Part II. Probabilistic Methods in Discrete Mathematics

The probabilistic method is a powerful tool in tackling many problems in discrete mathematics. Roughly speaking, the method works as follows: In order to prove that a structure with certain desired property exists, one defines an appropriate probability space of structures and then shows that the desired properties hold in this space with positive probability.

1° The Basic Method

The *Ramsey number* $R(k, \ell)$ is the smallest integer n such that in any 2-coloring of the edges of K_n (the complete graph on n vertices) by red and blue, there either is a *red* K_k (i.e. a complete subgraph on k vertices, all of whose edges are colored red), or there is a *blue* K_ℓ. For instance, $R(3, 3) = 6$.

not the chromatic problem

random 2-coloring

Ramsey showed that $R(k, \ell)$ is finite for any two integers k and ℓ. Let us derive a lower bound for the diagonal Ramsey numbers $R(k, k)$.

Theorem 21. *If* $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, *then* $R(k, k) > n$. *Thus* $R(k, k) > 2^{k/2}$ *for all* $k \geq 3$.

Proof. Consider a random 2-coloring of the edges of K_n obtained by coloring each edge independently either red or blue, where each color is equally likely. For any fixed set R of k vertices, let A_R be the event that the induced subgraph of K_n on R is *monochromatic* (i.e. that either all its edges are red or they are all blue). Clearly, $Pr(A_R) = 2\cdot(\frac{1}{2})^{\binom{k}{2}}$. Since there are $\binom{n}{k}$ possible choices for R, the probability that at least one of the events A_R occurs is at most $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$. Thus, with positive probability, no event A_R occurs and so there is a 2-coloring of K_n without monochromatic K_k, that is, $R(k, k) > n$. Note that if $k \geq 3$ and we take $n = \lfloor 2^{k/2} \rfloor$, then

why $\binom{k}{2}$, because in a complete graph with k vertices, there are $\binom{k}{2}$ edges in total

so $\mathbb{P}(A_R) = 2\cdot(\frac{1}{2})^{\binom{k}{2}}$

opposite $\overline{A}_R$

$\mathbb{P} = 1 - \binom{n}{k}\cdot 2 \cdot (\frac{1}{2})^{\binom{k}{2}}$

$\binom{n}{k} = \frac{n(n-1)\cdots\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!}$

$2^{1-\binom{k}{2}} = 2^{1-\frac{k(k-1)}{2}} = 2^{\frac{2-k^2+k}{2}}$

$$\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < \frac{2^{1+k/2}}{k!} \cdot \frac{n^k}{2^{k^2/2}} < 1$$

$k! \geq \frac{2^{1+\frac{k}{2}} \cdot n^k}{2^{\frac{k^2}{2}}}$

$= n^k \cdot 2^{1+\frac{k}{2}-\frac{k^2}{2}}$

let $n = \lfloor 2^{\frac{k}{2}} \rfloor$

as $k! > 2^{1+k/2}$ (exercise !). Hence $R(k, k) > 2^{k/2}$ for all $k \geq 3$. □

Remark. This simple example demonstrates the essence of the probabilistic method. To prove the existence of a good coloring we do not present one explicitly, but rather show, in a nonconstructive way, that it exists.

A *tournament* on a set V of n players is an *orientation* $T = (V, E)$ of the edges of the complete graph on the set of vertices V. Thus for every two distinct elements x and y of V, either (x, y) or (y, x) is in T, but not both. The name tournament is natural, since one can think of the set as a set of players in which each pair participates in a single match, where (x, y) is in the tournament iff x beats y.

We say that tournament T has the property S_k if for every set of k players there is one who beats them all. Is it true that for every finite k there is a tournament with property S_k?

Theorem 22. *If $\binom{n}{k}(1-2^{-k})^{n-k} < 1$, then there is a tournament on n vertices that has the property S_k.*

Proof. Consider a *random tournament* on the set $V = \{1, 2, \ldots, n\}$, which is obtained by choosing, for each $1 \le i < j \le n$, independently, either the arc (i, j) or the arc (j, i), where each of these two choices is equally likely. For every fixed subset K of size k of V, let A_K be the event that there is no vertex that beats all the members of K. Clearly, $Pr(A_K) = \left(1-(\frac{1}{2})^k\right)^{n-k}$. This is because for each fixed vertex $v \in V - K$, the probability that v does not beat all the members of K is $1 - 2^{-k}$, and all these $n - k$ events corresponding to the various possible choices of v are independent. It follows that

notation arc(i,j) means edge $i \to j$

for $\mathbb{P}(A_k)$

$1-(\frac{1}{2})^k$ means a point can not beat all the k players and then we do $\left(1-(\frac{1}{2})^k\right)^{n-k}$

$$Pr\Big(\bigcup_{\substack{K\subseteq V\\|K|=k}} A_K\Big) \le \sum_{\substack{K\subseteq V\\|K|=k}} Pr(A_K) = \binom{n}{k} \cdot (1-2^{-k})^{n-k} < 1.$$

means opposite $\bar{A}_k$ happen with a positive probability

Therefore, with positive probability no event A_K occurs, i.e., there is a tournament on n vertices that has property S_k. □

Remark 1. In the probabilistic method, Stirling's formula $n! = (\frac{n}{e})^n\sqrt{2\pi n}e^{\alpha_n}$, where $\frac{1}{12n+1} < \alpha_n < \frac{1}{12n}$, and the inequality $\binom{n}{k} < (\frac{en}{k})^k$ are frequently used.

Remark 2. Since $\binom{n}{k} < (\frac{en}{k})^k$ and $(1-2^{-k})^{n-k} < e^{-\frac{n-k}{2^k}}$, we have $\binom{n}{k}\cdot(1-2^{-k})^{n-k} \to 0$ as $n \to \infty$. So there always exists a tournament with the property S_k on n vertices as long as n is sufficiently large.

Let us now give a probabilistic proof of Bollabás Theorem on set systems (Theorem 14). Let $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^m$ be a family of pairs of subsets of an arbitrary set. Call $\mathcal{F}$ an (r, s)-*system* if $|A_i| = r$, $|B_i| = s$, $A_i \cap B_i = \emptyset$ for all $1 \le i \le m$, and $A_i \cap B_j \ne \emptyset$ for all $1 \le i \ne j \le m$.

Theorem 23 (Bollabás). *If $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^m$ is an (r, s)-system then $m \le \binom{r+s}{r}$.*

Proof. Put $X = \bigcup_{i=1}^{m}(A_i \cup B_i)$. Let $n = |X|$ and $C_i = X - (A_i \uplus B_i)$ for each i. Consider a random order π of X. For each i with $1 \le i \le m$, let X_i be the event that all elements of A_i precede those of B_i in this order. Clearly

there is no limit for C_i

$$Pr(X_i) = \frac{\overbrace{n(n-1)\ldots(n-|C_i|+1)}^{\text{there is no limit for } C_i}\cdot|A_i|!\cdot|B_i|!}{\underbrace{\boxed{n!}}_{n \text{ position}, |X|=n \text{ and } \pi \text{ is a random order}}} = \frac{1}{\binom{r+s}{r}}. \quad \binom{|A_i|+|B_i|}{|A_i|}$$

n position, $|X| = n$ and π is a random order

Now let us show that the events X_i are mutually exclusive. Assume the contrary: let π be an order in which all the elements of A_i precede those of B_i, and all the elements of

pairwise disjoint

for remark 1 : recall that $e^x > 1+x$ $\forall x \ne 0$, so

$\cdot\ (e^x)^n = e^{nx} > (1+x)^n > \binom{n}{k}x^k$ when $x > 0$

let $x = \frac{k}{n}$ $\cdot\ e^k > \binom{n}{k}(\frac{k}{n})^k$

$\Rightarrow \binom{n}{k} < (\frac{en}{k})^k$

for remark 2 .

$$\binom{n}{k}\cdot(1-2^{-k})^{n-k} \le \left(\frac{en}{k}\right)^k \cdot e^{-\frac{n-k}{2^k}}$$

$$= \frac{1}{k^k}\cdot e^{\left[k(1+\ln n) - \frac{n-k}{2^k}\right]} \longrightarrow 0 \quad \text{as } n \to \infty$$

$$(en)^k = e^{k\cdot\ln(en)}$$

$$\mathbb{P}(X_i) = \frac{n(n-1)\cdots\cdots(n-|C_i|+1)\cdot|A_i|!\cdot|B_i|!}{n!}$$

$$= \frac{|A_i|! \quad |B_i|!}{(n-|C_i|)(n-|C_i|-1)\cdots\cdots 3\cdot 2\cdot 1}$$

$$= \frac{r!\cdot s!}{(r+s)!} = \binom{r+s}{r}^{-1} = \frac{(r+S)(r+S-1)\cdots\cdots(r+S-r+1)}{r!}$$

A_j precede those of B_j. Suppose, without loss of generality, that the last element of A_i does not appear after the last element of A_j. Then all elements of A_i precede all those of B_j, contradicting the assumption that $A_i \cap B_j \neq \emptyset$. So all the events X_i are mutually exclusive.

It follows that $1 \geq Pr(\biguplus_{i=1}^{m} X_i) = \sum_{i=1}^{m} Pr(X_i) = m/\binom{r+s}{r}$. So we have $m \leq \binom{r+s}{r}$, completing the proof. □

相交

A pair of families $\mathcal{A}$, $\mathcal{B}$ is *cross-intersecting* if every set in $\mathcal{A}$ meets every set in $\mathcal{B}$. The *degree* of a point x in $\mathcal{A}$, denoted by $d_{\mathcal{A}}(x)$, is the number of sets in $\mathcal{A}$ containing x. The *rank* of $\mathcal{A}$ is the maximum cardinality of a set in $\mathcal{A}$.

$d_{\mathcal{A}}(x) = |\mathcal{A}|$

$d_{\mathcal{B}}(x) = \frac{|\mathcal{B}|}{a}$

If $\mathcal{A}$ has rank a, then, by the pigeonhole principle, each set in $\mathcal{A}$ contains a point x which is "popular" for the sets in $\mathcal{B}$, that is, $d_{\mathcal{B}}(x) \geq |\mathcal{B}|/a$. Similarly, if $\mathcal{B}$ has rank b, then each member of $\mathcal{B}$ contains a point y for which $d_{\mathcal{A}}(y) \geq |\mathcal{A}|/b$. However, this alone does not imply that we can find a point which is popular in both families $\mathcal{A}$ and $\mathcal{B}$. It turns out that if we relax the "degree of popularity" by one-half, then such a point exists.

Theorem 24 (Razborov-Vereshchagin). *Let $\mathcal{A}$ be a family of rank a and let $\mathcal{B}$ be a family of rank b. Suppose that the pair $\mathcal{A}, \mathcal{B}$ is cross-intersecting. Then there exists a point x such that*

$$d_{\mathcal{A}}(x) \geq \frac{|\mathcal{A}|}{2b} \quad \text{and} \quad d_{\mathcal{B}}(x) \geq \frac{|\mathcal{B}|}{2a}.$$

random variable

random set

$\mathbb{A}$. $\mathbb{B}$

Proof. Assume the contrary: $d_{\mathcal{A}}(x) < \frac{|\mathcal{A}|}{2b}$ or $d_{\mathcal{B}}(x) < \frac{|\mathcal{B}|}{2a}$ for each $\overset{\forall}{\ } x \in X$, where $X = \bigcup_{A \in \mathcal{A}} A \cup \bigcup_{B \in \mathcal{B}} B$. Now let $\mathbb{A}$ and $\mathbb{B}$ be independent random sets that are uniformly 均匀分布 distributed in $\mathcal{A}, \mathcal{B}$ respectively; that is, for each $A \in \mathcal{A}$ and $B \in \mathcal{B}$, $Pr(\mathbb{A} = A) = 1/|\mathcal{A}|$ and $Pr(\mathbb{B} = B) = 1/|\mathcal{B}|$. Since the pair $\mathcal{A}, \mathcal{B}$ is cross-intersecting, we have

$$\sum_{x \in X} Pr(x \in \mathbb{A} \cap \mathbb{B}) \geq 1. \tag{27}$$

Let X_0 consist of those points x for which $\frac{d_{\mathcal{A}}(x)}{|\mathcal{A}|} = Pr(x \in \mathbb{A}) < \frac{1}{2b}$ and let $X_1 = X - X_0$. By the assumption, for any $x \in X_1$, $Pr(x \in \mathbb{B}) = \frac{d_{\mathcal{B}}(x)}{|\mathcal{B}|} < \frac{1}{2a}$. By double

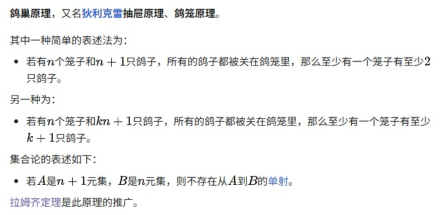
explain for pigeonhole principle

$\forall \tilde{A} \in \mathcal{A}$, we have $|A| \leq a$ (rank)

WTS $\forall A \in \mathcal{A}$, $\exists$ at least x s.t. $d_B(x) \geq \frac{|\mathcal{B}|}{a}$

so $\tilde{A} \cap B \neq \emptyset$, $\forall B \in \mathcal{B}$ which means

put a elements into $|\mathcal{B}|$ sets



鸽巢原理，又名狄利克雷**抽屉原理**、**鸽笼原理**。

其中一种简单的表述法为：

- 若有n个笼子和$n+1$只鸽子，所有的鸽子都被关在鸽笼里，那么至少有一个笼子有至少2只鸽子。

另一种为：

- 若有n个笼子和$kn+1$只鸽子，所有的鸽子都被关在鸽笼里，那么至少有一个笼子有至少$k+1$只鸽子。

集合论的表述如下：

- 若A是$n+1$元集，B是n元集，则不存在从A到B的单射。

拉姆齐定理是此原理的推广。

counting, $\sum_{x\in X} \boxed{d_{\mathcal{A}}(x)} = \sum_{x\in X} \boxed{\sum_{\substack{A\in\mathcal{A}\\ x\in A}} 1} = \sum_{A\in\mathcal{A}} \sum_{\substack{x\in X\\ x\in A}} 1 = \sum_{A\in\mathcal{A}} |A|$. Hence

independent and same distributed

$$\begin{aligned}
&\sum_{x\in X_1} Pr(x\in\mathbb{A}\cap\mathbb{B}) = \sum_{x\in X_1} Pr(x\in\mathbb{A})\cdot Pr(x\in\mathbb{B})\\
&< \frac{1}{2a}\sum_{x\in X_1} Pr(x\in\mathbb{A}) \le \frac{1}{2a}\cdot\sum_{x\in X} \boxed{Pr(x\in\mathbb{A})}\\
&= \frac{1}{2a}\cdot\sum_{x\in X} \boxed{\frac{d_{\mathcal{A}}(x)}{|\mathcal{A}|}} = \frac{1}{2a\cdot|\mathcal{A}|}\cdot\sum_{x\in X} d_{\mathcal{A}}(x)\\
&= \frac{1}{2a|\mathcal{A}|}\cdot\sum_{A\in\mathcal{A}} |A| \le \frac{a|\mathcal{A}|}{2a|\mathcal{A}|} = \frac{1}{2}.
\end{aligned}$$

Similarly, we can prove $\sum_{x\in X_0} Pr(x\in\mathbb{A}\cap\mathbb{B}) < \frac{1}{2}$. These two inequalities contradict (27) and thus the proof is complete. □

2° Linearity of Expectation

Let $X_1, X_2, \ldots, X_n$ be random variables and let $X = c_1X_1 + c_2X_2 + \ldots + c_nX_n$. *Linearity of expectation* states that

$$E[X] = c_1\cdot E[X_1] + c_2\cdot E[X_2] + \ldots + c_n\cdot E[X_n].$$

The power of this principle comes from there being no restrictions on the dependence or independence of the X_i's. In many situations, $E[X]$ can be easily calculated by a judicious decomposition into simple random variables X_i.

In applications, we often use the fact that there is a point in the probability space for which $X \ge E[X]$ and a point for which $X \le E[X]$.

Theorem 25 (Szele). *There is a tournament with n players and with at least $n!\cdot 2^{-(n-1)}$ Hamiltonian paths.*

K_n

Proof. Consider the random tournament on $V = \{1, 2, \ldots, n\}$, and let X be the number of Hamiltonian paths in it. For each permutation σ of $1, 2, \ldots, n$, let $X_\sigma = 1$ if σ gives a Hamiltonian path, that is, $(\sigma(i), \sigma(i+1))$ is an arc for each $i = 1, 2, \ldots, n-1$ and let $X_\sigma = 0$ otherwise. Then $X = \sum_\sigma X_\sigma$ and $E[X_\sigma] = 1\cdot Pr[X_\sigma = 1] = (\frac{1}{2})^{n-1}$. Hence

random variable on permutation

expectation → number of Hamilton roads

Bernoulli experiment with $\mathbb{P} = \frac{1}{2}$

$$E[X] = \sum_\sigma E[X_\sigma] = \underset{\text{number of n-permutation}}{n!}\cdot 2^{-(n-1)}.$$

and we only need to control n-1 edges

It follows that some tournament has at least $E[X] = n!2^{-(n-1)}$ Hamiltonian paths. □

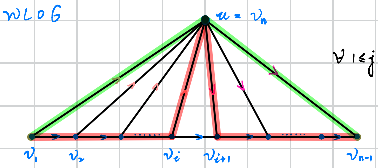
Theorem 26. *Let $G = (V, E)$ be a graph with e edges. Then G contains a bipartite subgraph with at least $e/2$ edges.*

for theorem 25, obtained by Szele in 1943, is considered the first use of the probabilistic method

thm: every tournament contains a Hamiltonian path
K_n

proof (sketch) by induction on # n of vertices

WLOG



$\forall 1 \le j \le i$: let i be the largest index s.t. arcs between u and v_j are all directed to u

after permutation

for thm 26. let (V, E) be a graph with e edges, then $G = (V, E)$ contains a bipartite subgraph with at least $e/2$ edges

Let $H = (X, Y; F)$ be a bipartite subgraph of $G = (V, E)$ s.t.

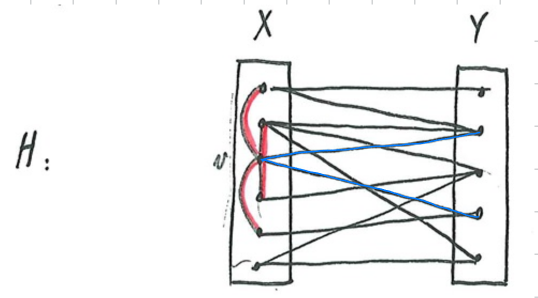
(a). $|X| + |Y| = |V|$

(b) the # of edges contained in H is maximized

then $d_H(v) \geq \frac{d_G(v)}{2}$, $\forall v \in V$

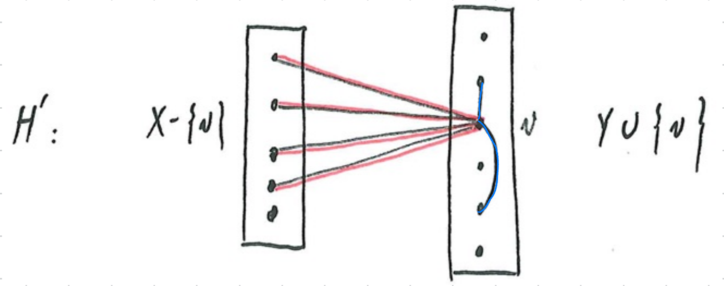
proof by contradiction, otherwise $\exists v \in V$ s.t. $d_H(v) < \frac{d_G(v)}{2}$

H:



let H' be the bipartite graph obtained from H by moving v from X to Y

H':



$\Longrightarrow$ # of edges in H' - # of edges in H

$= d_{H'}(v) - d_H(v)$

$= (d_G(v) - d_H(v)) - d_H(v)$

$= d_G(v) - 2 \cdot d_H(v) \geq 1$

random variable (indicator function)

consider number : expectation

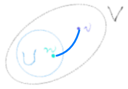
random variable over edge set

random subset (vertex)

Proof. Let $T \subseteq V$ be a random subset given by $Pr(x \in T) = 1/2$; these choices are mutually independent. Call an *edge xy crossing* if exactly one of x and y is in T. Let X be the number of crossing edges and let X_{xy} be the *indicator random variable* for xy being crossing; that is $X_{xy} = 1$ if xy is a crossing edge and 0 otherwise. Then $X = \sum_{xy \in E} X_{xy}$ and

$$E[X_{xy}] = Pr(X_{xy} = 1) = Pr(x \in T) \cdot Pr(y \notin T) + Pr(x \notin T) \cdot Pr(y \in T) = \frac{1}{2}.$$

Thus $E[X] = \sum_{xy \in E} E[X_{xy}] = e/2$. Hence $X \geq e/2$ for some choice of T, and the set of those crossing edges form a bipartite graph. □



A *dominating set* of a graph $G = (V, E)$ is a set $U \subseteq V$ such that each vertex $v \in V - U$ has at least one neighbor in U.

Theorem 27 (Alon). *Let $G = (V, E)$ be a graph on n vertices with minimum degree $\delta > 1$. Then G has a dominating set of at most $n[1 + \ln(\delta + 1)]/(\delta + 1)$ vertices.* $\frac{n[1+\ln(\delta+1)]}{\delta+1} \geqslant |U|$

Proof. Let us pick, randomly and independently, each vertex of V with probability p (to be determined). Let X be the random set of all vertices picked and let $Y = Y_X$ be the random set of all vertices in $V - X$ that do not have any neighbor in X. Then the expected value $E[|X|]$ of $|X|$ is np. 二项分布的期望

$Y = Y_X \subseteq V - X$

For each fixed vertex $v \in V$, $Pr(v \in Y) = Pr(v$ and its neighbors are all outside $X) \leq (1-p)^{\delta+1}$. Define $\chi_v = 1$ if $v \in Y$ and 0 otherwise. Then $E[\chi_v] = Pr(v \in Y) \leq (1-p)^{\delta+1}$. Thus $E[|Y|] = E[\sum_{v \in V} \chi_v] = \sum_{v \in V} E[\chi_v] \leq n(1-p)^{\delta+1} \leq ne^{-p(\delta+1)}$, as $1 + x \leq e^x$ for all x. Hence $E[|X| + |Y|] \leq np + ne^{-p(\delta+1)}$; this bound is minimized when $p = \frac{\ln(\delta+1)}{\delta+1}$. It follows that $\frac{n\ln(\delta+1)}{\delta+1} + ne^{-\ln(\delta+1)}$

here χ_V is the random variable what we define just like $\mathbb{1}_Y$

$$E[|X| + |Y|] \;\leq\; n \cdot \frac{1 + \ln(\delta + 1)}{\delta + 1}. \qquad (28)$$

From (28) we see that there is at least one choice of $X \subset V$ such that $|X| + |Y_X| \leq n \cdot \dfrac{1 + \ln(\delta + 1)}{\delta + 1}$. The set $U = X \cup Y_X$ is clearly a dominating set of G whose cardinality is at most $n[1 + \ln(\delta + 1)]/(\delta + 1)$. □

An *independent set* of a graph G is a set of pairwise nonadjacent vertices of G. The *independence number* $\alpha(G)$ of G is the maximum size of an independent set.

Theorem 28 (Turán). *Let G be a graph on n vertices and let d_i be the degree of the i^{th} vertex. Then $\alpha(G) \geq \sum_{i=1}^{n} \frac{1}{d_i + 1}$.*

Proof. Let $V = \{1, 2, \ldots, n\}$ and let $\pi : V \to V$ be a random permutation taking its values uniformly and independently with probability $1/n!$. This permutation corresponds for thm 27 , let $f(p) = p + e^{-p(\delta+1)}$ then WTS optimizer as $p = \frac{\ln(\delta+1)}{\delta+1}$

$$f'(p) = 1 - (\delta+1) \cdot e^{-p(\delta+1)}$$

$$f''(p) = (\delta+1)^2 \cdot e^{-p(\delta+1)} > 0$$ means $f'(p)$ monotone increasing

$f(p)$ is a convex function over $\mathbb{R}$ and hence minimized when $f'(p) = 0 \Longrightarrow f'(p) = 1 - (\delta+1) \cdot e^{-p(\delta+1)} = 0$

$$e^{-p(\delta+1)} = \frac{1}{\delta+1}$$

$$-p(\delta+1) = -\ln(\delta+1) \longrightarrow p = \frac{\ln(\delta+1)}{\delta+1}$$

three simple but important ideas are incorporated in the proof

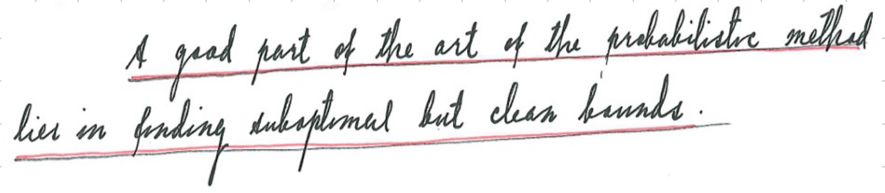
1. linearity of expectation

2. optimal choice of p

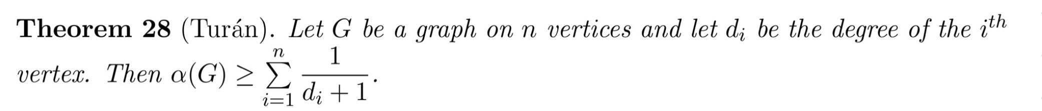
3. asymptotic calculus

we want the asymptotics of min $np + n(1-p)^{\delta+1}$, where p ranges over $[0, 1]$

the actual minimum $p = 1 - (\delta+1)^{-\frac{1}{\delta}}$ is difficult to deal with and in many similar cases precise minima are impossible to find in closed form. Rather, we give away a little bit, bounding $1 - p \leq e^{-p}$, yielding a clean bound


A good part of the art of the probabilistic method lies in finding suboptimal but clean bounds.

for thm 28


Theorem 28 (Turán). *Let G be a graph on n vertices and let d_i be the degree of the i^{th} vertex. Then $\alpha(G) \geq \sum_{i=1}^{n} \frac{1}{d_i + 1}$.*

let G be a graph on n vertices and let d be the average degree i.e. $d = \frac{\sum_{i=1}^{n} d_i}{n}$, then $\alpha(G) \geq \frac{n}{d+1}$

proof : note that $f(x) = \frac{1}{x+1}$ is a convex function when $x > 0$, so

$$f(d) = f\left(\sum_{i=1}^{n} d_i / n\right) \leq \frac{1}{n} \cdot \sum_{i=1}^{n} f(d_i)$$

$$\frac{1}{d+1} \leq \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{1}{d_i+1}$$

$$\boxed{\sum_{i=1}^{n} \frac{1}{d_i+1} \geq \frac{n}{d+1}}$$ for thm 29

to a random ordering of vertices in V. Let A_i be the event that $\pi(j) > \pi(i)$ for all neighbors j of i. There are $\binom{n}{d_i+1}$ possibilities to choose a (d_i+1)-element set $S \subseteq V$ of possible π-images of i and all its d_i neighbors. After that there are $(|S|-1)! = d_i!$ possibilities to arrange the π-images of neighbors of i within S (the place of $\pi(i)$ is fixed – it must come first), and $(n-|S|)! = (n-d_i-1)!$ possibilities to arrange the vertices outside S. Thus

$$Pr(A_i) = \binom{n}{d_i+1}\frac{d_i!(n-d_i-1)!}{n!} = \frac{1}{d_i+1}. \quad \frac{n(n-1)\cdots(n-d_i)}{(d_i+1)!} \quad \frac{d_i!(n-d_i-1)!}{n!}$$

event A_i means $\pi(j) > \pi(i)$ for all neighbour j of i so in this subset $S \cup \{i\}$ occur only when i is arranged at first

Let U be the set of those vertices i for which A_i holds. By linearity of expectation

$$E[|U|] = \sum_{i=1}^{n} Pr(A_i) = \sum_{i=1}^{n} \frac{1}{d_i+1}.$$

Thus, for some specific ordering π, $|U| \geq \sum_{i=1}^{n} \frac{1}{d_i+1}$. Now let ij be an edge of G. Then either $\pi(i) < \pi(j)$ or $\pi(j) < \pi(i)$. In the first case $j \notin U$ and in the second case $i \notin U$. So U is an independent set. □

3° The Deletion Method

As described in a previous section, the probabilistic method works as follows: In order to prove that a structure with certain desired property exists, one defines an appropriate probability space of structures and then shows that the desired properties hold in this space with positive probability. In this section, we consider the situations where the "random" structure does not have all the desired properties but may have a few "blemishes". With a small alternation, we remove the blemishes, getting the desired structure.

It follows from Theorem 28 that if G has n vertices and $nd/2$ edges, then the independence number $\alpha(G) \geq n/(1+d)$. We can get "halfway" to this result with the deletion method.

Theorem 29. *If a graph $G = (V, E)$ has n vertices and $nd/2$ edges, $d \geq 1$, then $\alpha(G) \geq \frac{n}{2d}$.*

random take $v \in V \longrightarrow S$

Proof. Let $S \subseteq V$ be a random subset defined by $Pr[v \in S] = p$, where p is to be determined, and the events $v \in S$ are mutually independent. Let $X = |S|$ and let Y be the number of edges contained in S. For each edge $e = ij$ of G, let Y_e be the indicator random variable for the event $i, j \in S$; that is, $Y_e = 1$ if $i, j \in S$ and 0 otherwise. Then $E[Y_e] = Pr[i,j \in S] = Pr[i \in S] \cdot Pr[j \in S] = p^2$. Since $Y = \sum_{e \in E} Y_e$, by linearity of expectation

mutually independent same distributed

$$E[Y] = \sum_{e \in E} E[Y_e] = \frac{nd}{2}p^2.$$

二项分布
伯努力实验

Clearly $E[X] = np$, so, again by linearity of expectation.

$$E[X - Y] = np - \frac{nd}{2}p^2.$$

to maximize means we can calculate derivative

We set $p = 1/d$ (here using $d \geq 1$) to maximize the quantity, giving

$$E[X - Y] = \frac{n}{2d}.$$

$X = |S|$

Thus there exists a specific S for whom the number of vertices in S minus the number Y of edges in S is at least $n/2d$. Select one vertex from each edge of S and delete it. This leaves a set S^* with at least $n/2d$ vertices. All edges having been destroyed, S^* is an independent set. □

Recall that the *girth* of a graph G, denoted by girth(G), is the length of its shortest cycles. Again, we let $\alpha(G)$ stand for the independence number of G and $\chi(G)$ for the chromatic number of G.

Theorem 30 (Erdös). *For any integers $k, \ell \geq 1$, there exists a graph G with*

$$\text{girth}(G) > \ell \quad \text{and} \quad \chi(G) > k.$$

G是一个n-vertices的随机图
每一对顶点都以prob p
独立连成边 edge

Proof. Fix $0 < \theta < 1/\ell$ and let $G \sim G(n, p)$ with $p = n^{\theta - 1}$; that is, G is a random graph on n vertices chosen by picking each pair of vertices as an edge randomly and independently with probability p. Let X be the number of cycles of length at most ℓ. How many cycles of length $i, v_1v_2 \ldots v_iv_1$, can G have?

1 There are $n(n-1)\ldots(n-i+1)$ sequences $v_1, v_2, \ldots, v_i$ of distinct vertices, and each cycle is identified by $2i$ of these sequences: there are two possibilities to choose the "direction" and i possibilities to choose the first vertex of the cycle. Thus, for $3 \leq i \leq \ell$ there are $n(n-1)\ldots(n-i+1)/2i$ potential cycles of length i, each of which is in G with probability p^i. By linearity of expection

in a cycle with i vertices has i edges and each edge corresponding prob p

$$E[X] = \sum_{i=3}^{\ell} \frac{n(n-1)\ldots(n-i+1)}{2i} \cdot p^i \leq \sum_{i=3}^{\ell} \frac{\overset{n^{\theta i} \ll n}{n^{\theta i}}}{2i} = o(n). \tag{29}$$

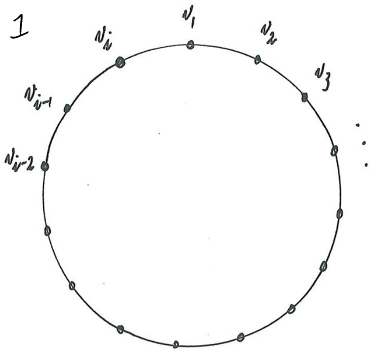
as $\theta\ell < 1$. Recall Markov's inequality: if X is a nonnegative random variable, then for $a > 0$, $Pr[X \geq a] \leq E[X]/a$. Thus by (29), we have

$$Pr[X \geq n/2] \leq \frac{o(n)}{n/2} = o(1). \tag{30}$$

Setting $x = \lceil \frac{3}{p} \ln n \rceil$ gives

$$\begin{aligned} Pr[\alpha(G) \geq x] &\leq \binom{n}{x}(1-p)^{\binom{x}{2}} \leq n^x e^{-p(x-1)x/2} \qquad (\text{as } 1 - p \leq e^{-p}) \\ &= e^{[\ln n - p(x-1)/2]\cdot x} \\ &\leq e^{[\ln n - p(\frac{3}{p}\ln n - 1)/2]x} \\ &= e^{(-\frac{1}{2}\ln n + \frac{p}{2})x}. \end{aligned} \tag{31}$$

1



cycle with length i

<u>Clockwise Direction</u> 顺时针方向

- $v_1\ v_2\ v_3\ \cdots\ v_{i-2}\ v_{i-1}\ v_i\ v_1$
- $v_2\ v_3\ v_4\ \cdots\ v_{i-1}\ v_i\ v_1\ v_2$
- $v_3\ v_4\ v_5\ \cdots\ v_i\ v_1\ v_2\ v_3$

$\cdots$

- $v_i\ v_1\ v_2\ \cdots\ v_{i-3}\ v_{i-2}\ v_{i-1}\ v_i$

$\Big\}\ i$

<u>Counterclockwise Direction</u> 逆时针方向

- $v_i\ v_{i-1}\ v_{i-2}\ \cdots\ v_3\ v_2\ v_1\ v_i$
- $v_{i-1}\ v_{i-2}\ v_{i-3}\ \cdots\ v_2\ v_1\ v_i\ v_{i-1}$
- $v_{i-2}\ v_{i-3}\ v_{i-4}\ \cdots\ v_1\ v_i\ v_{i-1}\ v_{i-2}$

$\cdots$

- $v_1\ v_i\ v_{i-1}\ \cdots\ v_4\ v_3\ v_2\ v_1$

$\Big\}\ i$

Markov's inequality

if X is a random variable that takes only nonnegative values, then for any value $a>0$

$$\Pr(X \geq a) \leq \frac{E(X)}{a}$$

proof for any $a>0$ let $I = \begin{cases} 1, & \text{if } X \geq a \\ 0, & \text{else} \end{cases}$

since $X \geq 0$, we have $I \leq \frac{X}{a}$ so $E(I) \leq \frac{E(X)}{a}$

which, as $E(I) = \Pr(X \geq a)$ proves the result

formula 31
$$\begin{aligned} Pr[\alpha(G) \geq x] &\leq \binom{n}{x}(1-p)^{\binom{x}{2}} \leq n^x e^{-p(x-1)x/2} \qquad (\text{as } 1-p \leq e^{-p}) \\ &= e^{[\ln n - p(x-1)/2]\cdot x} \\ &\leq e^{[\ln n - p(\frac{3}{p}\ln n - 1)/2]x} \\ &= e^{(-\frac{1}{2}\ln n + \frac{p}{2})x}\ . \end{aligned}$$

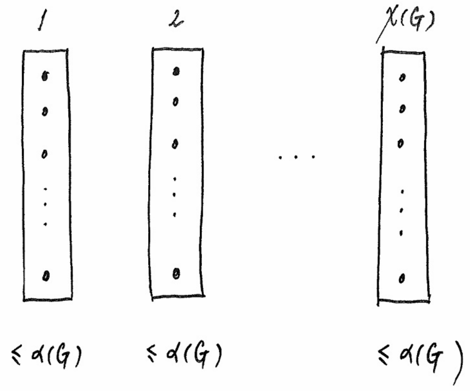
substitute $x = \left\lceil \frac{3}{p} \ln n \right\rceil$

for any 2 points, can't form an edge in total $\binom{x}{2}$ edges

since $x \doteq 3 \cdot n^{1-\theta} \ln n$, then $e^{(-\frac{1}{2}\ln n + \frac{p}{2})x} = e^{(-\frac{1}{2}\ln n + \frac{p}{2})(3\cdot n^{1-\theta}\ln n)} \xrightarrow{n \to +\infty} -\infty$

$|G| \leq \chi(G) \cdot \alpha(G)$ for any graph G

proof by definition, G has a $\chi(G)$-coloring of vertices. Since each color class is an independent set which contains at most $\alpha(G)$ vertices, then $|G| \leq \chi(G) \cdot \alpha(G)$



Since $x \doteq 3n^{1-\theta} \ln n$, $(-\frac{1}{2} \ln n + \frac{p}{2})x \to -\infty$ as $n \to \infty$. By (31) we get

$$Pr[\alpha(G) \geq x] = o(1) \quad \text{as} \quad n \to \infty. \tag{32}$$

$\mathbb{P}(X > \frac{n}{2}) \xrightarrow{n \to \infty} 0$

Now let n be sufficiently large so that $Pr[X \geq n/2] < 0.5$ and $Pr[\alpha(G) \geq x] < 0.5$ (recall (30) & (32)). Then there is a specific G on n vertices with less than $n/2$ cycles of length at most ℓ and with $\alpha(G) \leq \frac{3}{p} \ln n = 3 \cdot n^{1-\theta} \ln n$. Remove from G a vertex from each cycle of length at most ℓ. This gives a graph G^* with at least $\frac{n}{2}$ vertices. Now girth$(G^*) > \ell$ and $\alpha(G^*) \leq \alpha(G)$. Thus

$$\chi(G^*) \geq \frac{|G^*|}{\alpha(G^*)} \geq \frac{n/2}{3 \cdot n^{1-\theta} \ln n} = \frac{n^\theta}{6 \ln n}, \quad > k$$

where $|G^*|$ is the number of vertices in G^*. The proof is complete by taking n so that $\frac{n^\theta}{6 \ln n} > k$. □

4° The Lovász Local Lemma

Let $A_1, A_2, \ldots, A_n$ be events in a probability space. In combinatorial applications the A_i are "bad" events. We wish to show that $Pr(\bigcap_{i=1}^{n} \overline{A}_i) > 0$ so that there is a point (e.g. coloring, tournament, configuration) which is "good". The basic probabilistic method can be written as the following.

question: is it possible to give an explicit construction of the dependency graph?

Counting Sieve. *If $\sum_{i=1}^{n} Pr(A_i) < 1$, then $Pr(\bigcap_{i=1}^{n} \overline{A}_i) > 0$.*

in many applications, i and j are adjacent iff A_i and A_j are dependent. But it is not always true,

There are other simple conditions that ensure $Pr(\bigcap_{i=1}^{n} \overline{A}_i) > 0$.

in general, A_i is mutually independent of $\{A_j,\ ij \notin E\}$

Independence Sieve. *If $A_1, A_2, \ldots, A_n$ are mutually independent and all $Pr(A_i) < 1$, then $Pr(\bigcap_{i=1}^{n} \overline{A}_i) > 0$.*

that is. A_i is independent of all conbinations of $\cup, \cap$ these A_j and $\overline{A}_j$

The Lovász local lemma is a sieve method which allows for some dependence among the A_i. A graph $G = (V, E)$ with vertex set $V = \{1, 2, \ldots, n\}$ (the indices for the A_i) is called a *dependency graph* for $A_1, A_2, \ldots, A_n$ if for all i, A_i is mutually independent of all A_j with $ij \notin E$; that is, A_i is independent of any combinations of $\cup, \cap$, these A_j and $\overline{A}_j$.

Theorem 31 (Lovász Local Lemma; The General Case). *Let $A_1, A_2, \ldots, A_n$ be events with dependency graph G. Suppose there exist $x_1, x_2, \ldots, x_n \in [0, 1)$ with*

$$Pr(A_i) \leq x_i \cdot \prod_{ij \in E(G)} (1 - x_j). \tag{32}$$

in general $P(A|Bc) = \frac{P(AB|c)}{P(B|c)}$

proof : $\frac{P(AB|c)}{P(B|c)}$

conditional probability $= \frac{P(ABc)}{P(c)} \div \frac{P(Bc)}{P(c)}$

$= \frac{P(ABc)}{P(Bc)} = P(A|Bc)$

for all i. Then

$$Pr(\bigcap_{i=1}^{n} \overline{A}_i) \geq \prod_{i=1}^{n} (1 - x_i) > 0$$

Proof. We show by induction on s that for all i, S with $|S| = s$ and $i \notin S$

$$\text{WTS} \quad Pr[A_i | \bigcap_{j \in S} \overline{A}_j] \leq x_i.$$

For $s = 0$, $Pr[A_i] \leq x_i \cdot \prod_{ij \in E(G)} (1 - x_j) \leq x_i$ is immediate. Let us proceed to the induction step. Clearly we may assume i is adjacent to some $j \in S$. Now renumber the indices so that $i = n$, $S = \{1, 2, \ldots, s\}$ and among $j \in S$, $ij \in E(G)$ for $1 \leq j \leq d$. Write

设在S中与i相邻的点的个数为d

$$Pr[A_n | \overline{A}_1 \overline{A}_2 \ldots \overline{A}_s] = \frac{Pr[A_n \overline{A}_1 \ldots \overline{A}_d \mid \overline{A}_{d+1} \ldots \overline{A}_s]}{Pr[\underbrace{\overline{A}_1 \ldots \overline{A}_d}_{\text{adjacency with } A_n} \mid \underbrace{\overline{A}_{d+1} \ldots \overline{A}_s}_{\text{independent with } A_n}]}.$$

$P(X|Y \cap Z) = \frac{P(X \cap Y|Z)}{P(Y|Z)}$

We bound the numerator as follows:

$$Pr[A_n \overline{A}_1 \ldots \overline{A}_d \mid \overline{A}_{d+1} \ldots \overline{A}_s] \leq Pr[A_n \mid \overline{A}_{d+1} \ldots \overline{A}_s] = Pr[A_n]$$

as A_n is mutually independent of $A_{d+1}, A_{d+2}, \ldots, A_s$. The denominator can be bounded by the induction hypothesis:

$$\boxed{Pr[\overline{A}_1 \ldots \overline{A}_d \mid \overline{A}_{d+1} \ldots \overline{A}_s] = \prod_{i=1}^{d} Pr[\overline{A}_i \mid \overline{A}_{i+1} \ldots \overline{A}_s]} \quad \text{conditional prob'y}$$
$$\geq \prod_{i=1}^{d} (1 - x_i) \quad \text{(by induction hypothesis)}$$

$$\boxed{\begin{aligned} &\prod_{i=1}^{d} Pr[\bar{A}_i \mid \bar{A}_{i+1} \cdots \bar{A}_s] \\ &= \frac{Pr[\bar{A}_1 \cdots \bar{A}_s]}{Pr[\bar{A}_2 \cdots \bar{A}_s]} \cdot \frac{Pr[\bar{A}_2 \cdots \bar{A}_s]}{Pr[\bar{A}_3 \cdots \bar{A}_s]} \cdots \frac{Pr[\bar{A}_d \cdots \bar{A}_s]}{Pr[\bar{A}_{d+1} \cdots \bar{A}_s]} \\ &= \frac{Pr[\bar{A}_1 \cdots \bar{A}_s]}{Pr[\bar{A}_{d+1} \cdots \bar{A}_s]} \\ &= Pr[\bar{A}_1 \cdots \bar{A}_d \mid \bar{A}_{d+1} \cdots \bar{A}_s]. \end{aligned}}$$

Thus we have the quotient

$x_n \cdot \prod_{nj \in E(G)} (1 - x_j)$ by assumption

$$Pr[A_n \mid \overline{A}_1 \ldots \overline{A}_s] \leq \frac{Pr[A_n]}{\prod_{ni \in E(G)} (1 - x_i)} \leq x_n,$$

completing the induction. Hence

$$\boxed{Pr[\overline{A}_1 \ldots \overline{A}_n] = \prod_{i=1}^{n} Pr[\overline{A}_i \mid \overline{A}_1 \ldots \overline{A}_{i-1}]}$$
$$\geq \prod_{i=1}^{n} (1 - x_i) > 0,$$

$$\boxed{\begin{aligned} &\prod_{i=1}^{n} Pr[\bar{A}_i \mid \bar{A}_1 \cdots \bar{A}_{i-1}] \\ &= \cancel{Pr[\bar{A}_1]} \cdot \frac{\cancel{Pr[\bar{A}_1 \bar{A}_2]}}{\cancel{Pr[\bar{A}_1]}} \cdot \frac{\cancel{Pr[\bar{A}_1 \bar{A}_2 \bar{A}_3]}}{\cancel{Pr[\bar{A}_1 \bar{A}_2]}} \\ &\qquad \cdots \frac{Pr[\bar{A}_1 \cdots \bar{A}_n]}{\cancel{Pr[\bar{A}_1 \cdots \bar{A}_{n-1}]}} \\ &= Pr[\bar{A}_1 \cdots \bar{A}_n]. \end{aligned}}$$

as desired. □

Theorem 32 (Lovász Local Lemma; The Symmetric Case). *Let $A_1, A_2, \ldots, A_n$ be events with dependency graph G such that $Pr[A_i] \leq p$ and degree $(i) \leq d$ for all i. If $ep(d+1) \leq 1$, then $Pr[\bigcap_{i=1}^{n} \overline{A}_i] > 0$.*

Remark. e is best possible.

Proof. If $d = 0$, then $A_1, A_2, \ldots, A_n$ are mutually independent. So the result follows from the independent sieve.

If $d \geq 1$, then $\frac{1}{e} \leq (1 - \frac{1}{d+1})^d$ (exercise!). Hence

we already know

$\lim_{n \to \infty} (1 - \frac{1}{n})^n = \frac{1}{e}$

so $\lim_{d \to \infty} (1 - \frac{1}{d+1})^d = \frac{1}{e}$

for all finite d, we have $(1 - \frac{1}{d+1})^d > \frac{1}{e}$

$$\begin{aligned} Pr[A_i] &\leq p \leq \frac{1}{e(d+1)} \\ &\leq \frac{1}{d+1}(1 - \frac{1}{d+1})^d \\ &\leq x_i \cdot \prod_{ij \in E(G)} (1 - x_j), \end{aligned}$$

where $x_j = \frac{1}{d+1}$ for $j = 1, 2, \ldots, n$, and so (32) holds. By virtue of the preceding theorem, we have $Pr[\bigcap_{i=1}^{n} \overline{A}_i] > 0$. □

Let us now give some applications of the Lovász local lemma. There is no known proof of any of these results, which does not use the local lemma.

A *hypergraph* is a pair $H = (V, E)$, where V is a finite set whose elements are called *vertices* and E is a family of subsets of V, called *edges*. We say that *H has property B*, or that it is *2-colorable* if there is a 2-coloring of V such that no edge is monochromatic.

Theorem 33. *Let $H = (V, E)$ be a hypergraph in which each edge has at least k elements and intersects at most d other edges. If $e(d+1) \leq 2^{k-1}$, then H has property B.*

Proof. Color each vertex of H, randomly and independently, by either blue or red (with equal probability). For each edge $f \in E$, let A_f be the event that f is monochromatic. Clearly, $Pr[A_f] = 2/2^{|f|} \leq 1/2^{k-1}$. Moreover, each event A_f is clearly mutually independent of all the other events $A_{f'}$ for all edges f' that do not intersect f. The result now follows from Theorem 32. □

Theorem 34. *Let $D = (V, E)$ be a simple digraph with minimum outdegree δ and maximum indegree Δ. If $e(\Delta\delta + 1)(1 - \frac{1}{k})^\delta < 1$, then D contains a (simple directed) cycle of length $0 \pmod{k}$.*

Proof. Clearly we may assume that every outdegree is precisely δ, since otherwise we can consider a subgraph of D with this property.

Let $f : V \to \{0, 1, \ldots, k-1\}$ be a random coloring of V, obtained by choosing, for each $v \in V$, $f(v) \in \{0, 1, \ldots, k-1\}$ independently, according to a uniform distribution. For each $v \in V$, let A_v denote the event that there is no $u \in V$, with $(v, u) \in E$ and $f(u) \equiv (f(v) + 1) \pmod{k}$. Clearly $Pr[A_v] = (1 - \frac{1}{k})^\delta$. It is easy to check that each event A_v is mutually independent of all the events A_u but those satisfying

$$N^+(v) \cap (\{u\} \cup N^+(u)) \neq \emptyset,$$

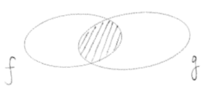
hyper graph

thm 33. construction of the dependency graph

1. each vertex f stands for A_f. $\forall f \in E$

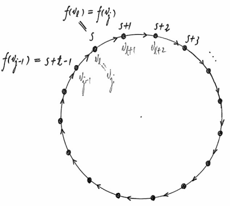
2. two vertices f and g are adjacency iff A_f and A_g are dependent

$\forall f, g \in E \iff f \cap g \neq \phi$



by def of j. $\nexists$ $\ell \leq p < q \leq j-1$

s.t. $v_p = v_q$. so v_ℓ, $v_{\ell+1}$ ...... v_{j-1} are distinct

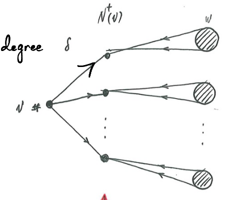


implies $s \equiv (s+t-1)+1 \mod k$

$t \equiv 0 \mod k$

$t = c \cdot k$. $c \in \mathbb{R}$

assume outdegree δ



$\leq \Delta - 1$

$\leq \Delta - 1$

$\leq \Delta - 1$

maximal indegree is Δ

number of such u's $\leq \delta + (\Delta-1)\delta = \Delta\delta$

where $N^+(v) = \{w \in V : (v,w) \in E\}$. The number of such u's is at most $\Delta\delta$ and hence, by our assumption and by Theorem 32, $Pr(\bigcap_{v \in V} \overline{A}_v) > 0$. Therefore there is an $f : V \to \{0, 1, \ldots, k-1\}$ such that for every $v \in V$, there is $u \in V$ with

$$(v,u) \in E \quad \text{and} \quad f(u) \equiv (f(v)+1) \pmod{k} \tag{33}$$

Starting at an arbitrary $v = v_0$ and applying (33) repeatedly, we obtain a sequence $v_0, v_1, v_2, \ldots$ of vertices of D so that $(v_i, v_{i+1}) \in E$ and $f(v_{i+1}) \equiv (f(v_i)+1) \pmod{k}$ for all $i \geq 0$. Let j be the smallest integer so that there is an $\ell < j$ with $v_\ell = v_j$. The cycle $v_\ell v_{\ell+1} \ldots v_j = v_\ell$ is a directed simple cycle of D whose length is divisible by k. □

Now let us apply the general form of the Lovász local lemma to give a lower bound for $R(3,k)$.

Theorem 35. *$R(3,k) > c \cdot \dfrac{k^2}{\ln^2 k}$ for some positive constant c.*

Proof. Color the edges of K_n independently with each edge red with probability p and blue with probability $1-p$. For each 3-set S let A_S be "S is red" and for each k-set T let B_T be "T is blue ". Then

$$Pr(A_S) = p^3, \quad Pr(B_T) = (1-p)^{\binom{k}{2}} \sim e^{-pk^2/2}.$$

Let S and S' be adjacent in the dependency graph if they have a common edge; the same holds for S, T or T, T'. Each S is adjacent to only $3(n-3) \sim 3n$ other S'. Each T is adjacent to less than $\binom{k}{2}n < k^2n/2$ of the S. We use only that each S or T is adjacent to at most $\binom{n}{k}$ – that is all – of T. Suppose that with each A_S we associate the same $x_S = x$, and with each B_T we associate the same $y_T = y$. According to Theorem 31, if there exist p, x, y with

$$p^3 \leq x(1-x)^{3n}(1-y)^{\binom{n}{k}}$$

and

$$(1-p)^{\binom{k}{2}} \leq y(1-x)^{k^2n/2}(1-y)^{\binom{n}{k}},$$

then $R(3,k) > n$.

Our objective is to find the largest possible $k = k(n)$ for which there is such a choice of p, x, y. An elementary but tedious computation shows that the best choice is

$$\begin{aligned} p &= c_1 n^{-1/2}, & k &= c_2 n^{1/2} \ln n \\ x &= c_3/n^{3/2}, & y &= c_4/\exp(n^{1/2} \ln^2 n), \end{aligned}$$

where c_1, c_2, c_3, c_4 are some positive constants. This implies that $R(3,k) > c_5 \dfrac{k^2}{\ln^2 k}$. □

$$\frac{k^{1/\alpha}}{\ln^{\beta/\alpha} k} = \frac{c^{1/\alpha}\, n \ln^{\beta/\alpha} n}{(\ln c + \alpha \ln n + \beta \ln\ln n)^{\beta/\alpha}}$$

$$\le n \cdot \frac{c^{1/\alpha}}{\alpha^{\beta/\alpha}}$$

$$\triangleq \frac{n}{c'}$$

$$\Rightarrow n \ge c' \cdot \frac{k^{1/\alpha}}{\ln^{\beta/\alpha} k}.$$

2nd Method

$$k^{1/\alpha} = c^{1/\alpha}\, n \ln^{\beta/\alpha} n \Rightarrow k^{1/\alpha} \ge \theta n$$

$$\Rightarrow n \ge c' \frac{k^{1/\alpha}}{\ln^{\beta/\alpha} k}$$

Remark. In general, we can show that if $k = c \cdot n^\alpha \ln^\beta n$ for some positive constants c, α, β, then there exists a positive constant c' such that $\boxed{n \ge c' \frac{k^{1/\alpha}}{\ln^{\beta/\alpha} k}}$.

5° The Second Moment Method

After the expectation, the most vital statistic for a random variable X is the variance, which is defined by $\text{var}\,(X) = E[(X - E[X])^2]$ and measures how spread out X is from its expectation. As usual, let μ denote the expectation of X and σ^2 denote the variance. The positive square root σ of the variance is called the *standard deviation*. With this notation, here is our basic tool.

Chebyschev's Inequality. *For any positive constant λ,*

$$Pr[|X - \mu| \ge \lambda\sigma] \le \frac{1}{\lambda^2}.$$

The use of Chebyschev's inequality is called the *second moment method.*

A set $\{x_1, x_2, \ldots, x_k\}$ of positive integers is said to *have distinct sums* if all sums $\sum_{i \in S} x_i, S \subseteq \{1, 2, \ldots, k\}$, are distinct. Let $f(n)$ denote the maximal k for which there exists a set $\{x_1, x_2, \ldots, x_k\} \subseteq \{1, 2, \ldots, n\}$ with distinct sums.

The simplest example of a set with distinct sums is $\{2^i : i \le \log_2 n\}$. It shows $f(n) \ge 1 + \lfloor \log_2 n \rfloor$. Erdös offered \$300 for a proof or disproof that $f(n) \le \log_2 n + c$ for some constant c.

Theorem 36. $f(n) < \log_2 n + \frac{1}{2} \log_2 \log_2 n + O(1)$.

Proof. Let $\{x_1, x_2, \ldots, x_k\} \subseteq \{1, 2, \ldots, n\}$ be a set with distinct sums. Let $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k$ be independent random variables with $Pr[\varepsilon_i = 1] = Pr[\varepsilon_i = 0] = 1/2$. Let $X = \varepsilon_1 x_1 + \varepsilon_2 x_2 + \ldots + \varepsilon_k x_k, \mu = E[X]$, and $\sigma^2 = \text{var}\,[X]$. Then

$$\mu = \frac{x_1 + x_2 + \ldots + x_k}{2}$$

and

$$\sigma^2 = \frac{x_1^2 + x_2^2 + \ldots + x_k^2}{4} \le \frac{n^2 k}{4}.$$

So $\sigma \le n\sqrt{k}/2$. By Chebyschev's inequality, for any constant $\lambda > 1$

$$Pr[|X - \mu| \ge \frac{\lambda n \sqrt{k}}{2}] \le \lambda^{-2},$$

which implies that

$$1 - \frac{1}{\lambda^2} \le Pr[|X - \mu| < \frac{\lambda n \sqrt{k}}{2}]. \tag{34}$$

Since $\{x_1, x_2, \ldots, x_k\}$ is a set having distinct sums, X has any particular value with probability either 0 or 2^{-k}. Thus

$$Pr[|X - \mu| < \frac{\lambda n\sqrt{k}}{2}] \leq 2^{-k}\lambda n\sqrt{k}. \tag{35}$$

Combining (34) and (35), we obtain

$$n \geq \frac{2^k}{\sqrt{k}} \cdot \frac{1 - \lambda^{-2}}{\lambda}. \tag{36}$$

Claim. If $n \geq c \cdot \frac{2^k}{k^\alpha}$ for some positive constants c and α, then

$$k \leq \log_2 n + \alpha \log_2 \log_2 n + O(1). \tag{37}$$

The desired inequality follows instantly from (36) and our claim. Let us now justify the claim by contradiction.

Assume to the contrary that (37) does not hold. Then there exists three sequences $\{\beta_i\}$, $\{n_i\}$, and $\{k_i\}$ such that

- $0 < \beta_i < \beta_{i+1}$ for each i and $\beta_i \to \infty$ as $i \to \infty$;
- $n_i < n_{i+1}$ and $\log_2 n_i \geq 2\beta_i$ for each i;
- $k_i < k_{i+1}$ for each i;
- $n_i \geq c \cdot \frac{2^{k_i}}{k_i^\alpha}$ and $k_i > \log_2 n_i + \alpha \log_2 \log_2 n_i + \beta_i$ for each i.

Since $c \cdot \frac{2^k}{k^\alpha}$ is a strictly increasing function of k when k is sufficiently large, we have

$$\begin{aligned}
n_i &\geq c \cdot \frac{2^{k_i}}{k_i^\alpha} \\
&\geq c \cdot \frac{2^{\log_2 n_i + \alpha \log_2 \log_2 n_i + \beta_i}}{(\log_2 n_i + \alpha \log_2 \log_2 n_i + \beta_i)^\alpha} \\
&= n_i \frac{2^{\beta_i} c \cdot (\log_2 n_i)^\alpha}{(\log_2 n_i + \alpha \log_2 \log_2 n_i + \beta_i)^\alpha} \\
&\geq (\frac{2^{\beta_i} c}{1.5^\alpha} - o(1)) n_i \qquad \text{as } \log_2 n_i \geq 2\beta_i \\
&> n_i \quad \text{as } \beta_i \to \infty \text{ when } i \to \infty,
\end{aligned}$$

a contradiction. □

As an application of the second moment method, let us now prove the famous Weierstrass approximation theorem, which asserts that the set of real polynomials over $[0, 1]$ is dense in the space of all continuous functions over $[0, 1]$.

Theorem 37 (Weierstrass Approximation Theorem). *For every continuous real function $f : [0, 1] \to R$ and every $\varepsilon > 0$, there is a polynomial $p(x)$ such that $|p(x) - f(x)| < \varepsilon$ for all $x \in [0, 1]$.*

Proof. Since a continuous function $f : [0,1] \to R$ is uniformly continuous, there is a $\delta > 0$ such that if x, $y \in [0,1]$ and $|x-y| \leq \delta$, then $|f(x) - f(y)| \leq \varepsilon/2$. In addition, since f is bounded, there is an $M > 0$ such that $|f(x)| \leq M$ in $[0,1]$.

For any $x \in (0,1)$, let $B(n,x)$ denote the binomial random variable with n independent trials and probability of success x for each of them. Recall that the probability that $B(n,x) = j$ is precisely $\binom{n}{j} x^j (1-x)^{n-j}$, the expectation of $B(n,x)$ is nx, and the standard deviation is $\sqrt{nx(1-x)} \leq \sqrt{n}$. In view of Chebyschev's inequality, for every integer n

$$Pr(|B(n,x) - nx| > n^{2/3}) \leq \frac{1}{n^{1/3}}.$$

Let n be a positive integer such that $1/n^{1/3} < \text{Min}\,\{\delta, \frac{\varepsilon}{4M}\}$. Then

$$Pr(|B(n,x) - nx| > n^{2/3}) < \frac{\varepsilon}{4M} \qquad \text{for all } x \in (0,1).$$

Define

$$P_n(x) = \sum_{i=0}^{n} \binom{n}{i} x^i (1-x)^{n-i} f(\frac{i}{n}).$$

We claim that for every $x \in [0,1]$, $|p_n(x) - f(x)| \leq \varepsilon$. Since $p_n(0) = f(0)$ and $p_n(1) = f(1)$, it remains to justify the claim for $x \in (0,1)$.

Indeed, since $\sum\limits_{i=0}^{n} \binom{n}{i} x^i (1-x)^{n-i} = 1$, we have

$$\begin{aligned}
|p_n(x) - f(x)| &\leq \sum_{i:\ |i-nx| \leq n^{2/3}} \tbinom{n}{i} x^i (1-x)^{n-i} |f(\frac{i}{n}) - f(x)| + \\
&\quad + \sum_{i:\ |i-nx| > n^{2/3}} \tbinom{n}{i} x^i (1-x)^{n-i} (|f(\frac{i}{n})| + |f(x)|) \\
&\leq \sum_{i:\ |\frac{i}{n}-x| \leq n^{-1/3} < \delta} \tbinom{n}{i} x^i (1-x)^{n-i} |f(\frac{i}{n}) - f(x)| + \\
&\quad + Pr(|B(n,x) - nx| > n^{2/3}) \cdot 2M \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4M} \cdot 2M = \varepsilon,
\end{aligned}$$

completing the proof. □