MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070J Lecture 9 Mar 13, SP17

Lecturer: Yury Polyanskiy Scribe notes by Ryan Zheyuan Shi

Isoperimetry

Content.

1 Introduction

Theorem 1 (Isoperimetric inequality in the plane)**.** *In $\mathbb{R}^2$, the circle encloses the maximum area among all curves of given length.*

One proof uses Steiner symmetrization. The proof roughly consists of the following steps.

Step 1: Without loss of generality, the region enclosed by the curve under consideration is convex. If not, we can always make it convex and decrease the length and increase the area enclosed.

Step 2: Given any convex set, we may pick a point in the enclosed region and a line passing through it. We divide the enclosed region into slices. From Minkowski's inequality, it can be shown that centering all slices around the line increases the area of the enclosed region.

Step 3: We repeatedly apply step 2. The only curve that is fixed under such symmetrization is a circle.

However, this proof breaks down if there exists no optimal closed curve.

2 Harper's Theorem

In this section, we consider the Hamming space $\mathcal{X}^n = \{0,1\}^n$ with the Hamming distance d_H.

Definition 1 (Blow up)**.** *For $A \subseteq \mathcal{X}^n$, the r-blow up of A is the set*

$$\Gamma^r A = \{x : d_H(x, A) \leq r\}.$$

And as a reminder, $d_H(x, A) = \min_{y \in A} d_H(x, y)$.

Definition 2 (Hamming ball)**.** *The Hamming ball centered at c with radius m is given by $\mathcal{B}(c, m) = \{x : d_H(x, c) \leq m\}$. A set C is called a quasiball if $\mathcal{B}(c, m) \subseteq C \subsetneq \mathcal{B}(c, m+1)$.*

Theorem 2 (Harper's theorem)**.** *For $A \subseteq \mathcal{X}^n$, if $|A| \geq \sum_{k=0}^{m} \binom{n}{k}$, then $|\Gamma^r A| \geq \sum_{k=0}^{m+r} \binom{n}{k}$. Hamming ball of radius k shows this is the best possible.*

This theorem follows trivially follows (by taking $B = (\Gamma^r A)^c$) from the following result:

Theorem 3. *For any given $A, B \in \mathcal{X}^n$ there exists a pair of quasiballs A' and B' centered at 0^n and 1^n respectively, such that $|A| = |A'|$, $|B| = |B'|$ and $d_H(A', B') \leq d_H(A, B)$.*

Proof (Part 1). We first introduce the idea of shifting. We define

$$S_i^+(B) = \bigcup_{b \in B} \{b' = (1_i, b_{\sim i}) \text{ if } b_i = 0 \text{ and } b' \notin B, \text{otherwise } b' = b\}$$

And similarly, define

$$S_i^-(A) = \bigcup_{a \in A} \{a' = (1_i, a_{\sim i}) \text{ if } a_i = 1 \text{ and } a' \notin A, \text{otherwise } a' = a\}$$

To illustrate the notion of shifting, consider the following example.

$$\text{If } B = \begin{pmatrix} [0, & 0] \\ [1, & 1] \end{pmatrix}, \text{ then, } S_1^+ B = \begin{pmatrix} [1, & 0] \\ [1, & 1] \end{pmatrix}, S_2^+ B = \begin{pmatrix} [0, & 1] \\ [1, & 1] \end{pmatrix}$$

Observe that in the above example, $S_2^+ S_1^+ B = S_1^+ B$, and in general $S_2^+ S_1^+ B \neq S_1^+ S_2^+ B$. Thus, the shifting operation is not commutative. Also note that shifting does not change the size of a set. That is, we have $|B| = |S_i^+ B|, |A| = |S_i^- A|$.

We take two sets A, B, and claim that $d_H(S_i^+ B, S_i^- A) \geq d_H(A, B)$. Suppose not, denote $A_1 = S_i^- A, B_1 = S_i^+ B$. Suppose there exist $a \in A_1, b \in B_1$, such that $d_H(a, b) < d_H(A, B)$. Since a, b are either obtained by shifting or remain unshifted, without loss of generality, we discuss the following cases.

Case 1: a and b are both shifted. We have $a_i^{old} = 1, a_i^{new} = 0, b_i^{old} = 0, b_i^{new} = 1$. Thus, we observe that $d_H(a^{old}, b^{old}) = d_H(a^{new}, b^{new})$. Thus, this case cannot happen.

Case 2: a is shifted, b is not shifted. We have $a_i^{old} = 1, a_i^{new} = 0$. If $b_i = 1$, then we are even increasing the distance, which in ruled out. If $b_i = 0$, then there exists some b' which prevents b's move. In particular, $b'_i = 1$ and $b'_{\sim i} = b_{\sim i}$. Thus, we have $d_H(a^{new}, b) = d_H(a^{old}, b')$. This also invalidates the assumption. □

We will now prove a few lemmas and return to the proof of Theorem 3 later.

Lemma 4. *If C is such that $C = S_i^+ C, \forall i$, then C is monotone increasing (up-set). That is, if $x \in C, y \geq x$ coordinate-wise, then $y \in C$.*

Proof. Take $x \in C$. Suppose we have $y > x$ coordinate-wise. Then, x and y must be of the following form

$$y = (1, 1, 1, \ldots, \text{same})$$

$$x = (0, 0, 0, \ldots, \text{same})$$

In $S_i^+ C$, x can be moved to $(1, 0, 0, \ldots, \text{same})$. But since $C = S_i^+ C$, we know $(1, 0, 0, \ldots, \text{same}) \in C$. Continue with this argument, all points on the shifting path from x to y must be in C. Thus, $y \in C$. □

We define a more general version of the shifting.

Definition 3. *Given $U, V \subseteq [n], |U| < |V|, U \cap V = \emptyset$, we define*

$$S^+_{U,V}(B) = \bigcup_{b \in B} \{b' = (1_V, 0_U, b_W) \text{ if } b_V = 0_V, b_U = 1_U, \text{ otherwise } b' = b\}$$

$$S^-_{U,V}(B) = \bigcup_{b \in B} \{b' = (1_U, 0_V, b_W) \text{ if } b_U = 0_U, b_V = 1_V, \text{ otherwise } b' = b\}$$

where $W = [n] - U - V$.

We illustrate what this operation does below

$S^+_{U,V}$	U	V	W
b	$(1, 1$	$0, 0, 0, 0$	Stuff)
b'	$(0, 0$	$1, 1, 1, 1$	Same Stuff)

$S^-_{U,V}$	U	V	W
b	$(0,0$	$1,1,1,1$	Stuff)
b'	$(1,1$	$0,0,0,0$	Same Stuff)

We state but not prove the following 2 lemmas.

Lemma 5. *If C is such that $S^+_{U,V}C = C, \forall U, V$, then C is a quasi-ball centered at 1^n (all-one vector).*

Proof. Indeed, if C is not a quasi-ball, then there must exist a pair $x \in C$ and $y \notin C$ such that $|x|_H < |y|_H$. Then take $U = \{i : x_i = 1, y_i = 0\}$ and $V = \{i : x_i = 0, y_i = 0\}$. Then $S^+_{U,V}C \neq C$, a contradiction. □

Lemma 6. *Given U, V, if A, B are fixed under $S^{\pm}_{U',V'}, \forall |U'|+|V'| < |U|+|V|$, then we have $d_H(A_1, B_1) \geq d_H(A, B)$, where $A_1 = S^-_{U,V}A, B_1 = S^+_{U,V}B$.*

Proof. The proof resembles the case-work in part 1 of the proof for Theorem 3. Namely, assume for contradiction that $d_H(A_1, B_1) < d_H(A, B)$ and let $a \in A_1$, $b \in B_1$ be the distance-minimizing pair. It is clear that only one of them could have been a "moved" point (moved by the transformation). Then, we again can show that if (e.g.) a moved, then there must exist $b' \in B$ such that $d_H(a^{old}, b') \leq d_H(a, b)$ – a contradiction. □

We now finish the proof of Theorem 3.

Proof of Theorem 3 (Part 2). Without loss of generality, we assume A is a down-set and B is an up-set. We apply $S^{\pm}_{U,V}$ until A, B do not change. By Lemma 5, we have $A = \mathcal{B}(0, m_1) \cup S_1$ and $B = \mathcal{B}(1, m_2) \cup S_2$. Then by Lemma 6, we have

$$D_H(A, B) \leq d_H(A_1, B_1) = n - m_1 - m_2 - k \leq n - m_1 - m_2$$

where $k = 0, 1$, or 2 depending on S_1 and S_2. □

3 The blow-up lemma

Definition 4. *For any measure metric space $(\mathcal{X}, \mu, d)$,*

$$\alpha(r, \lambda) = \max(1 - \mu(\Gamma^r A) : \mu(A) \geq \lambda)$$

is the concentration function.

We state the following theorem.

Theorem 7. *For the measure metric space $(\{0,1\}^n, Ber(\frac{1}{2})^{\partial n}, d_H)$, we have*

$$\alpha(r,\lambda) = 1 - F_n(F_n^{-1}(\lambda) + r),$$

where F_n is the CDF of Bino$(n, 1/2)$.

This theorem has the following corollary.

Corollary 8. *For any $A \subseteq \{0,1\}^n$ such that $\mu(A) \leq 1/2$, we have*

$$\mu(\Gamma^r A) \geq 1 - e^{-r^2/n}\frac{1}{\mu(A)}.$$

Equivalently, we have $\alpha(r,\lambda) \leq \frac{1}{\lambda}e^{-r^2/n}$.

Proof. Take m such that $\mu(A) = \mathbb{P}(\text{Bino}(n, \frac{1}{2}) \leq m)$. Let $m = \frac{n}{2} - \delta$ and $\lambda = \mu(A)$. Using the corollary to Hoeffding lemma, we have $\mu(A) \leq e^{-2\delta^2/n}$. From Harper's Theorem, we know that $1-\mu(\Gamma^r A) \leq \mathbb{P}(\text{Bino}(n,1/2) \geq n/2+r-\delta)$. When $r - \delta > 0$, we can apply Hoeffding and get $1 - \mu(\Gamma^r A) \leq e^{-2(r-\delta)^2/n}$. Since $|r-\delta|_+^2 \geq \frac{1}{2}r^2 - \delta^2$, we have

$$1 - \mu(\Gamma^r A) \leq e^{-\frac{2}{n}(\frac{1}{2}r^2 - \delta^2)} \leq e^{-\frac{r^2}{n}} e^{\frac{2\delta^2}{n}} \leq \frac{1}{\lambda} e^{-\frac{r^2}{n}}$$

. □

Theorem 9 (Blow-up lemma, Margulis '74, Ahlswede, Gacs and Korner '76)**.** *For any $(\mathcal{X}^n, \mu = \prod \mu_i, d_H)$, we have*

$$\mu(\Gamma^r A) \geq 1 - \frac{4}{\mu(A)} e^{-\frac{r^2}{4n}}.$$

Instead of proving this, we state and prove the following theorem, which implies the blow-up lemma through the bounded difference method.

Theorem 10. *(1) If $(\mathcal{X}, \mu, d_H)$ is such that every 1-Lipschitz function is (b, ν)-subgaussian, then*

$$\alpha(r,\lambda) \leq \frac{b^2}{\lambda} e^{-\frac{r^2}{4\nu}}.$$

(2) If

$$\alpha(r, \frac{1}{2}) \leq b' e^{-\frac{r^2}{2\nu}},$$

then every 1-Lipschitz function is $(b'e^{b'}, 2\nu)$-subgaussian.

Proof. (1) We take a set A and let $f(x) = d_H(x, A)$, which is 1-Lipschitz. We have

$$\mathbb{P}_\mu(f - \mathbb{E}f \leq -\mathbb{E}f) = \mathbb{P}(f = 0) = \mu(A).$$

Since f is (b, ν)-subgaussian, we have

$$\mathbb{P}_\mu(f - \mathbb{E}f \leq -\mathbb{E}f) \leq be^{-\frac{\mathbb{E}^2 f}{2\nu}}.$$

This gives us

$$\mathbb{E}f \leq \sqrt{2\nu \log \frac{b}{\mu(A)}}.$$

Also, we have

$$\begin{aligned}
\mu(\Gamma^r A) = \mathbb{P}(f \leq r) &= 1 - \mathbb{P}(f - \mathbb{E}f > r - \mathbb{E}f) \\
&\geq 1 - be^{-\frac{(r-\mathbb{E})_+^2}{2\nu}} \geq 1 - be^{-\frac{1}{4\nu}r^2 + \frac{\mathbb{E}^2 f}{2\nu}} \\
&\geq 1 - \frac{b^2}{\mu(A)} e^{-\frac{r^2}{4\nu}}
\end{aligned}$$

(2) By definition on concentration function, $\forall \mu(A) = \frac{1}{2}$, we have $\mu(\Gamma^r A) \geq 1 - b'e^{-\frac{r^2}{2\nu}}$. Now, let f be 1-Lipschitz and consider $A = \{x : f(x) \geq \text{med}(F)\}$, we have $\mathbb{P}(A) = \frac{1}{2}$. Since f is 1-Lipschitz, we have $\{f \geq \text{med}(f) + r\} \subseteq (\Gamma^r A)^c$. □

4 Gaussian concentration of measure

Theorem 11. *If f is 1-Lipschitz in l_2, then*

$$\mathbb{P}(|f(X) - \mathbb{E}f(X)| > t) \leq 2e^{-\frac{t^2}{2}}, \qquad X \sim \mathcal{N}(0, I_n)$$

This follows from the following stronger result:

Theorem 12. *For $(\mathbb{R}^n, \mathcal{N}(0, I_n), l_2)$, we have*

$$\alpha_n(r, \lambda) = \Phi(\Phi^{-1}(\lambda) + r),$$

where Φ is the CDF of $N(0, 1)$.

Proof. For $X^n \sim \mathcal{N}(0, I_n)$, take $N \gg n$. Let $Y \sim \text{Unif}(S^{N-1})$. Then, we have $\sqrt{N}(Y_1, \ldots, Y_n) \xrightarrow{d} X^n$ as $N \to \infty$. The set with minimal blow-up (and consequently minimal surface area) on the sphere is a spherical cap (Levy's

theorem). When projected to $\mathbb{R}^n$ the cap becomes a half-space $\{X_1 \leq a\}$, thus the set that minimizes

$$\min_S \{\mathbb{P}[d_{\ell_2}(X^n, S) \leq r] : \mathbb{P}[X^n \in S] \geq \lambda\}$$

is given by $\{X_1 \leq \Phi^{-1}(\lambda)\}$. $\square$