First moment method

Definition 1 A discrete probability space is a finite or countable set $\Omega$ together with nonnegative weights on the elements that sum to 1. An event is a subset of $\Omega$. The probability $P(A)$ of an event $A$ is the sum of weights of the elements of $A$. Events $A$ and $B$ are independent if $P(A \cap B) = P(A)P(B)$.

A random variable is a function $X : \Omega \to \mathbb{R}$. The expected value of $X$ is

$$E(X) = \sum_{a \in \Omega} X(a)w_a.$$

Here $w_a$ is the weight of an element $a$ in $\Omega$.

Proposition 1 If $X = \sum X_i$, then $E(X) = \sum E(X_i)$.

Let $R(s, t)$ denote the least number $N$ such that if each edge of $K_N$ is colored in either red or blue then there is a red clique $K_s$ or a blue clique $K_t$.

Examples: $R(3, 3) = 6$, $R(3, 4) = 9$, $R(3, 5) = 14$, $R(4, 4) = 18$, $R(4, 5) = 9$.

The following theorem is often referred as the the birth of combinatorial probabilistic methods.

Theorem 1 (Erdős, 1947)

$$R(n, n) > 1 + o(1)\frac{\sqrt{2}}{e}n^{2n/2}.$$ 

Proof: Color edges of $K_N$ in two colors randomly and independently. For any set $S \subset V(K_N)$ of order $n$, the probability that $S$ forms a monochromatic clique is

$$2^{1-\binom{n}{2}}.$$ 

Let $X$ be the number of monochromatic cliques of order $n$. Then we have

$$E(X) = \binom{N}{n}2^{1-\binom{n}{2}}. \quad (1)$$

If $E(X) < 1$, then there exist some colorings with $X = 0$. In another word, $R(n, n) > N$. 

We will choose a maximum $N$ so that $\E(X) < 1$. Stirling formula gives an approximation to $n!$.

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$ 

Note that

$$\binom{N}{n}2^{1-\left(\frac{1}{2}\right)} < \frac{N^n}{n!}2^{1-\left(\frac{1}{2}\right)} \approx \frac{2}{\sqrt{2\pi n}} \left(\frac{Ne}{n2^{(n-1)/2}}\right)^n.$$ 

Choose $N = \lfloor \frac{2}{e} 2^{(n-1)/2} \rfloor$. Then, $\E(X) < 1$. □

**Theorem 2**

$$R(s, t) \leq R(s - 1, t) + R(s, t - 1).$$

**Proof:** Let $N = R(s - 1, t) + R(s, t - 1)$. For any edge-coloring of $K_N$, we pick any vertex $v$. A vertex $u$ is called a red (or blue) neighbor of $v$ if edge $uv$ is a red (or blue) edge.

Since the degree of $v$ is $N - 1$. By pigeonhole principle, $v$ has either $R(s - 1, t)$ red neighbors or $R(s, t - 1)$ blue neighbors.

For the first case, we consider the induced subgraphs on the set $R(v)$ of $v$’s red neighbors. By the definition of $R(s - 1, t)$, $R(v)$ contains either a red $s - 1$-clique or a blue $t$-clique. Any red $s - 1$-clique together with $v$ forms a red $s$-clique. The argument is similar for the second case. Thus, $R(s, t) \geq N$. □

**Corollary 1** For all $s, t \geq 2$,

$$R(s, t) \leq \binom{s + t - 2}{s - 1}.$$ 

**Proof:** We use induction on $s + t$. If $s = t = 2$, $R(2, 2) = 2 = \binom{1}{1}$.

If one of $s, t$ is 2, it is true since $R(s, 2) = \binom{s}{2}$ and $R(2, t) = \binom{t}{2}$. Now we assume $s, t \geq 3$. We have

$$R(s, t) \leq R(s - 1, t) + R(s, t - 1) \text{ by Theorem 2} \leq \binom{s - 1 + t - 2}{s - 2} + \binom{s + t - 1 - 2}{s - 1} \text{ by inductive hypothesis} \leq \binom{s + t - 2}{s - 1}.$$ 

The inductive step is finished. □

**Lemma 1 (Markov’s Inequality)** If $X$ takes only nonnegative values, then

$$\Pr(X \geq t) \leq \frac{1}{t} \E(X).$$ 

In particular, if $X$ is integer-valued, then $\E(X) \to 0$ implies $\Pr(X = 0) \to 1$. 2
Proof: We have
\[
\begin{align*}
E(X) &= \sum_i x_i \Pr(X = x_i) \\
&= \sum_{0 \leq x_i < t} x_i \Pr(X = x_i) + \sum_{x_i \geq t} x_i \Pr(X = x_i) \\
&\geq \sum_{x_i \geq t} x_i \Pr(X = x_i) \\
&\geq \sum_{x_i \geq t} t \Pr(X = x_i) \\
&= t \Pr(X \geq t).
\end{align*}
\]
Thus,
\[
\Pr(X \geq t) \leq \frac{1}{t} E(X).
\]
If \(X\) is integer-valued, we have
\[
\Pr(X = 0) = 1 - \Pr(X \geq 1) \geq 1 - E(X) \to 1.
\]
The proof of this lemma is finished. \(\square\)

Theorem 3 (Caro 1979, Wei 1981) For any simple graph \(G\), the independent number satisfies
\[
\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d_G(v) + 1}.
\]

Proof: Number vertices of \(G\) from 1 to \(n\) in an arbitrary way. We get an directed graph \(D\) by orientating each edge from the smaller vertex to the larger vertex. Let \(S\) be the set of vertices whose indegree is 0 in \(D\). Then \(S\) forms an independent set of \(G\). For each \(v, v \in S\) if \(v\) is smaller than its neighbors. The probability of \(v \in S\) is exactly \(\frac{1}{1 + d_v}\). By linearity, we have
\[
E(|S|) = \sum_v \Pr(v \in S) = \sum_v \frac{1}{1 + d_v}.
\]
In particular, there is an order of vertices so that the resulting set \(S\) is at least the expected value of \(|S|\). Hence, \(\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d_G(v) + 1}\). \(\square\)

2 Deletion method

A random graph is a collection of graphs together with a probability distribution over it. Here are two classical Erdős-Rényi models of random graphs:

A uniform random graph model \(G_{n,m}\): Given \(n\) and \(m = m(n)\), let each graph with vertex set \([n]\) and \(m\) edges occur with probability \((\binom{n}{m})^{-1}\), where \(N = \binom{n}{2}\).
A edge-independent random graph model $G(n, p)$: Given $n$ and $p = p(n)$, generate graph with vertex set $[n]$ by letting each pair be an edge with probability $p$.

**Deletion Method:** when a randomly generated object is close to having a desired property, a slight alternation may produce it.

**Theorem 4 (Spencer)**

$$R(n, n) \geq \frac{1 + o(1)}{en} 2^{-n}.$$  

**Proof:**

Color edges of $K_N$ in two colors randomly and independently. For any set $S \subset V(K_N)$ of order $n$, the probability that $S$ forms a monochromatic clique is $2^{1-\binom{n}{2}}$.

Let $X$ be the number of monochromatic cliques of order $n$. Then we have

$$E(X) = \binom{N}{n} 2^{1-\binom{n}{2}}.$$  

For any $\epsilon > 0$, if $E(X) < \epsilon N$, then there exist some colorings with $X < \epsilon N$. We can destroy all monochromatic cliques of order $n$ by delete one vertex from each monochromatic clique. We delete at most $\epsilon N$ vertices. In another word, $R(n, n) > (1 - \epsilon)N$.

We will choose a maximum $N$ so that $E(X) < \epsilon N$. Choose $N = \lceil (1 - \epsilon) \frac{n}{2} 2^{n/2} \rceil$. We can show $E(X) < \epsilon N$ is satisfied for $n$ large enough. Because $\epsilon$ is arbitrary, we conclude that

$$R(n, n) \geq (1 - o(1)) \frac{n}{\epsilon} 2^{n/2}.$$  

The proof of this lemma is finished. \qed

**Theorem 5 (Erdős, 1959)** Given $m \geq 3$ and $g \geq 3$, there exists a graph with girth at least $g$ and chromatic number at least $m$.

**Proof:** Consider a random graph $G$ on $n$ vertices. For any pair of vertices, an edge is added with probability $p$ independently. Here we choose $p = n^{t-1}$ with a positive constant $t < \frac{1}{2}$.

Let $X_i$ be the number of cycles of length $i$ and $X = \sum_{i=3}^{g-1} X_i$. We have

$$E(X) = \sum_{i=3}^{g-1} E(X_i) = \sum_{i=3}^{g-1} P(n, i) p^i / (2i)$$

4
\[
\leq \sum_{i=3}^{g-1}(np)^i/(2i) \\
= \sum_{i=3}^{g-1}n^{ti}/(2i) \\
= O(n^{tg}).
\]

Since \(tg < 1\), \(E(X) = o(n)\). For \(n\) sufficient large, we have \(E(X) < \frac{n}{4}\).

By Markov’s inequality, we have

\[
\Pr(X \geq n/2) \leq \frac{2}{n} E(X) \leq \frac{1}{2}.
\]

For the independent number of \(G\), we have

\[
\Pr(\alpha(G) \geq r) \leq \binom{n}{r} (1-p)^{\binom{r}{2}} < n^r e^{-pr(r-1)/2}.
\]

Choose \(r = \lceil \frac{3}{p} \ln n \rceil\). We have \(\Pr(X \geq n/2) \rightarrow 0\) and \(\Pr(\alpha(G) \geq r) \rightarrow 0\).

With positive probability, \(G\) has at most \(\frac{n}{2}\) cycles with length at most \(g - 1\) and with independent number at most \(r\). For each small cycle, delete one vertex from it. Let \(G'\) be the remaining graph. Then we have

\[
\begin{align*}
n(G') &\geq \frac{n}{2} \\
\alpha(G') &\leq r.
\end{align*}
\]

The graph \(G'\) has girth at least \(g\). We also have

\[
\chi(G') \geq \frac{n(G')}{\alpha(G')} \geq \frac{n^t}{3 \ln n}.
\]

Since \(\frac{n^t}{3 \ln n} \rightarrow \infty\), it is larger than any given number \(m\) for \(n\) sufficient large. \(\Box\)

### 3 Lovász local lemma

**Definition 2** Given events \(A\) and \(B\), the conditional probability of \(A\) given \(B\) is defined as

\[
\Pr(A|B) = \frac{\Pr(AB)}{\Pr(B)}.
\]

If \(A\) and \(B\) are independent then, \(\Pr(A|B) = \Pr(A)\).

\[
\Pr(A|BC) = \frac{\Pr(ABC)}{\Pr(B|C)}.
\]

\[
\Pr(A_1A_2 \cdots A_n) = \prod_{i=1}^{n} \Pr(A_i|A_{i+1} \cdots A_n).
\]

In the scenario of showing the existence of certain good event, Lovász local lemma is a very powerful tool. Here is the symmetric version.
Lemma 2 (Lovász local lemma) Let $A_1, A_2, \ldots, A_n$ be events satisfying

1. For $1 \leq i \leq n$, each event $A_i$ is mutually independent of all but at most $d$ ($d \geq 1$) other events.
2. For $1 \leq i \leq n$, $\Pr(A_i) \leq p$.
3. $4dp < 1$.

Then

$$\Pr(\bigcap_{i=1}^n \bar{A}_i) > 0.$$ 

**Proof:** We show by induction on $s$ that if $|S| \leq s$, then for any $i \not\in S$

$$\Pr(A_i \cap \bigcap_{j \in S} \bar{A}_j) \leq 2p.$$ 

For $S = \emptyset$ this is true by assumption. Renumber for convenience so that $i = n$, $S = \{1, \ldots, s\}$ and $A_n$ is mutually independent of events $\{A_x\}_{x \geq s}$. We have

$$\Pr(A_n | \bar{A}_1 \cdots \bar{A}_s) = \frac{\Pr(A_n, \bar{A}_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s)}{\Pr(A_1 \cdots A_d | \bar{A}_{d+1} \cdots \bar{A}_s)}.$$ 

The numerator can be bounded as follows

$$\Pr(A_n, \bar{A}_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s) \leq \Pr(A_n | \bar{A}_{d+1} \cdots \bar{A}_s) \leq \Pr(A_n) \leq p.$$ 

We bound the denominator

$$\Pr(\bar{A}_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s) \geq 1 - \sum_{i=1}^d \Pr(A_i | \bar{A}_{d+1} \cdots \bar{A}_s) \geq 1 - \sum_{i=1}^d 2p = 1 - 2dp \geq \frac{1}{2}.$$ 

Hence we have the quotient

$$\Pr(A_n | \bar{A}_1 \cdots \bar{A}_s) \leq 2p.$$ 

The induction is finished. Finally,

$$\Pr(\bar{A}_1 \cdots \bar{A}_n) = \prod_{i=1}^n \Pr(A_i | \bar{A}_1 \cdots \bar{A}_{i-1}) \geq \prod_{i=1}^n (1 - 2p) > 0.$$ 

The proof of this lemma is finished.
Theorem 6

\[ R(n, n) \geq \sqrt{2 + o(1)} \cdot 2^{-n}. \]

Proof:
Color edges of \( K_n \) in two colors randomly and independently. For any set \( S \subset V(K_n) \) of order \( n \), let \( A_S \) be the bad event that \( S \) forms a monochromatic clique.

\[ \Pr(A_S) = 2^{1 - \binom{n}{2}}. \]

\( A_S \) and \( A_T \) are independent if \( |S \cap T| \leq 1 \). Let

\[ d = \sum_{k=2}^{n-1} \binom{n}{k} \frac{(N-n-k)}{(n-k)!}. \]

By Lovász local lemma, \( R(n, n) \geq N \) if

\[ 4^{\binom{n}{2}} \frac{N^{n-2}}{(n-2)!} 2^{1 - \binom{n}{2}} < 1. \]

We will choose a maximum \( N \) satisfying above equation. A similar estimation show \( N = \lfloor (1 - o(1)) \sqrt{\frac{2e}{n}} 2^{n/2} \rfloor \).

The proof of this lemma is finished. \( \square \)

Definition 3 A graph \( G \) on vertices \([n]\) is called a dependency graph for events \( A_1, \ldots, A_n \) if for all \( i \) \( A_i \) is mutually independent of all \( A_j \) with \( \{i, j\} \notin G \).

Lemma 3 (Lovász local lemma) (General case). Let \( A_1, \ldots, A_n \) be events with dependency graph \( G \). Assume there exist \( x_1, \ldots, x_n \in [0, 1) \) with

\[ \Pr(A_i) < x_i \prod_{ij \in E(G)} (1 - x_j) \]

for all \( i \). Then

\[ \Pr(\cap_{i=1}^{n} A_i) < \prod_{i=1}^{n} (1 - x_i) > 0. \]

Proof: We show by induction on \( s \) that if \( |S| \leq s \), then for any \( i \notin S \)

\[ \Pr(A_i | \cap_{j \in S} \bar{A}_j) \leq x_i. \]

For \( S = \emptyset \), \( \Pr(A_i) < x_i \prod_{ij \in E(G)} (1 - x_j) < x_i. \)

Renumber for convenience so that \( i = n, S = \{1, \ldots, s\} \) and among \( x \in S, nx \in E(G) \) for \( x = 1, 2, \ldots, d \). We have

\[ \Pr(A_n | \bar{A}_1 \cdots \bar{A}_s) = \frac{\Pr(A_n, \bar{A}_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s)}{\Pr(A_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s)} \]

The numerator can be bounded as follows

\[ \Pr(A_n, \bar{A}_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s) \leq \Pr(A_n | \bar{A}_{d+1} \cdots \bar{A}_s) \leq \Pr(A_n) \]
We bound the denominator
\[ \Pr(A_1 \cdots A_d|\overline{A}_{d+1} \cdots A_s) \geq \prod_{i=1}^{d} \Pr(A_i|\overline{A}_{i+1} \cdots A_s) \]
\[ \geq \sum_{i=1}^{d} (1 - x_i). \]
Hence we have the quotient
\[ \Pr(A_n|\overline{A}_1 \cdots \overline{A}_s) \leq \frac{\Pr(A_n)}{\prod_{i=1}^{d} (1 - x_i)} < x_i. \]

The induction is finished. Finally,
\[ \Pr(\overline{A}_1 \cdots \overline{A}_n) = \prod_{i=1}^{n} \Pr(\overline{A}_i|\overline{A}_1 \cdots \overline{A}_{i-1}) \geq \prod_{i=1}^{n} (1 - x_i). \]
The proof of this lemma is finished.

**Corollary 2** In the symmetric version of Lovász local lemma, the condition \( 4dp < 1 \) can be replaced by \( (d + 1)e p < 1 \).

**Proof:** We will apply Lemma 3. By symmetry, we choose all \( x_i = x \), for some \( x \). It is enough to show that for each event \( A_i \), there is an \( x \) so that
\[ \Pr(A_i) \leq x(1 - x)^d. \]
In another word, we need a sufficient condition that \( p = x(1 - x)^d \) has a positive solution \( x < 1 \).

Note that \( f(x) = x(1 - x)^d \) reaches the maximum at \( x = \frac{1}{d+1} \). Since \( f(\frac{1}{d+1}) > \frac{1}{(d+1)e} \), \( p = f(x) \) has a solution if \( p \leq \frac{1}{(d+1)e} \). \( \square \)

### 4 Comparison of three methods

We will use Ramsey number \( R(3,n) \) to compare three methods — first moment method, deletion method, and Lovász local lemma.

First we use first moment method. Consider that a random graph \( G(N, p) \).

The bad structures are triangles and independent set of order \( n \).

The expected number of triangles is
\[ \binom{N}{3} p^3. \]
The expected number of independent set of size \( n \) is
\[ \binom{N}{n} (1 - p)^\binom{n}{2}. \]
Thus $R(3, n) > N$, if
\[
\binom{N}{3} p^3 + \binom{N}{n} (1-p)^{\binom{n}{2}} < 1.
\] (2)

Note that
\[
\binom{N}{n} (1-p)^{\binom{n}{2}} < \frac{N^n}{n!} e^{-p \binom{n}{2}} < \left( \frac{eN}{n e^{p(n-1)/2}} \right)^n.
\]
In this case, the first moment method does not give any substantially better bound than the trivial bound $n$.

**Deletion method:** If the expected number of bad structures is less than $\frac{N}{2}$, we can destroy all bad structures by deleting one vertex from each bad structure. The remaining subgraph has at least $\frac{N}{2}$ vertices and has no bad structures. In particular, $R(3, n) > \frac{N}{2}$ if
\[
\binom{N}{3} p^3 + \binom{N}{n} (1-p)^{\binom{n}{2}} < \frac{N}{2}.
\] (3)

Choose $N = \frac{n^{1.5}}{\log^{1.5} n}$ and $p = N^{-2/3} = \frac{\log n}{n}$. We have
\[
\binom{N}{3} p^3 + \binom{N}{n} (1-p)^{\binom{n}{2}} < \frac{N}{6} + \frac{N^n}{n!} e^{-p n (n-1)/2}
\]
\[
< \frac{N}{6} + e^{n (\log N - \log n + 1 - p(n-1)/2)}
\]
\[
< \frac{N}{6} + e^{n (-\log n^{1.5} + 1 + \log n/(2n))}
\]
\[
< \frac{N}{2}
\]
Thus, $R(3, n) \geq \frac{n^{1.5}}{\log^{1.5} n}$.

**Lovász local lemma:** We consider the random graph $G(N, P)$. For any 3-set $S$, let $A_S$ be "$S$ is a triangle". For any $n$-set $T$, let $B_T$ be "$T$ is an independent set." Then
\[
\Pr(A_S) = p^3 \quad \Pr(B_T) = (1-p)^{\binom{n}{2}} = e^{-p n^2/2}.
\]

Let $S, S'$ be adjacent in the dependency graph if they have a common edge; the same for $S, T$ or $T, T'$. Each $S$ is adjacent to $3(n-3) \approx 3n$ of other $S'$ and to less than $3 \binom{n-2}{n-2}$ of $T$. Each $T$ is adjacent to $\binom{n}{2} N < n^2 N/2$ of $S$ and to at most $\binom{N}{n}$ of other $T'$. Lovász local lemma takes the following form:

If there exist $x, y$ with
\[
x(1-x)^3(1-y)^{\binom{n}{n-2}} < e^{-p n^2/2} < y(1-x)^{n^2/2}(1-y)^{\binom{N}{n}},
\]

then $R(3, n) \geq \frac{n^{1.5}}{\log^{1.5} n}$.
then $R(3, n) > N$.

Choose $y = \frac{1}{(\frac{n}{N})^{1}}$ to maximize $y(1 - y)(\frac{N}{n})$. We can simplify the system as follows

\begin{align*}
p^{3} &< (1 + o(1))x(1 - x)^{3N} \quad (4) \\
e^{-pn^2/2} &< (1 - x)^{n^2N/2} \frac{1}{e^{\frac{N}{n}}}. \quad (5)
\end{align*}

Choose $x = (1 + \epsilon)p^{3}$ so that equation 4 is satisfied. Take logarithm of equation 5. We have

\[-pn^2/2 < -(1 + \epsilon)p^{3}n^2N/2 - 1 - n \ln(eN/n).

Choose $p = \frac{(2+2\epsilon)\ln n}{n}$ and $N = \frac{2n^2}{(1+\epsilon)^2\ln^2 n}$. The above inequality is satisfied.

Choose $\epsilon = \frac{1}{3}$ to maximize $N$. We have

\[R(3, n) \geq \left(\frac{27}{128} - o(1)\right) \frac{n^2}{\ln^2 n}.\]

The best lower bound is due to Kim:

\[R(3, n) \geq c_1 \frac{n^2}{\ln n}.

It matches the best known upper bound up to a constant factor.

This lecture is partly based on [1] and [2]. This note is only for your convenience.

References
