The Probabilistic Method

Techniques

- **Union bound**
- Argument from expectation
- Alterations
- The second moment method
- The (Lovasz) Local Lemma

And much more

- Alon and Spencer, "The Probabilistic Method"
- Bolobas, "Random Graphs"

The Union Bound Technique: Main Idea

- A: event our structure exists, want Prob$[A] > 0$ or Prob$[\bar{A}] < 1$
- Suppose $\bar{A}$ implies one of $B_1, \cdots, B_n$ must hold
- (Think of the B_i as "bad" events)
- Then, by the union bound

$$\text{Prob}\left[\bar{A}\right] \leq \text{Prob}\left[\bigcup_i B_i\right] \leq \sum_i \text{Prob}[B_i]$$

- Thus, as long as

$$\sum_i \text{Prob}[B_i] < 1$$

our structure exists!

We have seen this used in d-disjunct matrix examples.

Example 1: Nice Tournaments

- A tournament is an orientation G of K_n
- Think of $u \to v$ as *"player u beats player v"*
- Fix integer k, G is *nice* if for every k-subset S of players there is another v who beats all of S
- Intuitively, nice tournaments may exist for large n
  (Remember the theme? "Sufficiently large space contains locally nice structures")

Existence of Nice Tournaments (Erdős, 1963)

- For every $\{u, v\}$, let $u \to v$ with probability $1/2$
- A: event that a random G is nice
- $\bar{A}$ implies $\bigcup_{|S|=k} B_S$ where $B_S =$ "S is not beaten by **any** $v \notin S$"

$$\text{Prob}[B_S] = \left(1 - \frac{1}{2^k}\right)^{n-k}$$

- Hence, nice tournaments exist as long as $\binom{n}{k}\left(1 - \frac{1}{2^k}\right)^{n-k} < 1$
- What's the order of n for which this holds?

$$\text{use } \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k \text{ and } \left(1 - \frac{1}{2^k}\right)^{n-k} < e^{-\frac{n-k}{2^k}}$$

- Nice tournaments exist as long as $\left(\frac{ne}{k}\right)^k e^{-\frac{n-k}{2^k}} < 1$.
- So, $n = \Omega\left(k^2 \cdot 2^k\right)$ is large enough!

Example 2: 2-coloring of uniform hypergraphs

- Given a k-uniform hypergraph $G = (V, E)$, i.e.
  - E is a collection of k-subsets of V
- G is 2-colorable iff each vertex in V can be assigned with red or blue such that there's no monochromatic edge
- Intuitively, if $|E|$ is small then G is 2-colorable!
- Question is: "how small?"
- An answer may be obtained along the line: "for n *small enough*, a random 2-coloring is good with positive probability"

Theorem (Erdős, 1963)

Every k-uniform hypergraph with $< 2^{k-1}$ edges is 2-colorable!

Example 3: Error-Correcting Codes

- Message $\mathbf{x} \in \{0,1\}^k$
- Encoding $f(\mathbf{x}) \in \{0,1\}^n$, $n > k$, f an injection
- $C = \{f(\mathbf{x}) \mid \mathbf{x} \in \{0,1\}^k\}$: codewords
- $f(\mathbf{x})$ is sent over noisy channel, few bits altered
- $\mathbf{y}$ is received instead of $f(\mathbf{x})$
- Find codeword $\mathbf{z}$ "closest" to $\mathbf{y}$ in Hamming distance
- Decoding $\mathbf{x}' = f^{-1}(\mathbf{z})$
- Measure of utilization: relative rate of C

$$R(C) = \frac{\log |C|}{n}$$

- Measure of noise tolerance: relative distance of C

$$\delta(C) = \frac{\min_{\mathbf{c}_1 \neq \mathbf{c}_2 \in C} \mathsf{Dist}(\mathbf{c}_1, \mathbf{c}_2)}{n}$$

Linear Codes

- For any $\mathbf{x} \in \mathbb{F}_2^n$, define

$$\text{WEIGHT}(\mathbf{x}) = \text{ number of 1-coordinates of } \mathbf{x}$$

- E.g., $\text{WEIGHT}(1001110) = 4$
- If C is a k-dimensional subspace of $\mathbb{F}_2^n$, then

$$\begin{aligned} |C| &= 2^k \\ \delta(C) &= \min\{\text{WEIGHT}(\mathbf{x}) \mid \mathbf{x} \in C\} \end{aligned}$$

- Every such C can be defined by a parity check matrix $\mathbf{A}$ of dimension $(n-k) \times n$:

$$C = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

- Conversely, every $(n-k) \times n$ matrix $\mathbf{A}$ defines a code C of dimension $\geq k$

A Communication Problem

Large rate and large distance are conflicting goals

Problem

Does there exist a family of codes C_k, $|C_k| = 2^k$, for infinitely many k, such that

$$R(C_k) \geq R_0 > 0$$

and

$$\delta(C_k) \geq \delta_0 > 0$$

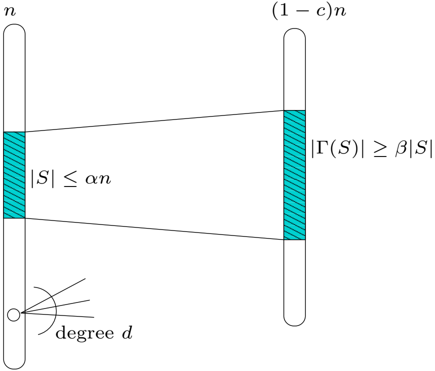
(Yes, using "magical graphs.")

Practicality

Design such a family explicitly, such that the codes are efficiently encodable and decodable.

Magical Graph

(n, c, d, α, β)-graph



c, d, α, β are constants, n varies.

From Magical Graphs to Code Family

- Suppose (n, c, d, α, β)-graphs exist for infinitely many n, and constants c, d, α, β such that $\beta > d/2$
- Consider such a $G = (L \cup R, E)$, $|L| = n, |R| = (1-c)n = m$
- Let $\mathbf{A} = (a_{ij})$ be the $m \times n$ 01-matrix, column indexed by L, and row-indexed by R, $a_{ij} = 1$ iff $(i, j) \in E$
- Define a linear code with $\mathbf{A}$ as parity check:

$$C = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

- Then, $\dim(C) = n - \text{rank}(A) \geq cn$, and

$$|C| = 2^{\dim(C)} \geq 2^{cn} \Rightarrow R(C) \geq c$$

- For every $\mathbf{x} \in C$, $\text{WEIGHT}(\mathbf{x}) \geq \alpha n$, hence

$$\delta(C) = \frac{\min\{\text{WEIGHT}(\mathbf{x}) \mid \mathbf{x} \in C\}}{n} \geq \alpha$$

Existence of Magical Graph with $\beta > d/2$

- Determine n, c, d, α, β later
- Let $L = [n], R = [(1-c)n]$.
- Choose each of the d neighbors for $u \in L$ uniformly at random
- For $1 \leq s \leq \alpha n$, let B_s be the "bad" event that some subset S of size s has $|\Gamma(S)| < \beta|S|$
- For each $S \subset L$, $T \subset R$, $|S| = s, |T| = \beta s$, define

$$X_{S,T} = \begin{cases} 1 & \Gamma(S) \subseteq T \\ 0 & \Gamma(S) \not\subseteq T \end{cases}$$

- Then,

$$\text{Prob}[B_s] \leq \text{Prob}\left[\sum_{S,T} X_{S,T} > 0\right] \leq \sum_{S,T} \text{Prob}[X_{S,T} = 1]$$

Existence of Magical Graph with $\beta > d/2$

$$
\begin{aligned}
\text{Prob}[B_s] &\leq \binom{n}{s}\binom{(1-c)n}{\beta s}\left(\frac{\beta s}{(1-c)n}\right)^{sd} \\
&\leq \left(\frac{ne}{s}\right)^s \left(\frac{(1-c)ne}{\beta s}\right)^{\beta s}\left(\frac{\beta s}{(1-c)n}\right)^{sd} \\
&= \left[\left(\frac{s}{n}\right)^{d-\beta-1}\left(\frac{\beta}{1-c}\right)^{d-\beta} e^{\beta+1}\right]^s \\
&\leq \left[\left(\frac{\alpha\beta}{1-c}\right)^{d-\beta}\cdot\frac{e^{\beta+1}}{\alpha}\right]^s
\end{aligned}
$$

Choose $\alpha = 1/100$, $c = 1/10$, $d = 32$, $\beta = 17 > d/2$,

$$\text{Prob}[B_s] \leq 0.092^s$$

Existence of Magical Graph with $\beta > d/2$

The probability that such a randomly chosen graph is **not** an (n, c, d, α, β)-graph is at most

$$\sum_{s=1}^{\alpha n} \mathsf{Prob}[B_s] \leq \sum_{s=1}^{\infty} 0.092^s = \frac{0.092}{1 - 0.092} < 0.11$$

Not only such graphs exist, there are **a lot** of them!!!