CS294 Markov Chain Monte Carlo: Foundations & Applications	**Fall 2009**

Lecture 16: October 29

Lecturer: Alistair Sinclair *Scribes:*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

16.1 Volume of a Convex Body

We will discuss algorithms for estimating the volume of convex bodies. We'll try not to get bogged down in technical details, but there are some things we'll need to sift through. The problem definition is the following.

Input: A convex body $K \subseteq R^n$.

Goal: Compute the volume of K, $\mathbf{vol}_n(K)$.

We assume throughout that the body K is specified by a *membership oracle*, i.e., a black box which, given a point $x \in R^n$, replies whether or not $x \in K$.

The problem is known to be $\sharp$P-hard [DF88, K93]. Moreover, it has been shown that there is no deterministic algorithm which approximates $\mathbf{vol}_n(K)$ within any "reasonable" factor [BF87]: for any deterministic sequence of oracle queries of length polynomial in n, there exist two different convex bodies consistent with the answers to the queries which have exponentially large volume differences. Nevertheless, as we shall see next, randomized algorithms allow us to approximate the volume of a convex body to arbitrary accuracy in polynomial time.

Claim 16.1 *There exists a f.p.r.a.s. for the volume problem in the oracle model.*

The first such algorithm was devised by Dyer, Frieze and Kannan [DFK91] and its running time was about $\tilde{O}(n^{26})$. (The notation $\tilde{O}$ indicates that factors of polylog(n) are omitted, in addition to constant factors.) A long sequence of subsequent improvements (mainly in the mixing time analysis, often involving major conceptual and technical advances) by various combinations of these three authors together with Lovász, Simonovits [LS93] and others, led to the current state of the art algorithm due to Lovász and Vempala [LV03] with a running time of $\tilde{O}(n^4)$.

The basic idea of the algorithm is to select an increasing sequence of concentric balls $B_0 \subseteq B_1 \subseteq \ldots \subseteq B_r$, such that $B_0 \subseteq K \subseteq B_r$, where, without loss of generality, B_0 can be assumed to be the unit ball and B_r to have radius of $O(\sqrt{n})$. (The fact that K can always be "rounded" in this way is a standard but non-trivial geometric fact that we will not prove here.)

Note that, since $\frac{\mathbf{vol} B_r}{\mathbf{vol} B_0}$ is exponential in n, a naive Monte-Carlo algorithm which samples from B_r and tests whether the sampled point belongs to K does not in general give an efficient algorithm.

Let us, however, express the volume of K as follows:

$$\mathbf{vol}(K) = \mathbf{vol}(K \cap B_r) = \frac{\mathbf{vol}(K \cap B_r)}{\mathbf{vol}(K \cap B_{r-1})} \times \ldots \times \frac{\mathbf{vol}(K \cap B_1)}{\mathbf{vol}(K \cap B_0)} \times \mathbf{vol}(K \cap B_0).$$

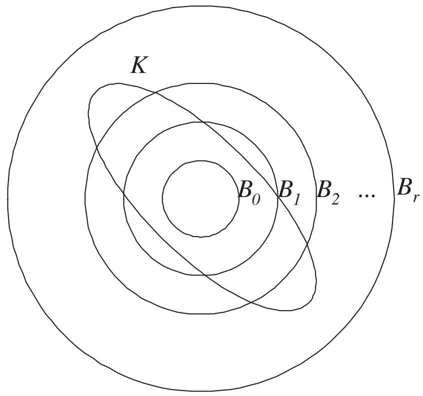

Figure 16.1: Sequence of concentric balls used in volume algorithm

Now note that $\mathbf{vol}(K \cap B_0) = \mathbf{vol}(B_0)$ is trivial, and that each ratio $\frac{\mathbf{vol}(K \cap B_i)}{\mathbf{vol}(K \cap B_{i-1})}$ can be estimated by sampling u.a.r. from $K \cap B_i$.

If we set $\frac{\mathbf{rad}(B_i)}{\mathbf{rad}(B_{i-1})} = 1 + \frac{1}{n}$ for each i, where $\mathbf{rad}(B)$ denotes the radius of B, it follows that $\frac{\mathbf{vol}(K \cap B_i)}{\mathbf{vol}(K \cap B_{i-1})} = (1 + \frac{1}{n})^n \leq$ e (a constant). Thus each ratio can be estimated accurately (say, within ratio $(1 + \frac{\epsilon}{r})$) using only $O(r^2\epsilon^{-2})$ random samples. And, since $\mathbf{rad}(B_r) = O(\sqrt{n})$, the number of balls we require is only $r = O(n \log n)$. Thus we deduce the following:

$\exists$ a f.p.r.a.s. for estimating $\mathbf{vol}(K)$ given a black box for sampling (almost) u.a.r. from $K \cap B_i$.

Note that $K \cap B_i$, being the intersection of two convex bodies, is itself convex. Thus we have a polynomial time reduction from computing the volume of a convex body to sampling u.a.r. in a convex body.

In the remainder of this lecture and next, we focus on algorithms for sampling from a convex body. This is the main content of work in this area.

16.1.1 Random Sampling

We consider two variants of a basic random-walk based sampling algorithm. Let $B(x, \delta)$ denote the ball of radius δ centered at x.

Heat-Bath Ball Walk: at point $x \in K$, move to a point in $B(x, \delta) \cap K$ u.a.r.

Metropolis Ball Walk: at point $x \in K$, pick $y \in B(x, \delta)$ u.a.r.;
if $y \in K$ then move to y, else stay at x.

Note that the heat-bath version of the ball walk is not easily implementable, since at every step it requires uniform sampling from the intersection of K and a ball around the current point. Moreover, as we shall soon see, it does not result in a uniform stationary distribution, since there is higher chance of moving away from the boundary. On the other hand, the Metropolis version is easy to implement and yields a uniform stationary distribution. Nevertheless, it turns out that it is easier to analyze the heat-bath version. What we shall do next will be to analyze the heat-bath version of the ball walk and then argue that the mixing time of the Metropolis version is similar.

16.1.2 Notation for Continuous Space Markov Chains

For a point $x \in R^n$ and a (measurable) set $A \subseteq K$, denote by $p(x, A)$ the transition probability from x to A, i.e.,

$$p(x, A) = \Pr[X_{t+1} \in A | X_t = x].$$

Then the t-step transition probability is recursively defined as follows:

$$p^t(x, A) = \int_K p^{t-1}(x, dy) p(y, A).$$

For the heat-bath ball walk we have

$$p(x, A) = \frac{\mathbf{vol}(B(x, \delta) \cap A)}{\mathbf{vol}(B(x, \delta) \cap K)}, \text{ where } A \subseteq K.$$

This walk is specified buy the infinitesimal generator

$$p(x, dy) = \begin{cases} \frac{dy}{\mathbf{vol}(B(x,\delta) \cap K)} & \text{if } y \in B(x, \delta) \cap K; \\ 0 & \text{otherwise.} \end{cases}$$

A probability distribution μ over K is *a stationary measure for P* if

$$\mu(A) = \int_K p(x, A) \mu(dx), \ \forall \text{ measurable } A \subseteq K. \tag{16.1}$$

Definition 16.2 *Define the density* $\ell(x) := \frac{\mathbf{vol}(B(x,\delta) \cap K)}{\mathbf{vol}(B(x,\delta))}$*, and set* $L := \int_K \ell(x) dx$.

Claim 16.3 *The distribution* $\mu(A) = \frac{\int_A \ell(x) dx}{L}$ *is stationary for the heat bath ball walk.*

Proof: With this definition of μ, the right-hand side of (16.1) becomes

$$\begin{aligned} \int_K p(x, A) \mu(dx) &= \int_A \int_K p(x, dy) \mu(dx) \\ &= \int_A dy \int_{B(y,\delta) \cap K} \frac{\mu(dx)}{\mathbf{vol}(B(y, \delta) \cap K)} \\ &= \frac{1}{L} \int_A dy \int_{B(y,\delta) \cap K} \frac{\ell(x) dx}{\mathbf{vol}(B(x, \delta) \cap K)} \\ &= \frac{1}{L} \int_A dy \int_{B(y,\delta) \cap K} \frac{dx}{\mathbf{vol}(B(x, \delta))} \\ &= \frac{1}{L} \int_A \ell(y) dy = \mu(A), \end{aligned}$$

which verifies the required condition (16.1). (In the second line here we have used the fact that $y \in B(x, \delta) \cap K \Leftrightarrow x \in B(y, \delta) \cap K$.) ■

Note that the density ℓ puts less weight near the boundary of K, as we had guessed earlier from the definition of the heat-bath walk. Also, at this stage we are not in a position to claim that the heat-bath ball walk always converges to μ as we have not developed a theory of Markov chains with continuous state spaces. However, the fact that it does converge to μ will emerge as a consequence of our analysis of its mixing time in the next lecture.

Exercise 16.4 *Verify that the uniform measure is stationary for the Metropolis version of the ball walk.*

16.1.3 Curvature Assumption

To simplify the analysis, we'll make the following curvature assumption, which is **not** without loss of generality:

$$\forall x \in K, \exists y \in K \text{ s.t. } B(y,1) \subseteq K \text{ and } d(x,y) \leq 1.$$

The assumption, which is illustrated in Figure 16.2, corresponds to a smoothness assumption for the boundary of K. It will eliminate a lot of technical complexity from the analysis of the ball walk while retaining most of the interesting ideas.

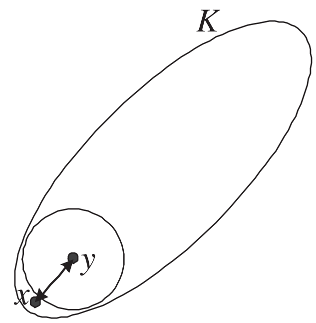

Figure 16.2: The curvature assumption

Claim 16.5 *Under the curvature assumption, if we also assume that $\delta \leq \frac{c}{\sqrt{n}}$, where c is a universal constant, then $\forall x \in K, \frac{2}{5} \leq \ell(x) \leq 1$.*

Observe that the above claim implies that we can efficiently implement the heat-bath ball walk by "rejection sampling" based on the Metropolis version. (The expected time to make one step of the heat-bath walk is at most $5/2$.) Also, given a sample produced by the heat-bath ball walk from its stationary distribution μ (with density ℓ), we can make this uniform by accepting with probability $\frac{2/5}{\ell(x)} \leq 1$. (The latter is based of course on the fact that, because $\ell(x)$ is bounded from above and below by constants, it can be estimated within very high accuracy by sampling.)

Proof: Let $x \in K$, $y \in K$ as in the curvature assumption. Then by definition of $\ell(x)$ we have

$$\ell(x) \geq \frac{\mathbf{vol}(B(x,\delta) \cap B(y,1))}{\mathbf{vol}(B(x,\delta))} \geq \frac{1}{2} - \frac{\text{volume of shaded cylinder of Figure 16.3}}{\mathbf{vol}(B(x,\delta))},$$

where we have used the fact that $\mathbf{vol}(B(x,\delta) \cap B(y,1))$ is minimized for x lying on the boundary of $B(y,1)$.

The height of the cylinder is $h = \delta^2/2$ (basic geometry: exercise). The area of the base of the cylinder is $b = \mathbf{vol}_{n-1}(B(x,\delta)))$. Using the fact that $\frac{\mathbf{vol}_{n-1}(B(x,\delta))}{\mathbf{vol}_n(B(x,\delta))} \geq \frac{c'\sqrt{n}}{\delta}$ for a constant c', we get

$$\ell(x) \geq \frac{1}{2} - \frac{\delta^2}{2} \times \frac{c'\sqrt{n}}{\delta} = \frac{1}{2} - \frac{c'\delta\sqrt{n}}{2} \geq 2/5 \text{ by taking } \delta \leq \frac{1}{5c'\sqrt{n}}.$$

∎

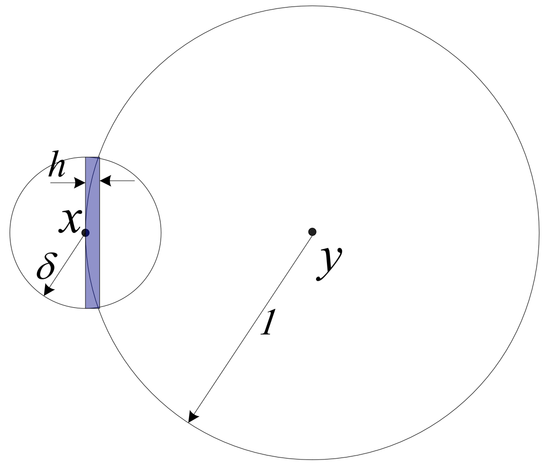

Figure 16.3: Proof of Claim 16.5

References

[1] I. Bárány and Z. Füredi. "Computing the volume is difficult," *Discrete and Computational Geometry*, 2:319–326, 1987.

[2] M.E. Dyer and A.M. Frieze. "The complexity of computing the volume of a polyhedron," *SIAM Journal of Computing*, 17:967–974, 1988.

[3] M. Dyer, A. Frieze and R. Kannan. "A random polynomial-time algorithm for approximating the volume of convex bodies," *J. Assoc. Comput. Mach.,* 38:1–17, 1991.

[4] L.G. Khachiyan. "Complexity of polytope volume computation," *New Trends in Discrete and Computational Geometry*, J. Pach editor, pages 91–101, Springer Verlag, Berlin, 1993.

[5] L. Lovász and M. Simonovits. "Random walks in a convex body and an improved volume algorithm," *Random Structures and Algorithms*, 4:359–412, 1993.

[6] L. Lovász and S. Vempala. "Simulated annealing in convex bodies and an $\tilde{O}(n^4)$ volume algorithm," *IEEE FOCS*, 2003.