Lecture 32

Last time

- concentration of measure 'on a slice'.

Lem Suppose $f: \{0,1,\cdots,q-1\}^n \to \mathbb{R}$ satisfies the bounded differences condition w/ parameter $(c_1,\cdots,c_n)$ and that η is drawn uniformly at random from $\{0,1,\ldots,q-1\}^n$ subject to $wt(\eta)=k$. Then

$$\mathbb{P}\left(|f(\eta)-\mathbb{E}f(\eta)|\geq t\right)\leq 2e^{-\frac{t^2}{68\sum c_i^2}} \quad \forall t\geq 0.$$

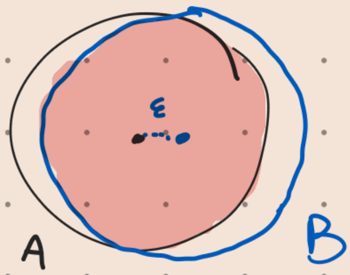
§ intersection volume

First

consider intersection of two balls in $\mathbb{R}^3$.

Take $\varepsilon>0$ suff. small and consider two unit balls A and B whose centers are of distance ε apart, then $\mathrm{vol}(A\cap B)\geq 99\%\ \mathrm{vol}(A)$

But this is not longer true in high dimension



Prop: Let $\varepsilon>0$. Then there exists $n_0=n_0(\varepsilon)$ s.t. the following holds for all $n\geq n_0$.

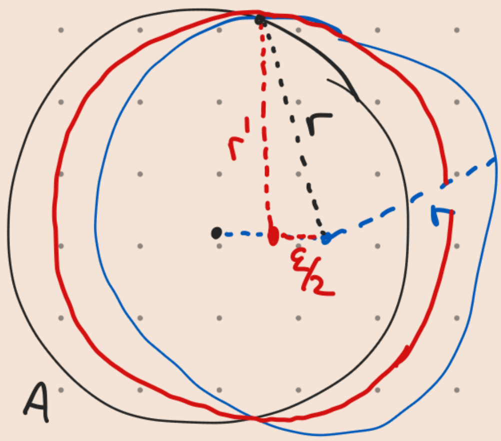
Let A and B be two unit balls whose centers are of distance ε apart, then

$$\frac{\mathrm{vol}(A \cap B)}{\mathrm{vol}(A)} < 1\%.$$

vol. of radius-r ball in $\mathbb{R}^n$

<u>Fact</u>: $\mathrm{vol}_n(r) \sim \frac{1}{\sqrt{n\pi}} \left(\frac{2\pi e}{n}\right)^{n/2} \cdot r^n$

<u>Pf</u>: $r' < r - \varepsilon/10$



$\Rightarrow \mathrm{vol}(A \cap B) < \mathrm{vol}_n(r')$

$$\frac{\mathrm{vol}(A \cap B)}{\mathrm{vol}(A)} < \frac{\mathrm{vol}_n(r')}{\mathrm{vol}_n(r)} = \left(\frac{r'}{r}\right)^n$$

$$< \left(1 - \frac{\varepsilon}{10r}\right)^n \longrightarrow 0 \text{ as } n \to \infty. \quad \blacksquare$$

- This idea of using a ball of smaller radius to bound the intersection in the continuous $\mathbb{R}^n$ setting fails in many discrete settings.

<u>Example</u>: $\{0,1\}^n$ w/ Hamming metric.

<u>Def</u>: $\Delta(x,y) = \#$ coordinates where they differ.

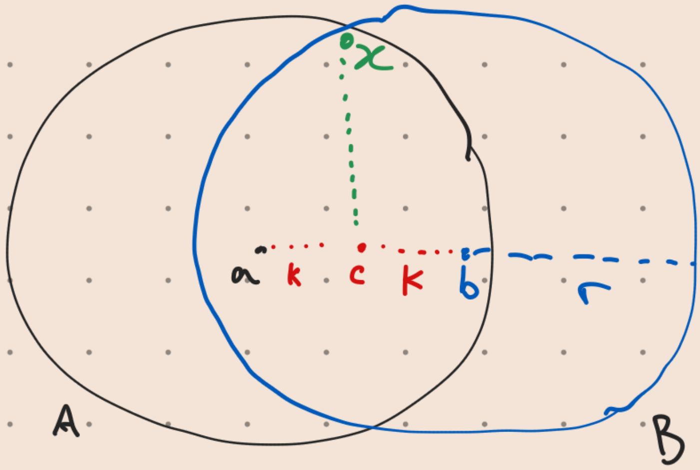
Take $2k \le r$.

$a = 00\cdots0\cdots0 = 0^n$

$b = 111111\,00\cdots0 = 1^{2k}0^{n-2k}$

$c = 111000\cdots0 = 1^k 0^{n-k}$

$x = 000111\cdots100\cdots0 = 0^k 1^r 0^{n-k-r}$



$$\Delta(x,a) = \Delta(x,b) = r$$
$$\implies x \in A \cap B$$

BUT... $\Delta(x,c) = r+k > r$.

Recall the result from last time. It gives natural suff. cond. for (X,d) to guarantee intersection volume is small.

Thm (Kim-L.-Tran) Let (X,d) be a finite metric space w/ d taking values in $\mathbb{N}\cup\{0\}$ and let $k,r\in\mathbb{N}$. Suppose
(A1) (X,d) has exponential growth at radius r w/ rate $c>0$;
(A2) (X,d) is (r,k)-dispersed w/ constant $\alpha>0$;
(A3) certain centered r.v. defined on a 'slice' is K-subgaussian.

$\forall\ d(a,b)=k$

$$\implies \frac{\mathrm{vol}(B(a,r)\cap B(b,r))}{\mathrm{vol}(B(a,r))} \le 2e^{-\Omega_{c,\alpha}\left(k+\frac{k^2}{K}\right)}$$

Thm (*) Let $0 < p < \frac{q-1}{q}$ and let $k \in \mathbb{N}$.

Consider $X = \{0, 1, \dots, q-1\}^n$ endowed w/ the Hamming metric Δ. Then (X, Δ) satisfies the following

(A1) (X, Δ) has exponential growth at radius pn w/ rate $c = \Omega_{p,q}(1)$.

(A2) (X, Δ) is (pn, k)-dispersed w/ constant $\alpha = \frac{1}{2}\left(1 - \frac{pq}{q-1}\right) > 0$

(A3) For any $a, b \in X$ w/ $\Delta(a,b) = k$ and any $0 \le i \le \alpha k$, $\underbrace{\ell_{a,b}(x)}_{\Delta(x,b) - \Delta(x,a)} - \mathbb{E}\,\ell_{a,b}(x)$ is $400k$-subgaussian

where x is drawn uniformly from $S(a, pn - i)$.

Consequently, $\forall\, a, b \in X$,

$$\frac{\mathrm{vol}(B(a,r) \cap B(b,r))}{\mathrm{vol}(B(a,r))} = 2e^{-\Omega_{p,q}(\Delta(a,b))}$$

The conditions (A1)-(A3) can be easily verified. In particular, (A3) follows from the concentration of measure on a slice.

This exponential decay on intersection volume for balls in Hamming space has applications on
- list-decodability of random codes.
- improvements on Gilbert-Varshamov type bounds.

— Error correction code:

By encoding messages w/ codewords that are pairwise far apart, we can recover the message even if some bits are corrupted by noise.

$YES = 000100$

$NO = \underbrace{111111}_{\text{length}=6}$

Receive $= 011100 = x$

$\Delta(x, YES) = 2$

$\Delta(x, NO) = 3.$

$\Rightarrow$ Alice said "YES".

In general, if the noise could corrupt up to t bits, then as long as all messages are encoded by strings w./ distance $\geq 2t+1$.

<u>Def</u>: $A(n,d) =$ max # ~~messages~~ codewords in $\{0,1\}^n$ w./ min. distance d.

<u>Thm</u> Gilbert-Varshamov bound.

$$A(n,d+1) \geq \frac{2^n}{\mathrm{vol}(n,d)},$$ where

$$\mathrm{vol}(n,d) = \sum_{i=0}^{d} \binom{n}{i}$$ is the volume of radius-d ball.

<u>Thm</u> (Turán) G N-vx. D-regular

$$\Rightarrow \alpha(G) \geq \frac{N}{D+1}.$$

K_{D+1} . . . . .

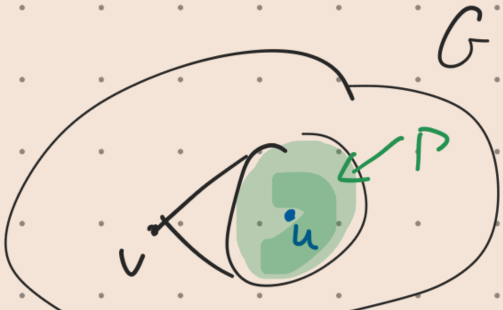
<u>Exer</u>. Prove G-V bound using Turán.

We can improve GV bound, if we get a better bound on indep. #.

- Ajtai-Komlós-Szemerédi '80s proved that if G $\begin{cases} \cdot\ \Delta\text{-free} \\ \cdot\ \text{locally sparse} \end{cases}$ $\Rightarrow$ $\alpha(G) \geq_c \frac{N}{D} \log D$.

<u>Thm</u> Let G be an N-vx graph w/ max. deg D and min deg $\geq D/2$. Let $K \in [D]$ and let $\Gamma \subseteq G$ be a subgraph induced by the neighbourhood of an arbitrary vertex.

Suppose there is a partition $V(\Gamma) = B \cup I$ s.t.



- $\forall u \in B$, $\deg_\Gamma(u) \leq D/K$; and
- $|I| \leq D/K$

$\Rightarrow$ $\alpha(G) \geq \left(1 - o_K(1)\right) \frac{N}{D} \cdot \log K$, and

$\#$ indep. sets in G is $\geq e^{\left(\frac{1}{8} + o_K(1)\right) \frac{N}{D} \log^2 K}$.

To apply the above thm to improve GV bound, we use Thm (*) to check that certain auxiliary graph is locally sparse.