18.409 An Algorithmist's Toolkit November 3^{rd}, 2009

Lecture 15

Lecturer: Jonathan Kelner *Scribe: Justin Thaler*

1 Outline

In this lecture we will design a method to randomly sample from a convex body, and this method will be a subroutine in approximately computing volume

2 Reminder from Last Time

In the last lecture, we showed that no deterministic algorithm that queries a membership oracle only a polynomial number of times, can approximate the volume of a convex body to within a factor of $\frac{m}{2^n}$ where m is the number of membership queries. However, a randomized algorithm can approximately compute the volume of a convex body and the procedure will be similar to the Jerrum and Sinclair method for approximating the permanent. The algorithm will construct a series of convex bodies, and the ratio of successive convex bodies in the series can be well-approximated. These approximations will be used to approximate the volume of the original convex body K, even if K has exponentially small volume compared to the bounding sphere that is guaranteed to contain K.

Last time, we presented the following sketch of how our randomized algorithm will work.

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

Note that Step 3 works because $K_0 = K \cap B = B$. Also note that the first part isn't too hard using the ellipsoid algorithm, which we mention briefly near the end of these notes.

3 Grid Walk

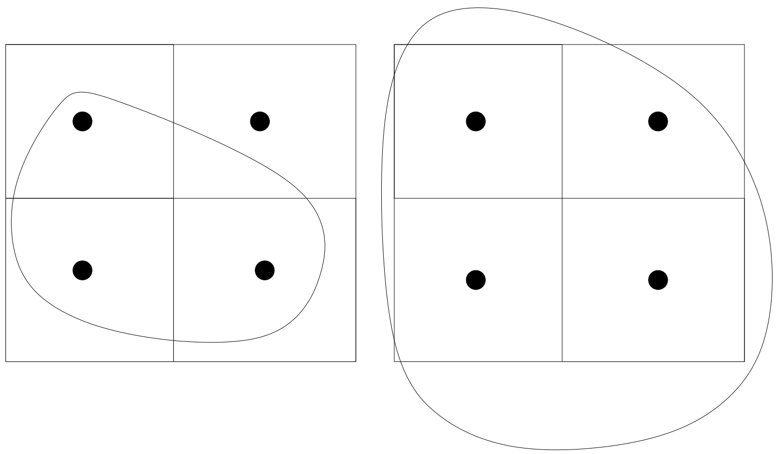
To approximately compute volume in Step 2 of our sketch above, we need a method to sample randomly from a convex body K. To do this, we will use a random walk and we will need to bound the mixing time of the random walk. There are a number of random walks that can be used to sample randomly from a convex body, and the most basic is the Grid Walk.

Define a grid graph H, such that nodes are points in δZ^n. The define the graph $G = H \cap K$ as the subgraph of nodes in H that are also contained in the convex body K. For sufficiently small δ, a random vertex in G is roughly a random point in K

The graph G can contain an exponential number of vertices. Consider the convex body K, the hyper-cube $[-1, 1]^n$. We must show that even if the number of vertices is exponential, that the random walk mixes in polynomial time. Returning to the example of the hyper-cube, the hyper-cube is the graph product of n line graphs. Each line graph has a mixing time c^2 if there are c nodes. Then a random walk on the grid graph is a choice of the direction to walk in, and a random step on the corresponding line graph. The walk mixes in $O(nc^2 logn)$ steps, because in expectation $O(nlogn)$ steps are needed to ensure that a step is taken in all directions (this is an instance of the coupon-collector problem), and when $O(c^2)$ steps have been taken in each direction, a random walk on the hyper-cube is mixed.

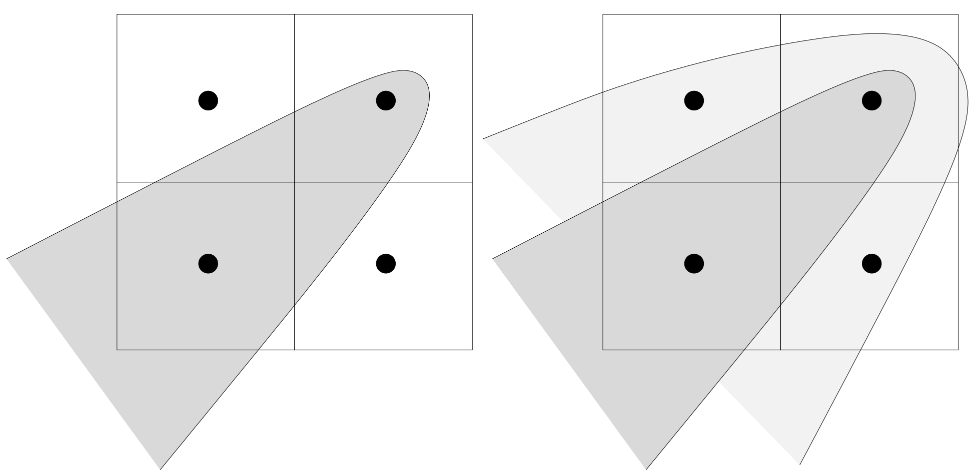
4 Problems with our Approach

We only sample grid points. A first-cut fix is to generate a random vertex p in G, and then add a random vector k in the cube $[-\frac{\delta}{2}, \frac{\delta}{2}]$ to the point p. However $p+k$ is not necessarily in K. We could generate another k in the cube, and try again - but if the nodes in G are generated uniformly at random, then a cube containing points in K and points not in K will generate the points in K disproportionately often compared to points in K that are in a cube that only contains points in K. This problem can be avoided by restarting the random sample procedure when a point $p+k$ is generated that is not in K. A more serious problem is that not all points in K can be generated.



Also, the graph G is bipartite because the full grid H is bipartite, and G is a subgraph of H. Then a random walk will be periodic. Also, not all nodes in G have the same degree and the limiting distribution is not necessarily uniform on the nodes in G. These problems can be fixed by adding self-loops at each node, and adding extra self-loops at any node that is not connected to $2n$ nodes in G. The degrees in the graph can be made equal, and the limiting distribution will be uniform on the nodes.

The graph G need not even be connected.



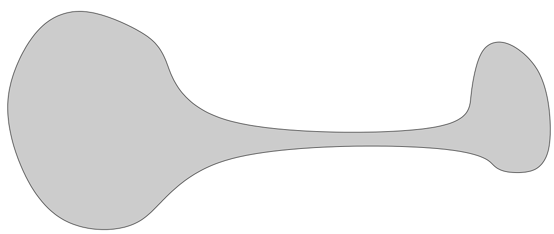
Intuitively, this problem arises when K contains sharp boundaries and these problems can be removed by rounding out K. A possible approach is to set $K(\alpha) = K \oplus \alpha B_2^n$, or to set $K' = (1+\alpha)K$. The approaches can both be made to work, but consider K'. By assumption $B_2^n \subset K$ and $\alpha B_2^n + K \subset K'$. Choose $\alpha = \delta\sqrt{n}$ equal to the diameter of a cube. Then all cubes contained in K have a neighbor in all $2n$ directions that is contained in K', and for all points p in K there is a cube in K' that contains p.

Run a random walk on K', and at any cube generate a random point $p + k$. If this point is not in K, then start the random walk over. Provided that there are not too many cubes near the boundary or entirely in K', in expectation we will not need to run the random walk many times before obtaining a random point in K. Then this defines a random walk on the graph $G' = H \cap K'$, and again self-loops can be added to each vertex to ensure that G' is regular and that the random walk is aperiodic.

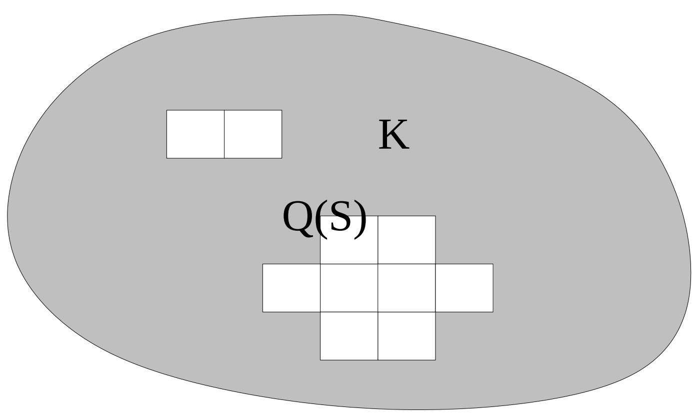
Consider an arbitrary cube C that contains points in K, and points not in K. A point in $K \cap C$ is generated with probability equal to the fraction of volume in C also in K, once the node corresponding to the cube C is reached. Because all points in K can be generated, then this random walk generates points in K uniformly at random.

5 Mixing

The above random walk will generate a random sample from any body K. Convexity is not needed to ensure that this random walk produces a random sample, but is needed to ensure that the random walk mixes quickly. Consider the body K given in the figure below. Intuitively, this walk mixes slowly for the same reasons that a random walk on a graph containing two cliques connected by a long path does.



To bound the mixing time for a random walk on a convex body K, we need to bound the isoperimetric number or conductance of G. Then for any set S of nodes in G such that $|S| \leq \frac{|V(G)|}{2}$, the isoperimetric number for S is $\frac{|E(S)|}{|S|}$. Each cube contains the same volume, and the size of S is proportional to the volume of $Q(S)$ - $Vol_n(Q(S))$, the space enclosed by the cubes corresponding to nodes in S. Similarly, each edge leaving S corresponds to a face on the surface of $Q(S)$, and each face has the same surface area. Then the number of edges leaving S is proportional to $Vol_{n-1}(dQ(S))$. If the space $Q(S)$ does not intersect the boundary of K, then this is exactly the isoperimeteric number of the graph. To incorporate the boundary, we need a Relative Isoperimetric Inequality.



Theorem Let $K \subset R^n$ be a convex body with diameter d. Let S be an $n-1$ dimensional surface that cuts K into two pieces A and B. Then

$$min\{Vol_n(A), Vol_n(B)\} \leq dVol_{n-1}(S)$$

Again, if A does not intersect the boundary (and is round enough) then this is approximately the standard isoperimeteric inequality. Also, we can define the isoperimeteric constant (or Cheeger constant) for any body X (not necessarily convex) as the minimum ϕ such that

$$min\{Vol_n(A), Vol_n(B)\} \leq \phi Vol_{n-1}(S)$$

Isoperimetric inequalities arise naturally in bounding the mixing time of any "diffusion" process.

6 Approximate Proof

The $n-1$-dimensional volume is more subtle to work with, and this theorem can be proven by proving a related theorem.

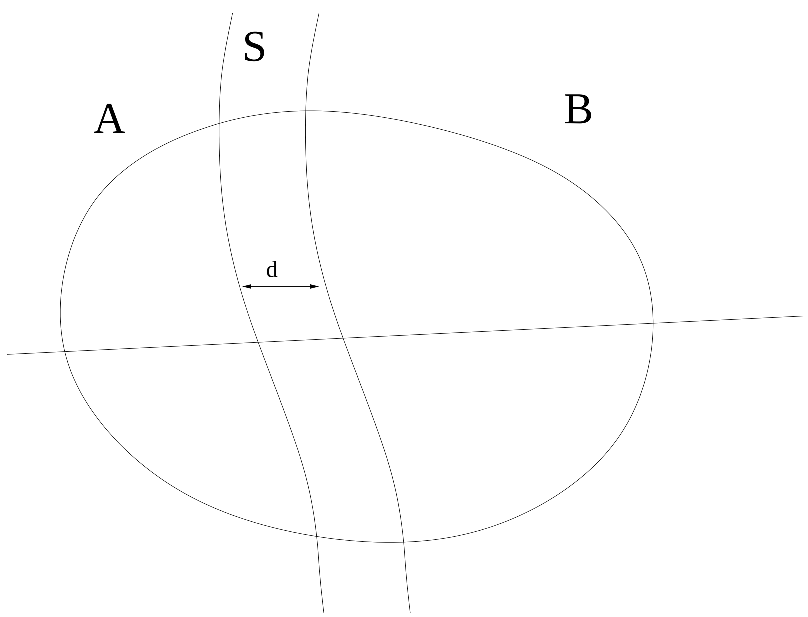
Theorem Let $K \subset R^n$ be a convex body with diameter d. Decompose K into $A \cup B \cup S$, where $dist(A, B) \geq t$. Then

$$min\{Vol_n(A), Vol_n(B)\} \leq \frac{d}{t} Vol_n(S)$$

The original theorem is proven by decreasing t to zero. Let E be the smallest volume ellipse containing K. Then there are two cases to consider.

Case 1: The ellipse E is needle-like, which we define to mean all but at most 1 axis of E is of radius $\leq \epsilon t$ for a small enough ϵ. In this case, the theorem is true by inspection.

Case 2: The ellipse E is not needle-like. Then we can apply a symmetrization procedure until the ellipse is needle-like. Suppose that there exists a counter-example to our theorem, then by the Ham Sandwhich Theorem there exists a hyperplane that simultaneously cuts A into A_1, A_2 and B into B_1, B_2 such that A_1 and A_2 have equal volume, and so do B_1 and B_2. Then the hyperplane cuts S into S_1, S_2 and one of the convex bodies $A_1 \cup B_1 \cup S_1$ OR $A_2 \cup B_2 \cup S_2$ is a counter-example that is closer to needle-like.



Iterating this procedure, we can eventually reduce all but at most one dimension to $\leq \epsilon t$, and this produces a contradiction because the theorem is true when the bounding ellipse is needle-like.

7 The Rest of the Details

1. We need to make sure we can arrange for our convex body K to be polynomially well-rounded to make sure the diameter isn't too big. The rough idea is that if K is far from isotopic (i.e. not well-rounded) we can find a point far from the origin using the ellipsoid algorithm and use this to construct a better John Ellipse; see the problem set for details.
2. We need to show that isoperimetry of our graph is properly related to isoperimetry of the body near the boundary. This is where we use rounding of the corners of K.
3. Finally, we need to show that we don't reject too many samples.
4. Once we've done all of the above, we get an algorithm for sampling from any convex body K, and can use this to estimate the volume as per our sketch at the beginning of these notes.

8 Concentration of Measure and Geometric Probability Theory

8.1 The Chernoff (Hoeffding-Azuma-Bernstein-...) Bound

The question here is how to think of a convex body relating in some way to probability theory. We've been doing a lot of things with convex bodies and probabiliies and there should be a lot of overlap. We sort of did this last time with isoperimetry but now we will be much more concrete. We'll think of points in the convex body as being points of a probability distribution. We'll have interesting, very strong theorems that go both ways in implications, that will appear unlikely. The main point is that we keep coming up with the phenomenon that volume in convex bodies is counterintuitively distributed, more or less pervasively for high-dimensional spaces. We've seen this over and over but we'll make it more concrete now. We already have the first concentration of measure theorem from earlier in the semester: the Chernoff bound.

Theorem 1 *Let $x \in \{\pm 1\}^n$ be independent random variables with $p[x_i = 1] = .5$, and $a_1, \ldots, a_n$ satisfying $\sum a_i^2 = 1$ (some can be negative). Then*

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

I assert this is the same bound we already did. Now, let's change it so that the x_i are anywhere in $[-1/2, 1/2]$. The bound is still true up to some constants. Let's see what this means geometrically.

Claim 2 $\sum a_i x_i = a \cdot x =$ *distance of x from hyperplane $H_a = \{x | a \cdot x = 0\}$.*

Pictorially, that means I can take the unit cube in R^n, pick any hyperplane at all, and we cut the cube with it, and that gives us some intersection with the interior of the cube. What it says is that no matter how we choose this hyperplane, almost all of the cube is pretty close to this hyperplane. I claim this is our first concentration of measure theorem.

Another way we can phrase this, so you get a hint of where I'm going with this, is to say that

$$\frac{\text{Vol}(S)}{\text{Vol}([-1/2, 1/2]^n)} \geq 1 - 2e^{-6t^2}$$

Here, we define S to be the set of all points within distance t of H_a.

So what is a little neighborhood of S? I've shown that it's "pretty big" as a function of the volumes involved. This is not exactly the isoperimetric inequality because we're looking at big sets, not small sets, and the set, not its complement. It's a different parameter regime but the same kind of question. Somehow we have three phenomena in this course that all come out to be the same thing: isoperimetric inequalities, Chernoff bounds, and this phenomenon of volume in convex bodies.

18.409 Topics in Theoretical Computer Science: An Algorithmist's Toolkit
Fall 2009