18.409: An Algorithmist’s Toolkit
Lecture 15

Jonathan Kelner
Today

- Finish approximating the volume of a convex body
- Start concentration of measure and geometric probability theory
APPROXIMATING THE VOLUME OF A CONVEX BODY
Sketch of Our Approach

- **Recall from last time:** can’t approximate volume deterministically, but will present a randomized algorithm
- **Main technical tool:** a method for sampling from a convex body

1) Change coordinates s.t. $K$ is well-rounded, $B \subseteq K \subseteq n \cdot B$
2) Let $\rho = 1+1/n$, and let $K_i = K \cap (\rho^i \cdot B)$. Compute
   \[
   \gamma_i = \frac{\text{Vol}(K_{i-1})}{\text{Vol}(K_i)}
   \]
3) Return $\text{Vol}(B) \cdot \prod \gamma_i$

- Last part works because $K_0 = K \cap B = B$
- First part isn’t too hard using ellipsoid algorithm (might talk about this later)
- Just need to do second one
  - Not hard if we can sample
The Grid Walk

- Last time, introduced *grid walk*
- Put a grid on $\mathbb{R}^n$ with width $\delta$
- Gives us a graph $H$:
  - Verts are points in the grid, i.e., points in $\delta \cdot \mathbb{Z}^n$
  - Edges connect $p$ to $p + \delta e_i$ for each $i$
- Let $G = H \cap K$
  - i.e., keep verts in $K$ and edges between them
- Given a starting point, can simulate a walk on this graph using a membership oracle
- If $\delta$ is small enough, a random vertex of $G$ is pretty much a random point in $K$
- Graph has exponentially many vertices, so we need to show this mixes really quickly, but convex bodies are nice
  - Cube works
- Main goal will be to show this walk mixes in poly time
- Last time, noted a couple of problems we have to deal with
Minor Problems From Last Time

- **First problem:** We’re only sampling lattice points. Never get points with coords that aren’t multiples of $\delta$.
  - **Fix:** If we get a point not in $K$, start our whole random walk over

- **Second problem:** Graph might be bipartite or almost bipartite
  - **Fix:** Take a lazy random walk (equivalently, add self loops)

- **Third problem:** The graph we’re walking on doesn’t have constant degree, so stationary distribution isn’t uniform
  - Most vertices have degree $2n$, but vertices near the boundary have lower degree
  - **Fix:** add self-loops to equalize degrees
A Worse Problem

• Our graph G might not be connected!
  ◦ So clearly our walk won’t mix quickly....
  ◦ You might say “choose a finer grid”, but this doesn’t help
    • The problem depends on the angle, not the size
  ◦ Also, some points are impossible

• Intuitively, need to get rid of “sharp corners”
  ◦ It’s a plausible guess that somehow well-roundedness rules this out, but it doesn’t

• Any guesses at a fix?
Smoothing out $K$

- Could try throwing in every grid point whose cube intersects $K$
  - But not obvious how to use a membership oracle to test this
- Even better would be to run the walk on $K(\alpha)$, defined as follows:
  - For any $\alpha$, let $K(\alpha) = \{x \in \mathbb{R}^n \mid \text{dist}(x,K) \leq \alpha\}$
    - Equals $K \oplus \alpha \cdot B_{2^n}$
  - $K(\alpha)$ has no sharp corners though, as long as our grid is small enough relative to $\alpha$
- Still have tricky question of oracles and such
  - Can be resolved, but better to avoid them
  - This body provides the right intuition for what we’re actually going to do, though
- Instead, let $K’=(1+\alpha)K$, and walk on $K’$
- See blackboard for why this helps
  - Need to be a little careful about what we ask for…
Now we have a reasonable walk that converges to the distribution we’re trying to sample from.

Now just need to show that it mixes.

For now, let’s try to get an intuition.

What does a body look like on which the grid walk mixes slowly?

- See blackboard.

We would like to bound the conductance of the graph G. What does this mean geometrically?

- See blackboard and next slide.
What We Need to Show

- We need to show that the conductance of our graph is not too small
- Let:
  - $S \subseteq G$ be a set of vertices with $|S| < n/2$
  - $E(S)$ be the edges leaving $S$
- Need to show $|E(S)| / |S|$ can’t be too small
- Further let:
  - $Q = Q(S)$ be the union of the cubes of the grid around the vertices of $S$
  - $\partial Q = \partial Q(S)$ be the $(n-1)$-dimensional boundary of $Q$
- Suppose (for now) that none of the verts in $S$ hit boundary of body
- $|S|$ is proportional to $\text{vol}_n(Q)$
- $|E(S)|$ is proportional to $\text{vol}_{n-1}(\partial Q)$
- So bounding the following two things is equivalent:

\[
\frac{|E(S)|}{|S|} \quad \text{Isoperimetric inequality!!!}\]
More Precisely What We Need

- So clearly there’s a strong link between the geometry and the conductance
- We’ll need to do two things:
  1) Formulate a version of the isoperimetric inequality that deals with things hitting the boundary
  2) Show that this geometric theorem properly captures the graph’s properties
     - A little subtler here than on last slide since we have to deal with surface areas near the boundary, corners, etc.
- We’ll do 1) first
- 2) isn’t too hard, as we’ve already talked about all the complexities that arise
Relative Isoperimetric Inequality

- To incorporate the boundary, we will need the following theorem:

- **Theorem:** Let $K \subseteq \mathbb{R}^n$ be a convex body with diameter $d$. Let $S$ be a $(n-1)$-dimensional surface that cuts $K$ into two pieces, $A$ and $B$. Then

  $$\min\{\text{Vol}_n(A), \text{Vol}_n(B)\} \leq d \cdot \text{Vol}_{n-1}(S)$$

- If $A$ doesn’t hit the boundary of $K$ and $K$ is pretty round, this is basically the standard isoperimetric inequality

- See blackboard for some pictures

- Could formulate the same question for any set $X$ (in place of the convex body $K$) and ask what $\phi$ makes the following true for any partition of it:

  $$\min\{\text{Vol}_n(A), \text{Vol}_n(B)\} \leq \phi \cdot \text{Vol}_{n-1}(S)$$

- Up to some constants and normalizations (depending whom you ask), $\phi$ is called the *isoperimetric constant* or *Cheeger constant* of $X$.
  - Yup, same Cheeger. See next slide.
Previous slides show that we can relate mixing time of walk to geometric isoperimetry
  ◦ This is why we used the word “isoperimetric constant” in the graph part of the course
  ◦ Very deep connection

Can define a “diffusion” process on many types of sets (e.g., manifolds, graphs, convex bodies)

Can then define its eigenvalues and mixing time

Cheeger’s inequality in its original form said that, for manifolds, the isoperimetric constant was related to the mixing time
  ◦ With the same basic formula

This is where the graph one came from
  ◦ In graphs, vertices are “volume” and edges “(n-1)-dim volume”

If you ignore real analysis issues, you can essentially prove the manifold one by doing an appropriately defined grid walk and sending the grid size to zero.

General tool and deep theory: if you want to prove something about mixing, there’s a geometric isoperimetric inequality somewhere
**Theorem:** Let $K \subseteq \mathbb{R}^n$ be a convex body with diameter $d$. Let $S$ be a $(n-1)$-dimensional surface that cuts $K$ into two pieces, $A$ and $B$. Then
\[ \min\{\text{Vol}_n(A), \text{Vol}_n(B)\} \leq d \cdot \text{Vol}_{n-1}(S) \]

- $(n-1)$-volume is annoying, so we’ll expand $S$ a bit and show:
- **Theorem:** Let $K \subseteq \mathbb{R}^n$ be a convex body with diameter $d$. Decompose $K$ into $K_1 \cup K_2 \cup B$, where $\text{dist}(K_1,K_2) \geq t$. Then
\[ \min\{\text{Vol}_n(K_1), \text{Vol}_n(K_2)\} \leq \frac{d}{t} \cdot \text{Vol}_n(B) \]
- Get other theorem by sending $t$ to $0$
- **Proof:** See blackboard
The Rest of the Details

- Need to check that we can arrange that our body K is polynomially well-rounded (so that its diameter isn’t too big)
  - This is on the pset. Rough idea:
  - Given a separation oracle, we can optimize a linear function over C (e.g. using ellipsoid alg.)
  - If C is far from isotropic, can find a point far from the origin and use this to construct a better John ellipse.

- Need to show that isoperimetry of our graph is properly related to isoperimetry of the body near the boundary

- This is where we use the rounding of the corners we talked about

- A couple of choices on how to proceed

- See blackboard

- So sampling from K(α), but need samples from K. What do we do?

- Need to show that we don’t reject too many samples. How?

- All together, get algorithm for sampling, and thus for volume.
CONCENTRATION OF MEASURE AND GEOMETRIC PROBABILITY THEORY
Convex Bodies and Probabilities

• Here, we’ll see how to think of convex body theory in terms of probability theory
  ◦ We sort of did this last time by relating isoperimetry to mixing time
  ◦ Here it will be more direct

• We’ll then be able to prove interesting theorems about probability using convex bodies, and vice versa

• **The main point:** Our theorems about where volume is concentrated in convex bodies have very strong probabilistic consequences
Our First Concentration of Measure Theorem

- Already have one from earlier in the semester (slightly rephrased here)
- **The Chernoff (-Hoeffding-Azuma-Bernstein-...) bound:**
- Let
  - $x \in \{\pm 1\}^n$ be independent random variables with $p[x_i=1]=0.5$.
  - $a_1,...,a_n$ satisfy $\sum a_i^2 = 1$.
    - Some could be negative

Then

$$\Pr \left[ \left| \sum_{i=1}^{n} a_i x_i \right| > t \right] \leq 2e^{-t^2/2}$$

- Let’s change it so the $x_i$ are in $[-1/2,1/2]^n$.
- Bound still true with slightly different constants
- What does this say?
Our First Concentration of Measure Theorem

\[
Pr \left[ \left| \sum_{i=1}^{n} a_i x_i \right| > t \right] \leq 2e^{-6t^2} \quad x \in [-1/2, 1/2]^n \quad \sum_i a_i^2 = 1
\]

- \[ \sum a_i x_i = a \cdot x = \text{distance of } x \text{ from hyperplane } H_a = \{ x \mid a \cdot x = 0 \} \]
  - Since \[ \sum a_i^2 = 1 \]
- So theorem says "most of the mass of the cube lies near any hyperplane through the origin"
  - See blackboard for picture
- **Can also phrase as:** Let \( S = \{ \text{all points within distance } t \text{ of } H_a \} \)
  \[
  \frac{\text{Vol}(S)}{\text{Vol}([-1/2, 1/2]^n)} \geq 1 - 2e^{-6t^2}
  \]
- Feels like our statement that most of the volume of the n-sphere lies near the equator
- Also feels a bit like the relative isoperimetric inequality
  - Although dealing with big sets, not small ones, so different quantitatively
- Here we’ll show how all these are related
Let’s ask the isoperimetric question on the sphere $S=S^{n-1} \subseteq \mathbb{R}^n$

For distance, will use distance in $\mathbb{R}^n$ (not on sphere)

Volumes w.r.t. unique rotationally invariant measure on sphere normalized to have total = 1.

For any $A \subseteq S$, let $A_\varepsilon$ be the set of points within distance $\varepsilon$ of $A$

**Question:** Out of all sets with $\text{vol}(A)=q$, which ones gives min possible $\text{vol}(A_\varepsilon)$?

**The answer:** Spherical caps

**Definition:** Let the spherical cap $C(r,v)$ be given by

$$C(r,v) = \{ x \in S^{n-1} | d(x,v) \leq r \}$$

Just a ball in the metric on the sphere

- So like Euclidean isoperimetric inequality

Need a slightly more complicated formulation because volume ratio depends on radius
Spherical Caps

Figure 26. The Euclidean metric on the sphere. A spherical cap (right) is a ball for this metric.

- Sometimes easier to talk about “cap at height $t$”:
  $$c_t = c(t, v) = \{x \in S^{n-1} \mid x \cdot v \geq t\}$$
- $c_0 = \text{hemisphere}$
- $\text{vol}(c_t) \approx e^{-nt^2}/2$
What Does This Imply?

- Will prove approximate version of this later. Let’s first see why it’s interesting.
  - Actual proof is sort of tricky, but can get the asymptotics right very easily with what we already know
- **Corollary 1:** For any $A$ with $\text{vol}(A) = 1/2$,
  $$\text{vol}(A_\varepsilon) \geq 1 - e^{-n\varepsilon^2/2}$$
  - Why?
- So almost all of the sphere is near *any* set of volume $1/2$, even though there may be many points far away from it.

![Figure 27. An $\varepsilon$-neighbourhood of a hemisphere.](image)
**Definition:** A function $f:S^{n-1} \rightarrow \mathbb{R}$ is $1$-Lipschitz if for any $a, b \in S^{n-1}$,

$$|f(a) - f(b)| \leq ||a - b||_2$$

- Let $M$ be the median of a 1-Lipschitz function $f$, so $\text{vol}({f \leq M}) = \text{vol}({f \geq M}) \geq 1/2$
- How many points have $f(x) \leq M + \varepsilon$?
  - See blackboard

**Theorem:** $\text{vol}({x : |f(x) - M| > \varepsilon}) \leq 2e^{-n\varepsilon^2 / 2}$
- So even though $f$ can vary by up to 2 on the sphere, it is almost equal to $M$ almost everywhere
- So “1-Lipschitz functions are almost constant”
  - A little weird...
- This is what we’ll call concentration of measure
- Note we just need the right asymptotics for this, not the exact isoperimetric inequality.
We have two types of properties:

**Isoperimetric inequalities:** Every set $A$ with $\text{vol}(A) \geq 1/2$ has $\text{vol}(A_\epsilon \geq 1 - e^{-c\epsilon^2})$

**Concentration theorems:** For every 1-Lipschitz function $f$, $\text{vol} \{ x : |f(x) - M| > t \} \leq 2e^{-ct^2/2}$

We showed that isoperimetric inequalities imply concentration theorems

Can go the other way as well: apply the concentration theorem to the function $f(x) = \text{dist}(x, A)$ [exercise]

Works whenever we have a notion of volume and distance
Proof of Isoperimetric Inequality for the Sphere

- See blackboard.