Lecture 14

Last time

- Community detection continued
- Nets, coverings, and packings.

Today

- Covering numbers via volume
- Detour: error correcting codes

Covering numbers via volume

In most applications we will be interested in how big covering numbers are. Today we will see a simple argument to bound these numbers when

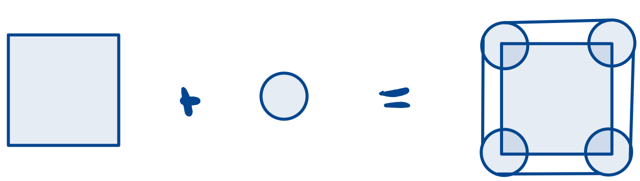
$$T = \mathbb{R}^n \quad \text{and} \quad d(x,y) = \|x-y\|_2.$$

We will relate them to volume.

We need a definition before that.

Def (Minkowski sum): Let A and B be subsets of $\mathbb{R}^n$. The Minkowski sum $A+B$ is given by

$$A+B = \{a+b \mid a \in A,\ b \in B\}.$$



Proposition ☺: Let K be a subet of $\mathbb{R}^n$ and $\varepsilon > 0$. Then,

Established last lecture

$$\frac{\mathrm{Vol}(K)}{\mathrm{Vol}(\varepsilon B_2^n)} \leq \mathcal{N}(K,\varepsilon) \leq \mathcal{P}(K,\varepsilon) \leq \frac{\mathrm{Vol}(K+\frac{\varepsilon}{2}B_2^n)}{\mathrm{Vol}(\frac{\varepsilon}{2}B_2^n)}.$$

$B_2^n = \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$.

We drop the dependency on d for brevity.

Proof: We start by proving the lower bound. Let $N = \mathcal{N}(K,\varepsilon)$, then K can be covered by N balls with radii ε. Comparing volumes gives

$$\mathrm{Vol}(K) \leq N \cdot \mathrm{Vol}(\varepsilon B_2^n).$$

For the upper bound, take $N = \mathcal{P}(K,\varepsilon)$. Notice that if we take $\varepsilon/2$ balls around the packing yields a set of nonintersecting balls contained in $K + \frac{\varepsilon}{2}B_2^n$, thus

$$N \cdot \mathrm{Vol}(\tfrac{\varepsilon}{2}B_2^n) \leq \mathrm{Vol}(K),$$

which establishes the result. $\square$

We leverage this argument to obtain upper and lower bounds for the ball.

Corollary ☺: The covering number of the unit ball B_2^n satisfies

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(B_2^n, \varepsilon) \leq \left(\frac{2}{\varepsilon}+1\right)^n \leq \left(\frac{3}{\varepsilon}\right)^n \quad \forall \varepsilon \in (0,1)$$

The same upper bound applies for S^{n-1}.

Proof: Recall that

$$\mathrm{Vol}(\varepsilon B_2^n) = \varepsilon^n \mathrm{Vol}(B_2^n)$$

thus the lower bound follows by Proposition ☺. By the same proposition

$$\mathcal{N}(B_2^n, \varepsilon) \leq \frac{\mathrm{Vol}((1+\varepsilon/2)B_2^n)}{\mathrm{Vol}(\frac{\varepsilon}{2}B_2^n)}$$

$$\leq \frac{(1+\varepsilon/2)^n}{(\varepsilon/2)^n}$$

$$= \left(\frac{2}{\varepsilon}+1\right)^n$$

$$\leq \left(\frac{3}{\varepsilon}\right)^n.$$

The same proof applies for the sphere. $\square$

Covering numbers measure the "complexi

ty of sets. We will see later that often $\log \mathcal{N}(K, \varepsilon)$ is a useful quantity to understand, and it is known as the metric entropy of K.

Error correcting codes

Suppose Ayush wants to send a message to Barbara with k letters

$$x := \text{"bring snacks"}$$

But an adversary corrups Ayush's message by changing $r = 2$ letter, and Barbara receives

$$y := \text{"bring snakes."}$$

How do we ensure that we can recover the correct message? A natural idea is to use redundancy: Ayush encodes his k-letter message into a longer n-letter message.

Example (Repeating Code): Ayush might just repeat the message

$$E(x) = \text{"bring snacks| bring snacks| bring snacks"}$$

Barbara can use majority decoding: check the received copies of each letter in $E(x)$ and pick the one that appears more often. If the message x is repeated $2r+1$ times, this strategy will recover x correctly (even if r letters of $E(x)$ are corrupted).

check!

The issue with this strategy is that it is very inefficient: it requires $n \geq (2r+1)k$.

Indeed, we shall show that n can be much smaller. First we formalize the notion of an error correcting code. For simplicity (an also for practical utility), consider a binary alphabet.

Def: An error correcting code that encodes k-bit strings into n-bit strings and can correct r errors consists of an encoding map

$E : \{0,1\}^k \to \{0,1\}^n$ and an decoding $D : \{0,1\}^n \to \{0,1\}^k$ such that

$D(y) = x$

for all $x \in \{0,1\}^k$ and $y \in \{0,1\}^n$ s.t. y differs from $E(x)$ in at most r bits. $\dashv$

In turn, the binary cube is a metric space.

Hamming cube

Lemma: The set $H = \{0,1\}^n$ and distance

Hamming distance $d_H(x,y) = \#\{ i \mid x_i \neq y_i \}$.

form a metric space. $\dashv$

Proposition $\square$ Consider $K = \{0,1\}^n$ and let $m \in [n]$, then

$$\frac{2^n}{\sum_{k=0}^{m} \binom{n}{k}} \leq \mathcal{N}(K, d_H, m) \leq \mathcal{P}(K, d_H, m) \leq \frac{2^n}{\sum_{k=0}^{\lfloor m/2 \rfloor} \binom{n}{k}}.$$

$\dashv$

Hint: Use the volume argument, now with cardinality.

Lemma $\boxtimes$: Suppose $k, n, r \in \mathbb{N}$ s.t.

(★) $\log_2 \mathcal{P}(\{0,1\}^n, d_H, 2r) \geq k$.

Then, there exists an error correcting code that encodes k-bit strings into n-bit strings and corrects up to r errors.

Proof: By (★) there exists a $2r$ separated set $\mathcal{N} \subseteq \{0,1\}^n$ s.t. $|\mathcal{N}| = 2^k$. Thus, the closed balls of radius r centered at the points in $\mathcal{N}$ are disjoint. Let $E: \{0,1\}^k \to \mathcal{N}$ be any 1-1 map, and let $D: \{0,1\}^n \to \{0,1\}^k$ be a nearest neighbor decoder.

$D(y) = x_0$ with $\min\limits_{x_0 \in \{0,1\}^k} d_H(E(x_0), y)$.

If $y \in \{0,1\}^n$ is within distance r of the true message, then this strategy returns the true message.

$\square$

Theorem: Suppose that $k, n, r \in \mathbb{N}$ s.t.

$$n \geq k + 2r \log_2\left(\frac{en}{2r}\right). \tag{11}$$

Then, there axists an error correcting code that encodes k-bit strings into n-bit strings and corrects up to r errors.

Proof: Invoking Lemma $\boxtimes$ together with Proposition $\square$ yields the conclusion follows since

$$
\begin{aligned}
\mathcal{P}(\{0,1\}^n, d_H, 2r) &\geq \mathcal{N}(\{0,1\}^n, d_H, 2r) \\
&\geq \frac{2^n}{\sum_{i=0}^{2r} \binom{n}{i}} \\
&\overset{\text{Sterling's approx}}{\geq} 2^n \left(\frac{2r}{en}\right)^{2r} \\
&\overset{(11)}{\geq} 2^k.
\end{aligned}
$$

$\square$

Thus, n grows linearly with r (ignoring log terms) as opposed to rk (recall our original naive idea).