Probabilistic Method

Benny Sudakov

Princeton University

Rough outline

The basic Probabilistic method can be described as follows:

In order to prove the existence of a combinatorial structure with certain properties, we construct an appropriate probability space and show that a randomly chosen element in this space has desired properties with positive probability.

Ramsey theory

Of three ordinary adults,
two must have the same sex.

D.J. Kleitman

Ramsey Theory refers to a large body of deep results in mathematics with underlying philosophy: in large systems complete disorder is impossible!

Theorem: *(Ramsey 1930)*

$\forall\, k, l$ there exists $N(k, l)$ such that any two-coloring of the edges of complete graph on N vertices contains either/or

- **Red complete graph of size k**
- **Green complete graph of size l**

Ramsey numbers

Definition:

$R(k,l)$ **is the minimal** N **so that every red-green edge coloring of** K_N **contains**

- **Red complete graph of size** k**, or**
- **Green complete graph of size** l

Theorem: (*Erdős–Szekeresh 1935*)

$$R(k,l) \leq \binom{k+l-2}{k-1}$$

In particular

$$R(k,k) \leq \binom{2k-2}{k-1} \approx 2^{2k}$$

Proof: part I

Induction on $k+l$. By definition, $R(2,l) = l$ and $R(k,2) = k$. Now suppose that

$$R(a,b) \le \binom{a+b-2}{a-1}, \ \forall\ a+b<k+l.$$

Let

$$N = R(k-1,l) + R(k,l-1)$$

and consider a red-green coloring of the edges of the complete graph K_N.

Fix some vertex v of K_N and let A, B be the set of vertices connected to v by red, green edges respectively. Since $|A|+|B| = N-1$ we have that

$$|A| \ge R(k-1,l) \quad \text{or} \quad |B| \ge R(k,l-1).$$

Proof: part II

If $|A| \geq R(k-1,l)$, then A must contain either a green clique of size l or a red clique of size $k-1$ that together with v gives red clique of size k and we are done. The case $|B| \geq R(k,l-1)$ is similar.

By induction hypothesis, this implies

$$\begin{aligned} R(k,l) &\leq N = R(k-1,l) + R(k,l-1) \\ &\leq \binom{k+l-3}{k-2} + \binom{k+l-3}{k-1} \\ &= \binom{k+l-2}{k-1}. \end{aligned}$$

Growth rate of $R(k,k)$

Example:

$k-1$ **parts**

of

size $k-1$

Conjecture: *(P. Turán)*

$R(k,k)$ **has polynomial growth in** k**, moreover**

$$R(k,k) \le c\, k^2$$

Erdős existence argument

Theorem: *(Erdős 1947)*

$$R(k,k) \geq 2^{k/2}$$

Proof:

Color the edges of the complete graph K_N with $N = 2^{k/2}$ red and green randomly and independently with probability $1/2$. For any set C of k vertices the probability that C spans a monochromatic clique is $2 \cdot 2^{-\binom{k}{2}} = 2^{1-\binom{k}{2}}$.

Since there are $\binom{N}{k}$ possible choices for C, the probability that coloring contains a monochromatic k-clique is at most

$$\binom{N}{k} 2^{1-\binom{k}{2}} \leq \frac{N^k}{k!} \cdot \frac{2^{k/2+1}}{2^{k^2/2}} = \frac{2^{k/2+1}}{k!} \ll 1$$

Open problem

Determine the correct exponent in the bound for $R(k,k)$

Best current estimates

$$\frac{k}{2} \leq \log_2 R(k,k) \leq 2k$$

Large girth and large chromatic number

Definitions:

• **The girth** $g(G)$ **of a graph is the length of the shortest cycle in** G**.**

• **The chromatic number** $\chi(G)$ **is the minimal number of colors which needed to color the vertices of** G **so that adjacent vertices get different colors.**

Note:

It is easy to color graph with large girth "locally" using only three colors.

Question:

If girth of G **is large, can it be colored by few colors?**

Surprising result

Theorem: *(P. Erdős 1959.)*

For all k and l there exists a finite graph G with girth at least l and chromatic number at least k.

Remark:

Explicit constructions of such graphs were not found until only nine years later in 1968 by Lovász.

Bound on $\chi(G)$

Definition:

A set of pairwise nonadjacent vertices of a graph G is called independent. The independence number $\alpha(G)$ is the size of the largest independent set in G.

Lemma:

For every graph G on n vertices

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

Proof:

Consider the coloring of G into $\chi(G)$ colors. Then one of the colors classes has size at least $n/\chi(G)$ and its vertices form an independent set. Thus $\alpha(G) \geq n/\chi(G)$, as desired.

Probabilistic tools

Lemma: *(Linearity of expectation.)*

Let $X_1, X_2, \ldots, X_n$ be random variables. Then

$$\mathbb{E}\left[\sum_i X_i\right] = \sum_i \mathbb{E}[X_i].$$

(*No conditions on random variables!***)**

Lemma: *(Markov's inequality.)*

Let X be a non-negative random variable and λ a real number. Then

$$\mathbb{P}[X \geq \lambda] \leq \frac{\mathbb{E}[X]}{\lambda}.$$

Proof: part I

Fix $\theta < 1/l$. Let n be sufficiently large and G be a **random graph** $G(n,p)$ with $p = 1/n^{1-\theta}$. Let X be the **number of cycles** in G of **length** at most l.

As $\theta \cdot l < 1$, by linearity of expectation,

$$\mathbb{E}[X] \leq \sum_{i=3}^{l} n^i \cdot p^i \leq O\left(n^{\theta l}\right) = o(n).$$

By Markov's inequality

$$\mathbb{P}[X \geq n/2] \leq \frac{\mathbb{E}[X]}{n/2} = o(1).$$

Set $x = \frac{3}{p} \log n$, so that

$$\begin{aligned}\mathbb{P}[\alpha(G) \geq x] \ &\leq \ \binom{n}{x} (1-p)^{\binom{x}{2}} \\ &< \ \left(ne^{-px/2}\right)^x = o(1)\end{aligned}$$

Proof: part II

For large n both of these events have probability less than $1/2$. Thus **there is** a specific **graph** G with less than $n/2$ short cycles, i.e., cycles of length at most l, and with

$$\alpha(G) < x \leq 3n^{1-\theta} \log n.$$

Remove a vertex from each short cycle of G. This gives G' with at least $n/2$ **vertices**, **girth greater than** l and $\alpha(G') \leq \alpha(G)$. Therefore

$$\chi(G') \geq \frac{|G'|}{\alpha(G')} \geq \frac{n/2}{3n^{1-\theta} \log n} = \frac{n^{\theta}}{6 \log n} \gg k.$$

Set-pair estimate

Theorem: (*Bollobás 1965.*)

Let $A_1, \ldots, A_m$ and $B_1, \ldots, B_m$ be two families of sets such that $A_i \cap B_j = \emptyset$ only if $i = j$. Then

$$\sum_{i=1}^{m} \binom{|A_i| + |B_i|}{|A_i|}^{-1} \leq 1.$$

In particular if $|A_i| = a$ and $|B_i| = b$, then $m \leq \binom{a+b}{a}$.

Example:

Let X be a set of size $a + b$ and consider pairs $(A_i, B_i = X - A_i)$ for all $A_i \subset X$ of size a. There are $\binom{a+b}{a}$ such pairs, so the above theorem is tight.

Proof: part I

Let $|A_i| = a_i, |B_i| = b_i$ and let

$$X = \bigcup_i (A_i \cup B_i).$$

Consider a random order π of X and let X_i be the event that in this order all the elements of A_i precede all those of B_i.

To compute probability of X_i note that there are $(a_i + b_i)!$ possible orders of element in $A_i \cup B_i$ and the number of such orders in which all the elements of A_i precede all those of B_i is exactly $a_i!b_i!$. Therefore

$$\mathbb{P}[X_i] = \frac{a_i!b_i!}{(a_i + b_i)!} = \binom{a_i + b_i}{a_i}^{-1}.$$

Proof: part II

We claim that events X_i are pairwise disjoint. Indeed suppose that there is an order of X in which all the elements of A_i precede those of B_i and all the elements of A_j precede all those of B_j. W.l.o.g. assume that the last element of A_i appear before the last element of A_j. Then all the elements of A_i precede all those of B_j, contradicting the fact that $A_i \cap B_j \neq \emptyset$.

Therefore events X_i are pairwise disjoint and so we get

$$1 \geq \sum_{i=1}^{m} \mathbb{P}[X_i] = \sum_{i=1}^{m} \binom{a_i + b_i}{a_i}^{-1}.$$

Sperner's lemma

Theorem: *(Sperner 1928.)*

Let $A_1, \ldots, A_m$ be a family of subsets of n element set X which is an antichain, i.e., $A_i \not\subseteq A_j$ for all $i \neq j$. Then

$$m \leq \binom{n}{\lfloor n/2 \rfloor}.$$

Proof:

Let $B_i = X - A_i$ and let $|A_i| = a_i$. Then $|B_i| = b_i = n - a_i$, $A_i \cap B_i$ is empty but $A_j \cap B_i \neq \emptyset$ for all $i \neq j$. Therefore by Bollobás' theorem

$$1 \geq \sum_{i=1}^{m} \binom{a_i + b_i}{a_i}^{-1} = \sum_{i=1}^{m} \binom{n}{a_i}^{-1} \geq \frac{m}{\binom{n}{\lfloor n/2 \rfloor}}.$$

Littlewood-Offord problem

Theorem: *(Erdős 1945.)*

Let $x_1, x_2, \ldots, x_n$ be real numbers such that all $|x_i| \geq 1$. For every sequence $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_i \in \{-1, +1\}$ let

$$x_\alpha = \sum_{i=1}^{n} \alpha_i x_i.$$

Then every open interval I in the real line of length 2 contains at most $\binom{n}{\lfloor n/2 \rfloor}$ of the numbers x_α.

Remark:

Kleitman (1970) proved that this is still true if x_i are vectors in arbitrary normed space.

Proof

Replacing $x_i < 0$ by $-x_i$ we can assume that all $x_i \geq 1$. For every $\alpha \in \{-1, 1\}^n$ let A_α be the subset of $\{1, \ldots, n\}$ containing all $1 \leq i \leq n$ with $\alpha_i = -1$. Note that if $A_\alpha \subset A_\beta$ then $\alpha_i - \beta_i$ is either 0 or 2. Hence

$$x_\alpha - x_\beta = \sum_i (\alpha_i - \beta_i) x_i = 2 \sum_{i \in A_\beta - A_\alpha} x_i \geq 2.$$

This implies that $\{A_\alpha \mid x_\alpha \in I\}$ form an antichain and by Sperner's lemma their number is bounded by $\binom{n}{\lfloor n/2 \rfloor}$.

Explicit constructions

Theorem: (*Erdős 1947*)

There is a 2-**edge-coloring of complete graph** $K_N, N = 2^{k/2}$ **with no monochromatic clique of size** k**.**

Problem: (*Erdős* \$100)

Find an "explicit" such coloring.

Explicit $\stackrel{\text{def}}{=}$ **constructible in polynomial time**

Theorem: (*Frankl and Wilson 1981*)

There is an explicit 2-**edge-coloring of complete graph** $K_N, N = k^{c\frac{\log k}{\log\log k}}$ **with no monochromatic clique of size** k**.**

Bipartite Ramsey

Problem:

Find "large" $0,1$ matrix A with no $k \times k$ homogeneous submatrices.

Submatrix $\stackrel{\text{def}}{=}$ intersection of k rows and columns

Homogeneous $\stackrel{\text{def}}{=}$ containing all 0 or all 1

Randomly:

There is $N \times N$ matrix A with $k = 2 \log_2 N$.

Explicitly:

There is $N \times N$ matrix A with $k = N^{1/2}$.

E.g., take $[N] = \{0,1\}^n$ and define

$$a_{x,y} = x \cdot y (\textbf{mod } 2)$$

Breaking $1/2$ barrier

Theorem: (*Barak, Kindler, Shaltiel, S., Wigderson*)
For every constant $\delta > 0$ there exists a polynomial time computable $N \times N$ matrix A with $0, 1$ entries such that none of its $N^\delta \times N^\delta$ submatrices is homogeneous.

Moreover, every $N^\delta \times N^\delta$ submatrix of A has constant proportion of 0 and of 1.