a); \end{aligned}$$

combining, we have $Pr(|X| > a) \leq \mathbb{E}(|X|)/a$.

So: with this theorem under our belt, we now have the tools to resolve the following graph theory question (which otherwise is fairly hard to surmount:)

Theorem 2 *There are graphs with arbitarily high girth and chromatic number.*

Proof. So: let $G_{n,p}$ denote a random graph on n vertices, formed by doing the following:

- Start with n vertices.
- For every pair of vertices $\{x, y\}$, flip a biased coin that comes up heads with probability p and tails with probability $1-p$. If the coin is heads, add the edge $\{x, y\}$ to our graph; if it's tails, don't.

Our roadmap, then, is the following:

- For large n and well-chosen p, we will show that $G_{n,p}$ will have relatively "few" short cycles at least half of the time.
- For large n, we can also show that G will have high chromatic number at least half the time.
- Finally, by combining these two results and deleting some vertices from our graph, we'll get that graphs with both high chromatic number and no short cycles exist in our graph.

To do the first: fix a number l, and let X be the random variable defined by $X(G_{n,p}) =$ the number of cycles of length $\leq l$ in $G_{n,p}$.

We then have that

$$X(G_{n,p}) \leq \sum_{j=3}^{l} \sum_{all j-tuples x_1 \ldots x_j} N_{x_1 \ldots x_j},$$

where $N_{x_1 \ldots x_j}$ is the event that the vertices $x_1 \ldots x_j$ form a cycle.

Then, we have that

$$\begin{aligned}
\mathbb{E}(X) &\leq \sum_{j=3}^{l} \sum_{j-tuples\ x_1 \ldots x_j} Pr(N_{x_1 \ldots x_j}) \\
&= \sum_{j=3}^{l} \sum_{j-tuples\ x_1 \ldots x_j} p^j \\
&= \sum_{j=3}^{l} n^j p^j.
\end{aligned}$$

To make our sum easier, let $p = n^{\lambda-1}$, for some $\lambda \in (0, 1/l)$; then, we have that

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{j=3}^{l} n^j p^j \\
&= \sum_{j=3}^{l} n^j n^{j\lambda - j} \\
&= \sum_{j=3}^{l} n^{j\lambda} \\
&< \frac{n^{\lambda(l+1)}}{n^\lambda - 1} \\
&= \frac{n^{\lambda l}}{1 - n^{-\lambda}}
\end{aligned}$$

We claim that this is smaller than n/c, for any c and sufficiently large n. To see this, simply multiply through; this gives you that

$$\begin{aligned}
&\frac{n^{\lambda l}}{1 - n^{-\lambda}} < n/c \\
\Leftrightarrow & n^{\lambda l} < n/c - n^{1-\lambda}/c \\
\Leftrightarrow & n^{\lambda l} + n^{1-\lambda}/c < n/c \\
&,
\end{aligned}$$

which, because both λl and $1 - \lambda$ are less than 1, we know holds for large n.

So: to recap: we've shown that

$$\mathbb{E}(|X|) < n/4.$$

So: what happens if we apply Markov's inequality? Well: we get that

$$Pr(|X| \geq n/2) \leq \frac{\mathbb{E}(|X|)}{n/2} < \frac{n/4}{n/2} = 1/2;$$

in other words, that more than half of the time we have relatively "few" short cycles! So this is the first stage of our theorem.

Now: we seek to show that the chromatic number of our random graphs will be "large," on average. Doing this directly, by working with the chromatic number itself, would be rather ponderous: as week two of our class showed, the chromatic number of a graph can be a rather mysterious thing to calculate.

Rather, we will work with the **independence number** $\alpha(G)$ of our graph, the size of the independent set of vertices[1] in our graph. Why do we do this? Well, in a proper

[1] A set of vertices is called **independent** if their induced subgraph has no edges. in it

k-coloring of a graph, each of the colors necessarily defines an independent set of vertices, as there are no edges between vertices of the same color; ergo, we have that

$$\chi(G) \geq \frac{|V(G)|}{\alpha(G)},$$

for any graph G.

So: to make the chromatic number large, it suffices to make $\alpha(G)$ small! So: look at $Pr(\alpha(G) \geq m)$, for some value m. We then have the following:

$$\begin{aligned} Pr(\alpha(G) \geq m) &= Pr(\text{there is a subset of } G \text{ of size } m \text{ with no edges in it}) \\ &\leq \sum_{S \subset V, |S|=m} Pr(\text{there are no edges in } S\text{'s induced subgraph}) \\ &= \binom{n}{m} \cdot (1-p)^{\binom{m}{2}} \end{aligned}$$

,

as there are $\binom{n}{m}$-many such subsets, and in order for there to be no edges in S's subgraph we need merely for all of the coin-flips that go into creating S to come up tails: i.e. this happens $(1-p)^{\binom{m}{2}}$ of the time.

So: note the following useful inequalities:

- $\binom{n}{m} \leq n^m$. To see this, expand the binomial coefficient into the form $\frac{n \cdot \ldots (n-m+1)}{m!}$, discard the $m!$, and bound the m terms in the numerator by n^m. It's a awful bound, but one that is really easy to work with!
- $(1-p) < e^{-p}$. There are a number of proofs of this, which can be regarded as a simplified form of Bernoulli's inequality; elementary calculus methods should suffice to give this to you!

Applying both of these inequalities, we have that

$$\begin{aligned} Pr(\alpha(G) \geq m) &< n^m \cdot e^{-p \cdot \binom{m}{2}} \\ &= n^m \cdot e^{-p \cdot m \cdot (m-1)/2}. \end{aligned}$$

So: motivated by a desire to make the above simple, let $m = \left\lceil \frac{3}{p} \ln(n) \right\rceil$. This then gives us that

$$\begin{aligned} Pr(\alpha(G) \geq m) &< n^m \cdot e^{-p \cdot \left\lceil \frac{3}{p} \ln(n) \right\rceil \cdot (m-1)/2} \\ &= n^m \cdot n^{-3(m-1)/2}, \end{aligned}$$

which goes to 0 as n gets large. So, in particular, we know that for large values of n and any m, we have

$$Pr(\alpha(G) \geq m) < 1/2.$$

So: let's combine our results! In other words, we've successfully shown that for large n,

$$Pr(G \text{ has more than } (n/2)\text{-many cycles of length } \leq l, \text{ or} \alpha(G) \geq m) < 1.$$

So: for large n, there is a graph G so that neither of these things happen! Let G be such a graph. G has less than $n/2$-many cycles of length $\leq l$; so, from each such cycle, delete a vertex. Call the resulting graph G'.

Then, we have the following:

- By construction, G' has girth $\geq l$.
- Also by construction, G' has at least $n/2$ many vertices, as it started with n and we deleted $\leq n/2$.
- Deleting vertices doesn't decrease the independence number of a graph! (If you're not sure why this is, look at the definition and work it out!) Thus, we have that

$$\begin{aligned} \chi(G') &\geq \frac{|V(G')|}{\alpha(G')} \\ &\geq \frac{n/2}{\alpha(G)} \\ &\geq \frac{n/2}{3\ln(n)/p} \\ &= \frac{n/2}{3n^{1-\lambda}\ln(n)} \\ &= \frac{n^{\lambda}}{6\ln(n)}, \end{aligned}$$

  which goes to infinity as n grows large.

So, for large n, this graph has arbitrarily large girth and chromatic number!