Approximating the volume of a convex body

Before we can talk about algorithms on convex bodies, we need to be more precise about what it means to be “given a convex body $K$”. A couple possible versions:

- **Membership oracle**: give it a point $p$, it tells you if $p \in K$.
- **Separation oracle**: give it a point $p$ it either:
  - Says $p \in K$, or
  - Gives a hyperplane that separates $p$ from $K$.

How would we construct a separation oracle for:
- A ball?
- A cube?
- A polytope given by inequalities?

We’re also going to assume we know a ball inside our body of radius 1 and a ball outside of it of radius $2^{\text{poly}(n)}$.

- WLOG both around 0
- How strong is this assumption?
Computing the Volume

- Our goal will be to approximate volume—exact computation is \#P-hard
- Let’s start with the 2-d case
- Suppose I give you a membership oracle for \( K \).
- How can you approximately compute its volume?
  - Really only one thing you can do: pick points, see if they’re in it
- Does this give a poly time algorithm?
  - See blackboard
- Okay, what if I tell you the body is “well-rounded” (i.e., \( B_2^n \subseteq K \subseteq n \cdot B_2^n \))
  - See blackboard
- Note: could even do this deterministically

Why Does the 2-d Algorithm Work?

- More formal statement of the algorithm:
  - Pick a bunch of points \( p_1, \ldots, p_m \)
  - Use membership oracle to see if each is in \( K \)
  - Let \( Q = \{ \text{set of } p_i \text{ that are in } K \} \)
  - Compute the area of the convex hull of \( Q \)
  - Return this as the approximate area of \( K \)
- Why does this work?
  - Theorem: For any \( \varepsilon > 0 \), there exists a set \( P = \{ p_1, \ldots, p_m \} \) s.t.
    - \( m = \text{poly}(1/\varepsilon) \)
    - For any well-rounded 2-d convex body \( K \), \( \text{area}(\text{conv}(P) \cap K) \geq \text{area}(K)/(1+\varepsilon) \)
- Could take a grid, for example

Higher Dimensions

- Clearly a grid won’t work in higher dimensions
  - Exponentially many points
- So we need to be a little trickier
- We also can’t compute volume of convex hull of a bunch of points so easily...
  - Suppose for now that we can
- And we need to get a well-rounded body somehow
  - Ignore this for now too
- We’ll need to construct a very carefully chosen set of points s.t. we know their convex hull has about the same volume as \( K \)

Higher Dimensions (cont.)

- New algorithm: same as old one, but pick the points depending on \( K \)
- Need new version of the theorem:
  - Theorem: For any well-rounded convex body \( K \) and any \( \varepsilon > 0 \), there exists a set \( P = \{ p_1, \ldots, p_m \} \) s.t.
    - \( m = \text{poly}(1/\varepsilon) \)
    - \( \text{vol}(\text{conv}(P) \cap K) \geq \text{vol}(K)/(1+\varepsilon) \)
- Only problem: theorem is very false!
- In fact, we’ll prove:
  - Theorem: There is no deterministic poly time algorithm that, given a membership oracle for \( K \), computes \( \text{vol}(K) \) within a polynomial factor
  - Could also prove with separation oracle (maybe will be a problem set question)
Impossibility of Volume Computation

- **Theorem:** There is no deterministic poly time algorithm that, given a membership oracle for \( K \), computes \( \text{vol}(K) \) within a better than exponential factor.
- How might we prove this?
  - Impossibility results are hard. What if we choose really good points, or use some smart theorem about convex bodies.
- Let’s scale s.t. our body is contained in the unit ball (and contains \( 1/n \) times the unit ball).
  - To get version of theorem when body is promised to be well-rounded, replace \( K \) with \( \text{conv}(K \cup \{ \pm e_1, \ldots, \pm e_n \}) \).
- Suppose we’ve picked \( p_1, \ldots, p_m \) by any method, and that they all lie in \( K \).
  - Body still could be either \( C = \text{conv}(p_i) \) or \( B_2^n \).
- **Theorem:**
  \[
  \frac{\text{vol}(C)}{\text{vol}(B_2^n)} \leq \frac{m}{2^n}.
  \]
- See blackboard.

A Few Comments

- **Self-Indulgent Philosophizing:** If we can do this (and we can, as I’m sure you’ve guessed), it’s pretty impressive. This is one of the biggest known gaps between deterministic and randomized computation.
  - Almost certainly, BPP=\( P \) (I can defend this if you disagree).
  - So, somehow, any body whose separation oracle can be created in poly time, can have its volume approximated deterministically, even though this is impossible for general bodies...
- **Note:** All that follows will aim for a poly time algorithm. Won’t make any effort to get a good polynomial.
- Roughly following original Dyer, Frieze, Kannan paper on approximating volume.
- Much lower polynomials now exist.
- Current best computes volume in \( O(n^4) \).

Sketch of Our Approach

- **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.
  - See blackboard.

Game Over?

- So does this mean we can’t hope to get a polynomial approximation of volume?
- Used the word “deterministic” in theorem, so I guess we have to use randomness.
- Maybe a Monte Carlo algorithm?
  - Pick random points in the sphere, see how many land in \( K \).
  - Won’t work since can make bodies where almost no points will end up in \( K \).
- Does this situation sound familiar?
  - Looks a lot like problems we had approximating the permanent.
- How did we get around them? (see blackboard).
Sampling from Convex Bodies

- So need a way of sampling from a convex body $K$
- You may have guessed that we’ll use a random walk
- Before this, let’s do a warmup. How would you sample from:
  - A cube?
  - A ball?
  - An ellipse?
  - A spherical cone?
  - A polytope (given by polynomially many constraints)?
- What might a good random walk be?
  - A few walks have been used for this
  - See blackboard for some sketches
- We’ll study the “grid walk”

The 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 $p \pm \delta e_i$ 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
  - How?
- 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
- What happens if $K$ is a cube?
  - So our main goal will be to show this walk mixes in poly time
- First, have a couple of problems to deal with

Problems with our Approach

- **First problem:** We’re only sampling lattice points. Never get points with coords that aren’t multiples of $\delta$.
- How could we fix this?
- Let $p$ be a random vertex of $G$
- Just add a random vector $v$ from the cube $[-\delta/2, \delta/2]^n$
- Does this work?
- What if $v+p$ isn’t in $K$?
- Could keep trying $v$’s until we get a point in $K$
  - What’s wrong with this?
  - See blackboard
- **(Somewhat unsatisfying) fix:** If we get a point not in $K$, start our whole random walk over

Some Not-Too-Bad Problems

- Graph might be bipartite
  - Or not actually bipartite, but might look bipartite until you take a lot of steps
  - So just take a lazy random walk
- The graph we’re walking on doesn’t have constant degree
  - Most vertices have degree $2n$
  - Vertices near the boundary have lower degree
- So the stationary distribution isn’t uniform
- How do we fix this?
- Just add self-loops to equalize degrees
- Can restate cleanly as:
  - Pick a random vector $v$ from $\{\pm e_i\}$
  - If $p+v \in K$, go to it
  - Otherwise, stay where you are
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

- 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?

Mixing

- 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

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 $K(\alpha) = \{ x \in \mathbb{R}^n | \text{dist}(x, K) \leq \alpha \}$
    - 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...

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:

\[
\text{Conductance for } S = \frac{|E(S)|}{|S|} = \frac{\text{Vol}_{n-1}(\partial Q(S))}{\text{Vol}(Q(S))}
\]
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.

Isoperimetry for Graphs and Bodies

- 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

Proof of Relative Isoperimetric Inequality

- **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(\alpha), 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.